

UNIVERSITY OF BIRMINGHAM

HIERARCHICALLY MODELLING MANY STARS TO IMPROVE INFERENCE WITH ASTEROSEISMOLOGY

by

ALEXANDER J. LYTTLE

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Sun, Stars and Exoplanets Group

School of Physics and Astronomy

College of Engineering and Physical Sciences

University of Birmingham

April 2023

UNIVERSITY OF BIRMINGHAM

University of Birmingham Research Archive
e-theses repository



This unpublished thesis/dissertation is under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:



Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.

Unless otherwise stated, any material in this thesis/dissertation that is cited to a third-party source is not included in the terms of this licence. Please refer to the original source(s) for licencing conditions of any quotes, images or other material cited to a third party.

ABSTRACT

Astronomers want accurate physical properties of stars like age, mass, and radius. We can estimate these by comparing observable parameters to those from models of stellar evolution. This has been made possible on a large-scale with recent astronomical surveys and the field of asteroseismology probing inside stars. In the first chapter of this thesis, I introduce asteroseismology — the study of stellar oscillations. I choose to focus on stars which oscillate like the Sun with masses from 0.8 to 1.2 $\mathrm{M}_{\odot}$ undergoing their main sequence and subgiant phases of evolution. To date, we have observed oscillations in hundreds of these stars. With the upcoming space-based *PLATO* mission, we anticipate observations of $\sim 10^4$ more solar-like oscillators. In this thesis, I aim to develop probabilistic modelling methods which can quickly and easily scale to such huge numbers of stars. Furthermore, we know our stellar models are wrong. It is important to accurately quantify this uncertainty if we are to use stellar parameters to understand stellar populations. In Chapters 2 and 3, I present a novel approach for improving the inference of fundamental stellar parameters using a hierarchical Bayesian model. I introduce a statistical treatment which 'pools' helium abundance (Y) and the mixing-length theory parameter (α_{MLT}) to incorporate information about their distributions in the population. Specifically, I model Y as a distribution centred on a linear enrichment law parametrised by $\Delta Y/\Delta Z$. I test our method on a sample of dwarfs and subgiants observed by *Kepler* with $0.8 < M/\mathrm{M}_{\odot} < 1.2$. Exploring various levels of pooling parameters, with and without the Sun as a calibrator, I report $\Delta Y/\Delta Z = 1.05^{+0.28}_{-0.25}$ when the Sun is included in the sample. Despite marginalising over uncertainties in Y and α_{MLT}, I am able to report statistical uncertainties of 2.5 per cent in mass, 1.2 per cent in radius, and 12 per cent in age. Moreover, my approach can be extended to larger samples. This will enable further uncertainty reduction in fundamental parameters and

data-driven insight into population-level distributions.

There is additional information on Y to be gained from detailed asteroseismology. Acoustic glitches, which arise from rapid changes in stellar structure (e.g. from helium ionisation), leave a periodic signature in the mode frequencies (ν_{nl}) of solar-like oscillators. I explore the theoretical background behind this effect in Chapter 4. Then, in Chapter 5, I present a new method for modelling glitch signatures in the radial mode frequencies using a Gaussian Process (GP). The GP provides a statistical treatment of uncertainty in the functional form of our model for ν_{nl}. Using a model star and 16 Cyg A, I compare this approach to another method which models the smooth component of the function using a 4th-order polynomial. My results show that the GP method accurately determines the strength and location of glitches caused by He II ionisation and the base of the convective zone. I find that using a prior to inform the glitch parameters in my method reduces the occurrence of extreme, unrealistic solutions in the posterior. Furthermore, I demonstrate that the GP approach outperforms the polynomial by marginalising over the lesser signature of He I ionisation. However, inclusion of the He I ionisation glitch in the model remains a question. Overall, my results suggest that the GP method should be further tested on more solar-like oscillators and then integrated into the hierarchical model presented in this work.

Dedicated to

Hannah



In Loving Memory of

Angela and Barry

ACKNOWLEDGEMENTS

I am tremendously grateful to my PhD supervisor Guy R. Davies for his support and excellent mentorship throughout my postgraduate research. Additionally, I thank my first co-supervisor Andrea Miglio for his guidance before his move to the University of Bologna. I extend my thanks to Amaury Triaud, who took over from Andrea as my co-supervisor and has since been an inspiring mentor and leader of our research group. I would like to thank in advance my PhD examiners Daniel Reese and Chris Moore, and viva chairperson Annelies Mortier.

It has been a pleasure to work with the Sun, Stars, and Exoplanets research group here at the University of Birmingham. I would like to thank all of my fellow current and former PhD students with whom I have exchanged ideas, advice, laughs, and stories. I also appreciate the support of the Head of School Bill Chaplin, and other postdoctoral researchers in our group, notably M. Nielsen, Tanda Li (李坦达), and W. Ball. I extend my thanks to the support staff in the School of Physics and Astronomy, particularly our office secretary Lou for her unwavering support over a challenging few years. I also thank my external collaborators, for example Nick Saunders at the University of Hawaii and Simon Murphy at the University of Southern Queensland, with whom I have worked during my graduate studies.

I acknowledge the support of the public who have indirectly funded the research in this work via various publicly funded agencies. Specifically, I thank the Science and Technology Facilities Council (STFC) who funded my PhD. I also acknowledge that this work is a part of a project that has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (CartographY; grant agreement ID 804752). Finally, I am grateful for the financial and academic support from the Alan Turing Enrichment scheme. This scheme provided me the opportunity to learn from and collaborate with the wider

world of machine learning in science.

Last but not least, I am hugely grateful for the support of my loving family and friends. Particularly, I thank my fiancé Hannah, my parents Katy and Paul, and my brother Dom.

CONTENTS

Abstract i

Acknowledgements v

Contents vii

List of Figures xi

List of Tables xv

1 Inferring Stellar Parameters with Asteroseismology 1

1.1 Understanding the Stars . . . 1

1.2 Asteroseismology of Solar-Like Oscillators . . . 3

1.2.1 A Brief History of Asteroseismology . . . 3

1.2.2 A Brief Theory of Solar-Like Oscillations . . . 5

1.3 Modelling Many Stars with Asteroseismology . . . 9

1.4 Modelling Stars the Bayesian Way . . . 12

2 Hierarchical Bayesian Models 17

2.1 Introduction . . . 17

2.2 Stellar Distances in an Open Cluster Analogue . . . 18

2.2.1 Simple Model . . . 19

2.2.2 Hierarchical Model . . . 21

2.2.3 Inferring the Model Parameters . . . 22

2.2.4 Comparing the Models . . . 23
2.2.5 Scaling with the Number of Stars . . . 26
2.3 Inferring Physical Properties of Stars . . . 28

3 Hierarchically Modelling Dwarf and Subgiant Stars 31
3.1 Introduction . . . 31
3.2 Data . . . 34
3.3 Methods . . . 37
3.3.1 Grid of Stellar Models . . . 38
3.3.2 Statistical Models . . . 40
3.3.3 The Sun as a Star . . . 46
3.3.4 Sampling . . . 46
3.4 Results . . . 48
3.4.1 Stellar Parameter Results . . . 48
3.4.2 Population Parameter Results . . . 51
3.5 Discussion . . . 51
3.5.1 Helium Enrichment . . . 54
3.5.2 Mixing-Length Theory . . . 56
3.5.3 Comparison with APOKASC Results . . . 58
3.5.4 Systematic Uncertainties . . . 62
3.5.5 Outliers . . . 62
3.6 Conclusion . . . 63

4 Acoustic Glitches in Solar-Like Oscillators 65
4.1 Introduction . . . 65
4.2 A One-Dimensional Example of a Glitch . . . 67
4.3 Acoustic Glitch Signatures in Stellar Oscillations . . . 75
4.3.1 Sun-Like Model Star . . . 77
4.3.2 Helium Ionisation Glitch . . . 77

4.3.3 Base of the Convective Zone Glitch . . . 87

5 Modelling Acoustic Glitches with a Gaussian Process **89**

5.1 Introduction . . . 89

5.2 Data . . . 91

5.3 Methods . . . 93

5.3.1 The V19 Method . . . 94

5.3.2 The GP Method . . . 95

5.4 Results . . . 99

5.4.1 Test Star . . . 99

5.4.2 16 Cyg A . . . 102

5.5 Discussion . . . 103

5.6 Conclusion . . . 107

6 Conclusion & Future Prospects **109**

Data Availability & Software **115**

References **117**

A Hierarchically Modelling Dwarf and Subgiant Stars **133**

A.1 Artificial Neural Network . . . 133

A.1.1 Train, Validation, and Test Data . . . 136

A.1.2 Tuning . . . 136

A.1.3 Testing . . . 140

A.2 Prior Distributions . . . 141

A.3 The Synthetic Population . . . 142

A.3.1 Stellar Parameters . . . 143

A.3.2 Population Parameters . . . 144

A.4 The Sun as a Star . . . 144

LIST OF FIGURES

1.1 Surface spherical harmonic oscillation modes for a few angular degrees (l) with azimuthal order $m = 0$. 5

1.2 The power spectral density of 16 Cyg A. 6

1.3 Signal-to-noise ratio as a function of frequency and echelle diagram for 16 Cyg A. 8

1.4 Colour-magnitude diagram highlighting the region occupied by dwarf and sub-giant solar-like oscillators. 11

2.1 Probabilistic graphical model for the simple (non-hierarchical) model 20

2.2 Probabilistic graphical model for the hierarchical model 22

2.3 Plots comparing results from the simple model to the HBM 24

2.4 A corner plot showing the marginalised and joint posterior distributions of the distance to the first three stars 25

2.5 The joint and marginalised posterior distributions for the mean and standard deviation of stellar distances in the cluster 26

2.6 Joint and marginalised posterior distributions from the HBM for 20, 80 and 320 stars. 27

3.1 The luminosity against effective temperature of the sample of 81 *Kepler* dwarfs and subgiants studied in this work. 37

3.2 A probabilistic graphical model of the partially-pooled hierarchical model. . . . 43

3.3 Corner plots showing the joint and marginalised sampled posterior distributions for the PP and PPS hyperparameters. 52

3.4 The same as Fig. 3.3 but for the MP (top) and MPS (bottom) models. 53

3.5 The results for initial helium fraction against initial heavy-element fraction for each star from the PPS model. 54

3.6 Kernel density estimates of the distribution of statistical uncertainties in the results from each model. 59

3.7 The mean and standard deviation in age, mass and radius results from the PPS model. 61

4.1 A diagram showing a one-dimensional medium with a small structural perturbation. 68

4.2 The change in mode frequency induced by a rapid change in sound speed for the 1D example. 71

4.3 The same as Figure 4.2 but showing the phase offset required to satisfy the boundary conditions. 73

4.4 Best fit of Equation 4.21 to $\delta\omega_n$ for $n = 1, \ldots, 50$, $L = 1$, and $c = 1$. 74

4.5 The sound speed gradient of model R plot against the acoustic depth. 76

4.6 The depressions in the first adiabatic exponent caused by the ionisation of hydrogen and helium. 78

4.7 Temperature-density profile for a Sun-like star. 80

4.8 The effect on the sound speed profile for three solar-like stars with initial helium mass fractions of 0.26, 0.28, and 0.3. 81

4.9 Structure kernels ($\mathcal{K}_{c^2,\rho}, \mathcal{K}_{\rho,c^2}$) scaled by the stellar radius as a function of fractional acoustic depth. 83

4.10 Discontinuity in the density gradient at the base of the convective zone for three solar-like model stars. 87

5.1 Echelle plot of simulated true and 'observed' radial mode frequency data from model star S for worst, better, and best case scenarios. 91

5.2 50 random draws from the V19 and GP results showing the total glitch signal as a function of frequency. 100

5.3 Samples from the V19 and GP methods for the acoustic depths of the helium glitch ($\tau_{\rm He}$) and base of the convective zone ($\tau_{\rm cz}$) 101
5.4 Samples of the helium glitch amplitude at a reference frequency of 3000 μHz fit with the V19 and GP methods for the worst, better, and best case test data. . . . 102
5.5 Results for the glitch fit to 16 Cyg A using the V19 method and GP method. . . 104
5.6 Echelle-like diagrams showing mode frequencies plot against their frequency modulo 133 μHz for the best case scenario. 105
5.7 The difference between the model without the glitch components predicted by theV19 method and the GP method. 106

A.1 An artificial neural network comprising H hidden layers with N neurons per layer. 134
A.2 The MAE as a function of epochs for the train and validation datasets. 139
A.3 The error between a given parameter in the test dataset and the ANN prediction for that parameter. 140
A.4 A beta distribution compared to a uniform distribution. 142
A.5 Kernel density estimates of the distributions of statistical uncertainties from each model for the sample of synthetic stars. 143
A.6 Corner plot showing the marginalised and joint posterior distributions between the PP model hyperparameters for 100 synthetic stars. 145
A.7 The same as Fig. A.6 but for the MP model. 146
A.8 A corner plot showing the sampled marginal and joint posterior distributions for the Sun as a part of the NP model. 148

LIST OF TABLES

3.1 The observables and their respective uncertainties for 5 stars in the sample of 81 stars. . . . 36

3.2 Stellar model grid parameters for train and test datasets. . . . 38

3.3 Solar input data. . . . 46

3.4 The median of the marginalised posterior samples for selected parameters output by the NP model. . . . 49

3.5 The same as Table 3.4, but for the PP model. . . . 49

3.6 The same as Table 3.4, but for the PPS model. . . . 49

3.7 The same as Table 3.4, but for the MP model. . . . 50

3.8 The same as Table 3.4, but for the MPS model. . . . 50

3.9 Hyperparameter results for each model with the omission of Y_P. . . . 51

5.1 Observations of radial mode frequency ν_n at radial order n for model star S and 16 Cyg A. . . . 92

5.2 The GP model parameters and their prior distributions. . . . 97

A.1 The median and standard deviation for each parameter in the training data, used to standardise the dataset. . . . 138

A.2 The median error and median absolute deviation of the error for each ANN output parameter from the test dataset. . . . 141

A.3 Solar results from the NP model. . . . 147

Chapter 1

INFERRING STELLAR PARAMETERS WITH ASTEROSEISMOLOGY

In this chapter, I introduce the current state of modelling stars with asteroseismology and the types of stars being studied in this thesis. I start with a brief history of understanding the stars spanning the last century. In Section 1.2, I introduce asteroseismology of stars which oscillate like the Sun. Then, I provide examples of asteroseismology being used to model large samples of dwarf and subgiant stars in Section 1.3. Finally, I introduce the concept of modelling stars the 'Bayesian way' with some examples of current methods and their limitations.

1.1 Understanding the Stars

Early efforts to understand the stars began by finding relations between their spectral classification and magnitude in different photometric bands. On what was later called a Hertzsprung-Russell (HR) diagram (e.g. Russell, 1914), astronomers plot the absolute magnitude (or luminosity, L) of a star against its spectral class (or effective temperature, $T_{\rm eff}$). They found that stars were not uniformly distributed on the HR diagram, but were instead grouped in distinct sequences. For example, the narrow diagonal band where most stars were found was called the *main sequence*. Astronomers wondered why stars adhered to these sequences and to what extent they evolved with time.

Initial insight into stellar evolution came about when astronomers studied open clusters on

the HR diagram (Trumpler, 1930). These are groups of stars found at a similar distance and close together on the sky. Assuming clusters formed at the same time with similar chemical abundances, the only expected difference between stars is their mass and multiplicity. Using stars of known mass (e.g. from orbital solutions to binary systems) and luminosity distributions, astronomers could trace lines of similar mass and composition from younger to older clusters (e.g. Sandage, 1957). Coupled with early arguments from stellar physics (e.g. Chandrasekhar, 1939), scientists formed stellar evolutionary tracks — the path a star takes on the HR diagram during its evolution. Deriving the stellar radius (R) from the relation $L \propto R^2 {T_{\rm eff}}^4$, they inferred that stars started life on the dimmer edge of the main sequence and became brighter and larger throughout most of their lifetime. At some point, stars would leave the main sequence, rapidly cool, expand, and then ascend a region of the HR diagram known as the *red giant branch*.

The chemical composition, energy sources and evolutionary physics of stars were still uncertain. Spectroscopy revealed relative abundances of elements excited in stellar atmospheres, but it was difficult to determine their absolute abundances. Payne (1925) proposed that stars comprised mostly hydrogen and helium, later reinforced by estimates of helium content in the Sun (e.g. Schwarzschild, 1946). The idea was radical at the time because it did not match the abundance of elements found on Earth. Meanwhile, advancements in nuclear science spawned the theory of stellar nucleosynthesis (Hoyle, 1946). Stars produce elements heavier than hydrogen and helium through nuclear fusion reactions. This discovery explained the production of many elements in the universe, tying the formation and evolution of stars to planetary formation and the conditions for life in the universe.

In the second half of the 20th century, astronomers advanced the theory of stellar evolution. With this came computational methods for simulating stars (e.g. Kippenhahn, Weigert, and Hofmeister, 1967). Astronomers could start to compare observations of stars and clusters with numerical solutions to stellar structure equations. We call this process *modelling stars*. Early models of the Sun benefited from independent geological age estimates and neutrino-production rates from observing cosmic rays. With these constraints, and the Sun as a calibrator, researchers could start to refine their models and test them on other stars. For example, Iben (1967) computed

stellar models for main sequence stars with similar masses to the Sun using observations of the clusters M67 and NGC 188.

On a similar timescale, a new field emerged which gave astronomers a model-independent way of studying the inside of stars. Identifying regular perturbations of the surface of the Sun, researchers realised that they pertained to high-order, stochastically-driven spherical harmonic oscillations. Measuring these oscillation modes, they could study the Sun similarly to how seismologists study the Earth. Astronomers studying multiple non-radial stellar oscillations distinguished this new field, called *asteroseismology*, from the well-known study of radially pulsating stars (e.g. Cepheid variables; Leavitt, 1908). In Section 1.2 we briefly elaborate on the history and theory of asteroseismology for stars like the Sun.

Today, astronomers have a huge abundance of data with which to determine stellar parameters and test theories of stellar evolution. In Section 1.3, we give some recent examples of Sun-like asteroseismic populations and their context in the wider HR diagram. With big datasets and complex, multi-parameter models, there has been a recent shift towards Bayesian statistical methods to model these stars. Therefore, we introduce the concept of modelling stars 'the Bayesian way' in Section 1.4. There, we introduce some tools used to model these stars and then highlight some common assumptions made.

1.2 Asteroseismology of Solar-Like Oscillators

1.2.1 A Brief History of Asteroseismology

Several decades ago, 5-minute oscillations in the radial velocity of the solar surface were observed by Leighton, Noyes, and Simon (1962), leading to the inference of acoustic waves trapped beneath the solar photosphere (Ulrich, 1970). A further decade of study culminated in the measurement of regular patterns of oscillation modes in the Sun; for example, from measurements of radial velocity (Claverie et al., 1979) and total irradiance (Woodard and Hudson, 1983). Initially thought to be short-lived irregularities on the surface, these modes were found to be compatible with stochastically excited standing waves penetrating deep into

the Sun. Later, Deubner and Gough (1984) introduced the word *helioseismology* to describe the study of the solar interior using observations of these modes. Helioseismology was soon responsible for breakthrough solar research, from measuring differential rotation (Deubner, Ulrich, and Rhodes, 1979) to solving the mismatch between predicted and measured solar neutrino production (Bahcall and Ulrich, 1988).

Astronomers initially debated the mechanism driving the modes of standing pressure waves (or p modes) in the Sun. Goldreich and Keeley (1977) suggested what became the prevailing theory, that the p modes were stochastically excited by near-surface convection. Therefore, astronomers expected solar-like oscillations to be present in other stars with a convective envelope like the Sun. Shortly thereafter, Christensen-Dalsgaard (1984) was among those to introduce the term *asteroseismology* — the study of the internal structure of stars with many observable oscillation modes. Subsequently, solar-like oscillations were discovered in a few bright stars. Among the first were Procyon and α Cen A (Gelly, Grec, and Fossat, 1986), with individual modes later resolved by Martić et al. (1999) and Bouchy and Carrier (2001) respectively.

Instrumental and atmospheric noise limited the progress of asteroseismology with ground-based equipment to studies of small number of bright dwarf stars. Asteroseismology of solar-like oscillators requires high-cadence ($\sim$ 1 to 10 min) brightness observations over long baselines ($\sim$ 1 yr) with parts-per-million precision (e.g. Schofield et al., 2019). The first dedicated space-based missions which met these requirements arrived in the late 2000s, accelerating progress in the field. Initially, the *CoRoT* mission (Baglin et al., 2006) detected solar-like oscillations in thousands of red giant stars (De Ridder et al., 2009; Mosser et al., 2010). Then, the *Kepler* mission (Borucki et al., 2010) yielded oscillations in thousands more red giants (Pinsonneault et al., 2014) and hundreds of main sequence stars similar to the Sun (Serenelli et al., 2017). Most recently, *TESS* (Ricker et al., 2015) added thousands more dwarf and giant stars to the roster of solar-like oscillators (Silva Aguirre et al., 2020; Hon et al., 2021; Hatt et al., 2023). For a more detailed review of modern asteroseismology, see Aerts (2021).

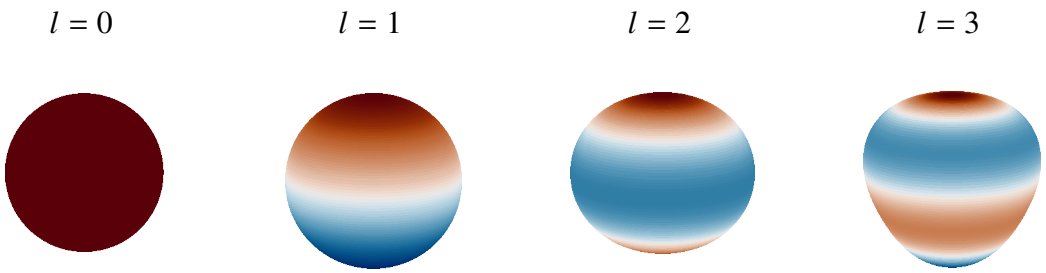


Figure 1.1: Surface spherical harmonic oscillation modes for a few angular degrees (l) with azimuthal order $m = 0$. The colour-map represents radial displacement at the surface, with *red* and *blue* corresponding to displacement inward and outward respectively. The coloured regions pulsate in and out, and the white regions represent stationary nodes on the surface.

1.2.2 A Brief Theory of Solar-Like Oscillations

We can approximate oscillations on the surface of a star by spherical harmonic functions with angular degree l and azimuthal order m,

$$Y_l^m(\theta, \varphi) \propto P_l^m(\cos\theta)\, e^{im\varphi}, \tag{1.1}$$

where θ is the polar angle, φ is the azimuthal angle, and P_l^m is the associated Legendre polynomial. The angular degree is the number of nodes on the surface of the star. We show a representation of the surface spherical harmonics for the first four angular degrees in Figure 1.1. For each l, there exists $2l + 1$ solutions with different azimuthal order (m) corresponding to the different orientations of the nodes over the spherical surface. Additionally, the oscillation modes have unique frequency solutions for different radial orders (n) proportional to the number of nodes radially throughout the star.

In solar-like oscillators, p modes are stochastically excited by near-surface convection. Typically, the timescale of this process drives high-order modes in main sequence stars ($n \sim 20$). We can identify these modes in a frequency-power spectrum derived from photometric or radial velocity time series observations. For instance, both stars in the 16 Cygni planet hosting system are solar-like oscillators with similar properties to the Sun (Metcalfe et al., 2012, 2014; Davies et al., 2015). Using 16 Cyg A as an example, we downloaded the power spectrum determined

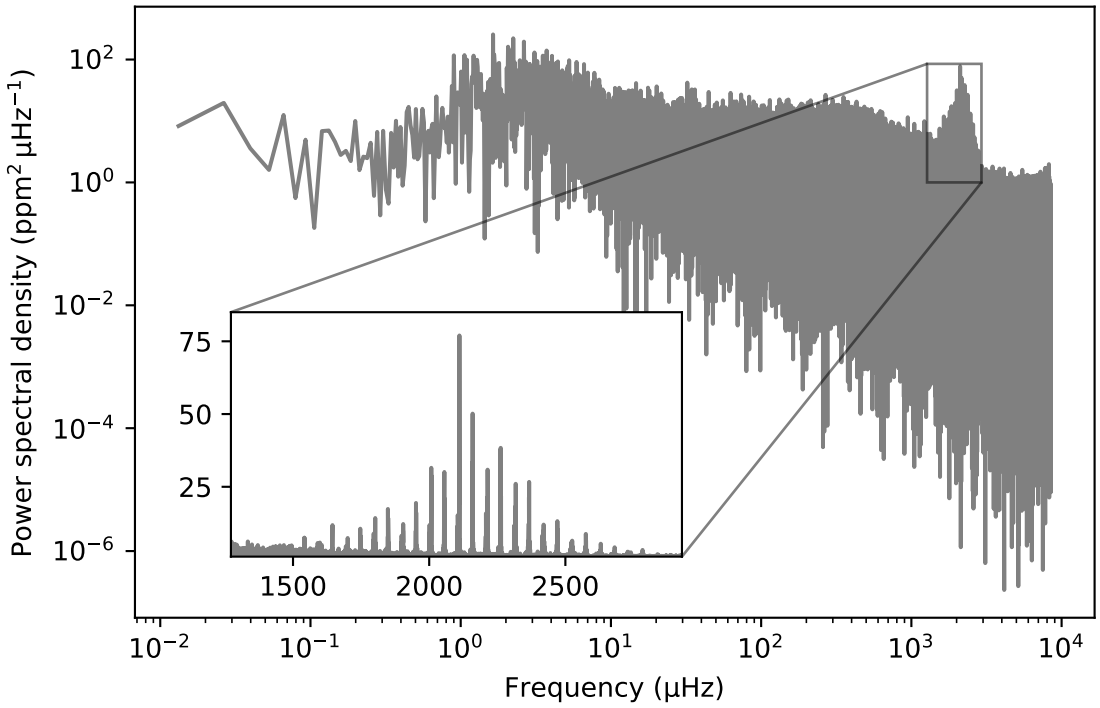


Figure 1.2: The power spectral density of 16 Cyg A. The inset plot highlights the Gaussian-like power excess in the larger plot.

by the *Kepler* Asteroseismic Science Operations Centre (KASOC) using *Kepler* observations[1]. Shown in Figure 1.2, the power spectrum of 16 Cyg A has a distinct power excess around 2000 μHz.

The power excess has a Gaussian-like shape around frequencies compatible with the near-surface convective timescale responsible for mode excitation. We call the centre of this Gaussian the *frequency at maximum power* ($\nu_{\rm max}$). Dependent on near-surface conditions, Brown et al. (1991) suggested $\nu_{\rm max}$ scales with the acoustic cut-off frequency — the highest frequency at which acoustic waves can reflect near the stellar surface. Subsequently, Kjeldsen and Bedding (1995) found that $\nu_{\rm max} \propto g{T_{\rm eff}}^{-1/2}$ where g and $T_{\rm eff}$ are the near-surface gravitational field strength and effective temperature. Hence, $\nu_{\rm max}$ for a solar-like oscillator decreases as it evolves, from $\sim 1000\,\mu$Hz on the main sequence to $\sim 10\,\mu$Hz near the tip of the red giant branch.

Looking closely at the power excess in Figure 1.2, we can see a comb of approximately equally spaced peaks. Each peak corresponds to an oscillation mode, with their collective

[1] https://kasoc.phys.au.dk

central frequencies and separations providing information about the internal stellar structure. Naturally, higher frequency modes correspond to higher radial order, n. However, the angular degree and azimuthal order are harder to identify. We saw in Figure 1.1 how modes of higher l have more anti-nodes on the surface. Therefore, the overall effect of the oscillations cancel out when measuring irradiance integrated over the stellar surface. Consequentially, observed mode amplitude decreases with l, leaving only $l \lesssim 3$ detectable. Therefore, we can assume the tallest peaks are $l = 0, 1$, and the smaller peaks are $l = 2, 3$, all modulated by the wider Gaussian-like envelope.

Modes with different m cannot be distinguished for a spherically symmetric, non-rotating star. In the case of a rotating or distorted (asymmetric) star, the observed mode frequencies will split for different m via the Doppler effect. Measuring this splitting can constrain the rotation rate of the star (e.g. Deheuvels et al., 2012; García et al., 2014; Davies et al., 2015). Such studies have lead to breakthrough research on the rotational evolution of stars (e.g. Angus et al., 2015; van Saders et al., 2016; Hall et al., 2021). However, we will hereafter consider only the case of a slowly rotating, spherically symmetric star.

For an acoustic wave in a one-dimensional homogeneous medium, we would expect each mode of oscillation to be an integer multiple of the fundamental mode. While the case for a star is more complicated, we can also approximate the frequencies for different modes as a multiple of a characteristic frequency. Tassoul (1980) found that the modes could be approximated by assuming the asymptotic limit, where $l/n \to 0$. This lead to the following expression for the frequency of a mode for a given n and l (cf. Gough, 1986),

$$\nu_{nl} \simeq \left(n + \frac{l}{2} + \varepsilon\right) \nu_0 + O(\nu_{nl}^{-1}), \tag{1.2}$$

where ε is some constant offset and $O(\nu_{nl}^{-1})$ represents higher order terms. The characteristic frequency (ν_0) is the inverse of the acoustic diameter (sound crossing time),

$$\nu_0 = \left(2 \int_0^R \frac{\mathrm{d}r}{c(r)}\right)^{-1}, \tag{1.3}$$

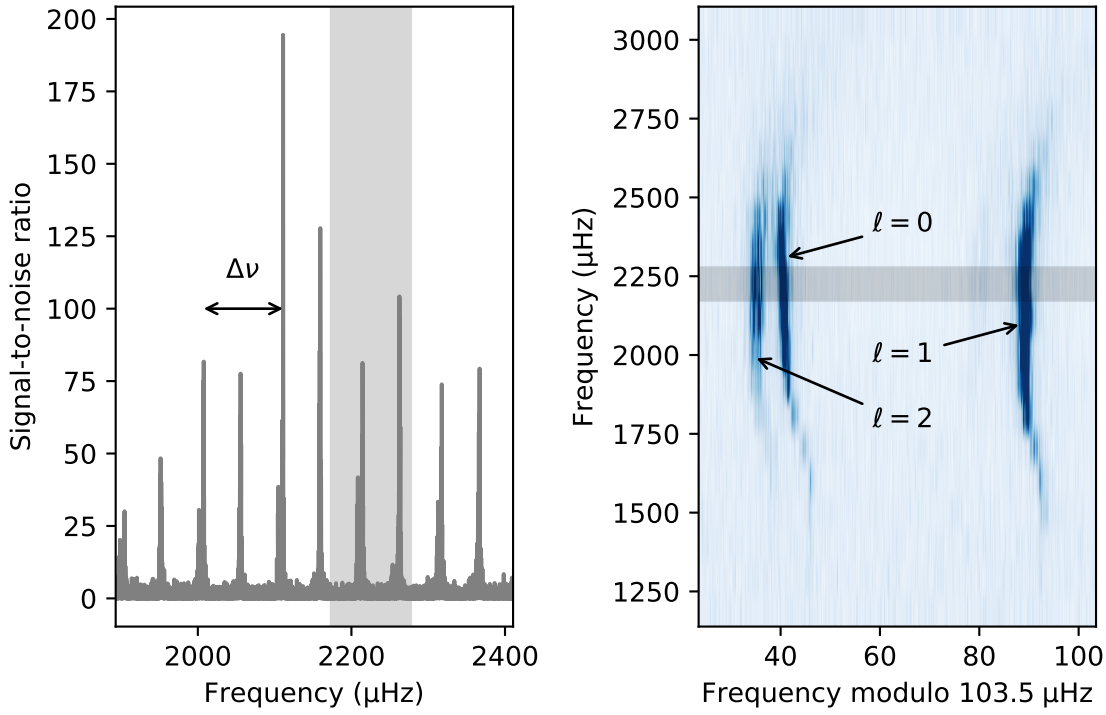


Figure 1.3: *Left:* A section of the spectral signal-to-noise ratio (SNR) against frequency for 16 Cyg A. The large frequency spacing ($\Delta\nu$) between two radial modes is annotated with a double-headed arrow. The shaded region corresponds to a single row (also highlighted) in the echelle plot. *Right:* The echelle plot shows the spectral SNR such that a darker colour represents a higher SNR. Each row spans 103.5 μHz and is stacked in order of frequency. The apparent ridges are labelled according to the angular degree (l) of the modes they represent.

where $c(r)$ is the sound speed as a function of radius (r), and R is the radius of the star. Similarly to other variable stars, Ulrich (1986) found that this characteristic frequency relates to mean stellar density by $\nu_0 \propto \overline{\rho}^{1/2}$. While ν_0 is not directly detectable in solar-like oscillators, we can approximate it by taking the difference between consecutive modes of the same angular degree, $\Delta\nu_{nl} = \nu_{nl} - \nu_{n-1\,l}$. Thus, estimates of a global (or average) *large frequency separation*, $\Delta\nu$, can also provide information about the density of a star. This leads to independent constraint on stellar mass and radius (e.g. from scaling relations to the Sun).

The asymptotic expression helps us identify modes in a star. If the first term of Equation 1.2 was exact, we would expect odd and even modes to be grouped together and separated by $\Delta\nu/2$. To show this, we revisit the power spectrum of 16 Cyg A and estimate a signal-to-noise ratio (SNR) by dividing out a moving median in log-frequency steps of 0.005 dex. We can see

the regular pattern predicted by Equation 1.2 in the left panel of Figure 1.3. Every other mode is approximately separated by $\Delta\nu$. To see this effect over the wider spectrum, we created an *echelle* plot in the right panel. Folding the spectrum by an estimate of $\Delta\nu$ reveals a sequence of ridges corresponding to modes of different angular degree. Odd and even angular degree are grouped together, although they do not lie on top of each other. The small difference between modes of different l is described by the higher order terms neglected from Equation 1.2. A faint ridge corresponding to $l = 3$ modes is also visible next to the $l = 1$ ridge. However, 16 Cyg A represents one of the highest SNR dwarf stars observed by *Kepler*, and the $l = 3$ ridge is otherwise not usually visible.

Repeating this kind of analysis on a large-scale can provide a wealth of observables from which to model stars. In the next section, we give some examples where asteroseismology was used to model many solar-like oscillators.

1.3 Modelling Many Stars with Asteroseismology

Recent space-based missions like *Kepler* and *TESS* enable asteroseismology with many stars. These missions were primarily designed to detect exoplanets using the transit method. However, their instrumental requirements were also well-suited to asteroseismology. This allowed astronomers to study regions of the HR diagram in more detail, from main sequence dwarf stars to red giants. We draw particular attention to *Kepler*, the first of the two missions. *Kepler*'s primary mission observed a single $\sim 100\,\mathrm{deg}^2$ patch of sky for around 4 years (Borucki et al., 2010), cut short by a mechanical failure which spawned the secondary *K2* mission (Howell et al., 2014). *Kepler* took photometric measurements in long (30 min) and short (60 s) cadences, with the Nyquist frequency of the latter high enough to detect main sequence oscillators. During its primary mission, large asteroseismic datasets were gathered. The long baseline of 4 years allowed for high frequency precision not since seen by *K2* or *TESS*.

Of the asteroseismic targets found by *Kepler*, we focus on dwarf and subgiant solar-like oscillators. These are stars with a similar mass to the Sun on the main sequence or shortly

after the turn-off before ascending the red giant branch. From an asteroseismic perspective, the power spectra of dwarf stars are relatively simple. As solar-like oscillators approach the red giant branch, their modes begin to couple with buoyancy-driven modes in the core (Bedding et al., 2011; Mosser et al., 2011). This leads to irregular patterns which can be used to separate evolutionary state and study properties of the core (e.g. Mosser et al., 2015). On the other hand, the relative simplicity of main sequence oscillators make them a suitable place to start our study. Furthermore, the work in this thesis anticipates the upcoming *PLATO* mission which aims to observe $\sim 10^4$ dwarf and subgiant oscillators (Rauer et al., 2016).

Modelling many stars with asteroseismology is complemented by other recent large-scale stellar surveys. High-precision astrometry from the space-based *Gaia* mission provides precise distances and orbital solutions for over a billion sources (Gaia Collaboration et al., 2016). Additionally, the APOGEE large-scale spectroscopic survey provides precise chemical abundances for over half a million targets to date (Majewski et al., 2017; Jönsson et al., 2020). When combined, these surveys have enabled studies of the assembly history of our galaxy (e.g. Helmi et al., 2018) which have since been strengthened by asteroseismology (Chaplin et al., 2020; Montalbán et al., 2021).

In Figure 1.4, we show a colour-magnitude diagram made using magnitudes and parallaxes from *Gaia* Data Release 3 (DR3; Gaia Collaboration et al., 2022). For this illustrative plot, we have neglected the effect of extinction from dust in the interstellar medium. The background distribution shows solar-neighbourhood *Gaia* sources with a parallax greater than 5 mas for context. Over-plot is the distribution of *Kepler* objects cross-matched with *Gaia* DR3[2]. The densest region lies in the low- to intermediate-mass main sequence (0.8 to 1.2 $M_\odot$) but we can also see a clear red giant branch and red clump in the upper-right of the distribution. The inset plot draws attention to the region occupied by dwarf and subgiant solar-like oscillators, where we give some examples.

Chaplin et al. (2011) identified the first large catalogue of $\sim$ 500 dwarf and subgiant solar-like oscillators (black circles in Figure 1.4) by measuring $\Delta\nu$ and $\nu_{\rm max}$ in *Kepler* data. Later,

[2]The cross-matched dataset was obtained from https://gaia-kepler.fun

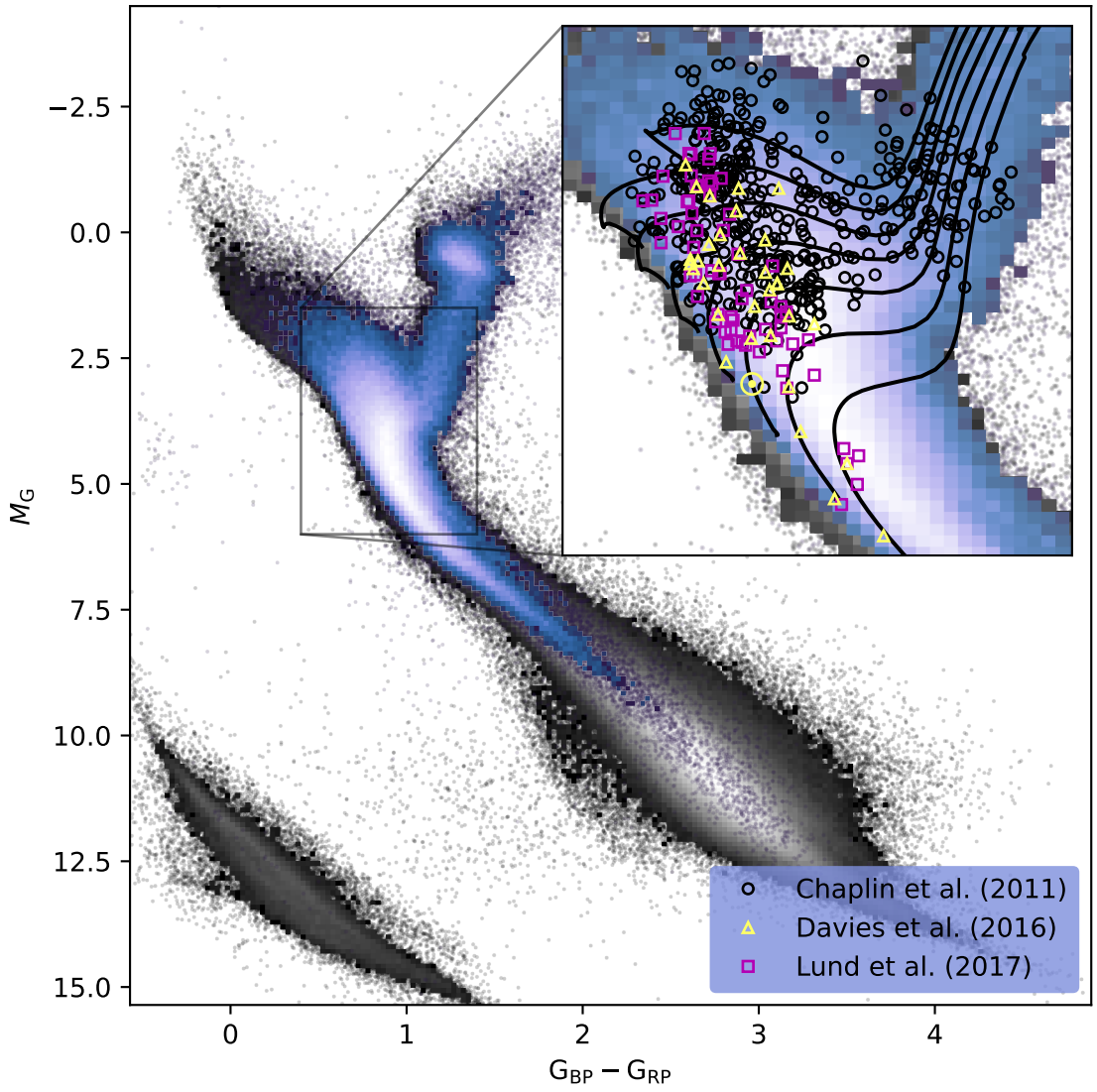


Figure 1.4: Colour-magnitude diagram highlighting the region occupied by solar-like oscillators. *Main plot:* The background 2D histogram (greyscale) shows the density of *Gaia* DR3 targets with parallax > 5 mas (distance $\lesssim 200$ pc). The foreground 2D histogram (blue-purple) shows the density of all *Kepler* targets in *Gaia* DR3. Scattered points are shown where the density is less than 10 points per bin. *Inset plot:* Stellar evolutionary tracks with solar metallicity are given by the black lines for 0.8 to 1.4 $M_\odot$ in steps of 0.1 $M_\odot$. The circles in the inset axes correspond to *Kepler* solar-like oscillators with references given in the legend. The Sun is given by the '$\odot$' symbol.

Chaplin et al. (2014) determined ages, masses and radii for these stars using $\Delta\nu$ and $\nu_{\rm max}$ complemented by photometry and ground-based spectroscopy where available. The subsequent arrival of APOGEE spectroscopy allowed Serenelli et al. (2017) to revisit this sample with a more consistent set of $T_{\rm eff}$ and metallicity. By comparing observations to models of stellar evolution, they found radii, masses and ages with uncertainties of around 3, 5 and 20 per cent respectively.

For a subset of these stars, the SNR was high enough to identify many individual modes. Appourchaux et al. (2012) were among the first to publish individual mode frequencies (ν_{nl}) for around 60 of these stars. This sample was later modelled by Metcalfe et al. (2014) who found observing individual modes doubled precision of radii, masses and ages over using $\Delta\nu$ and $\nu_{\rm max}$ alone. Around the same time, the number of confirmed exoplanets was increasing rapidly (Burke et al., 2014). This motivated a more detailed study of 35 exoplanet host stars by Silva Aguirre et al. (2015) with modes identified by Davies et al. (2016). We can see these as yellow triangles in Figure 1.4. The remaining best targets, referred to as the LEGACY sample, were modelled by Silva Aguirre et al. (2017) using modes identified by Lund et al. (2017). We show these as magenta squares in Figure 1.4.

1.4 Modelling Stars the Bayesian Way

Modelling stars 'the Bayesian way' involves estimating the probability distributions of stellar parameters ($\boldsymbol{\theta}$) given observed parameters ($\boldsymbol{y}$). For example, the parameters of interest might be the age, mass and initial fractional heavy-element abundance of the star ($\boldsymbol{\theta} = t_\star, M, Z$) and we might observe the effective temperature, luminosity, and large frequency spacing ($\boldsymbol{y} = T_{\rm eff}, L, \Delta\nu$). We can do this starting with Bayes' theorem for the *posterior* probability density,

$$p(\boldsymbol{\theta} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \boldsymbol{\theta})\, p(\boldsymbol{\theta})}{p(\boldsymbol{y})}, \tag{1.4}$$

where $p(\boldsymbol{y} \mid \boldsymbol{\theta})$ is the *likelihood*, $p(\boldsymbol{\theta})$ is the *prior*, and $p(\boldsymbol{y})$ is the *evidence*. We should choose a prior to represent our expectation for $\boldsymbol{\theta}$. The prior dominates if the observations are bad (i.e.

have a small likelihood) and we return our current 'belief' (or expectation). However, if the observations are good enough, the likelihood has more influence on the posterior and our belief is updated. As a result, we can use the Bayesian methodology to systematically update our beliefs. This provides a statistical scheme from which to build our stellar models.

Let us consider a stellar model with three parameters, $\boldsymbol{\theta} = t_\star, M, Z$. For instance, we might want the age of a star to help date a galactic merger, or its mass to characterise an exoplanetary system. To get the probability distribution of age given our observations, we must *marginalise* over the joint posterior with respect to all other parameters. The marginal posterior probability distribution for $t_\star$ is,

$$p(t_\star \mid \boldsymbol{y}) = \int_0^\infty \int_0^1 p(\boldsymbol{\theta} \mid \boldsymbol{y}) \, \mathrm{d}M \mathrm{d}Z, \tag{1.5}$$

where the bounds of the integral are chosen across all space for M and Z. In other words, the marginal distribution involves integrating over the uncertainty in all other model parameters.

The marginal posterior distribution depends on our choice of function which maps $\boldsymbol{\theta}$ to predict observable parameters, $\tilde{\boldsymbol{y}} = f(\boldsymbol{\theta})$. The likelihood then compares the predictions ($\tilde{\boldsymbol{y}}$) with observations ($\boldsymbol{y}$). When modelling stars, there is no analytical form to f. Instead, we use numerical simulations which take in $\boldsymbol{\theta}$ and approximate $\boldsymbol{y}$. For example, Modules for Experiments in Stellar Astrophysics (MESA; originally developed by Paxton et al., 2011) is a one-dimensional stellar evolution code. Assuming spherical symmetry, MESA simulates the structure and evolution of a star along a one-dimensional radial slice. The code numerically solves a series of differential equations which govern the interior dynamics of the star (see, e.g. Kippenhahn, Weigert, and Weiss, 2013, for an introduction to stellar structure and evolution). Often starting with a cloud of gas, MESA evolves the system over dynamical time steps recording physical properties at mesh points inside the star.

We can also predict oscillation mode frequencies (ν_{nl}) with simulations. For example, the GYRE code developed by Townsend and Teitler (2013) uses the output of MESA to compute oscillation modes for a given n and l. While the physics of p mode propagation in the star is relatively well-known, our understanding of the atmospheric boundary conditions are not. As

such, there is a known discrepancy between the simulated and observed p modes. The nature of this can lead to a systematic bias when modelled $\Delta\nu$ and ν_{nl}. Corrections for this effect exist (e.g. Ball and Gizon, 2014) but are still not fully understood.

The complexity of stellar models means that the marginalised posterior distributions are not analytically derivable. One solution is to estimate the posterior with numerical methods like Markov Chain Monte Carlo (MCMC). Typically, this involves exploring parameter space with multiple calls to f for different values of $\boldsymbol{\theta}$. In the case of MCMC-based algorithms like Hamiltonian Monte Carlo (HMC) and the No U-Turn Sampler (NUTS), the gradient of f is also required. There are several open-source Python packages widely used to implement these algorithms, for instance `pymc` (Salvatier, Wiecki, and Fonnesbeck, 2016) and `numpyro` (Phan, Pradhan, and Jankowiak, 2019).

There are some existing methods for determining stellar parameters using this Bayesian approach. As an early example, Bazot, Bourguignon, and Christensen-Dalsgaard (2008) used the MCMC algorithm to sample model parameters with on-the-fly stellar model calculation. While their method could be tailored to individual stars, it was very computationally expensive. Each proposed set of $\boldsymbol{\theta}$ spawns a stellar simulation which evolves to a given age. Steps prior to this age may be discarded, and the simulation can take minutes to hours for each set of $\boldsymbol{\theta}$. This was not a viable solution for modelling the large populations of stars being observed with high-precision today.

A more efficient solution is to sample a discrete grid of models (Gruberbauer, Guenther, and Kallinger, 2012; Gruberbauer et al., 2013). For example, a recent tool which does this is the BAyesian STellar algorithm (BASTA) developed by Aguirre Børsen-Koch et al. (2022). The authors pre-compute and interpolate a grid of simulated stars corresponding to a set of relevant input parameters. They weight the grid points according to the prior and then compute the marginalised posterior for each model parameter. We could extend this by interpolating a grid of stellar models to approximate f itself. This is done in the Asteroseismic Inference on a Massive Scale (AIMS) pipeline by Lund and Reese (2018) and Rendle et al. (2019). However, interpolation methods can be slow. They scale poorly with dimensionality and number of points

on the grid. A small subset of the grid can be used to mitigate this issue, but we want a solution which can be extended to model many stars at once.

Asteroseismology has allowed us to model stars more precisely than ever before. This has exposed systematic uncertainties in common assumptions used when computing these models (e.g. Tayar et al., 2022). For instance, fractional helium abundance (Y) cannot be determined spectroscopically in cool stars. Instead, we often assume it follows a linear enrichment law (Chiosi and Matteucci, 1982; Ribas et al., 2000; Casagrande et al., 2007). Such a law assumes that helium is enriched in the interstellar medium linearly with metallicity. However, Lebreton, Goupil, and Montalbán (2014) found that varying Y by ± 0.03 can have $\gtrsim \mp 20$ per cent effect on stellar age. Therefore, assuming Y instead of incorporating it in our model leads to underestimating our uncertainty on stellar age and other correlated parameters. In Chapters 4 and 5, we will explore one way to improve inference of Y using an asteroseismic signature of helium ionisation. In the meantime, we explore a better statistical treatment of our prior on such poorly characterised parameters.

Finally, it is important that our prior represents not only our belief for individual stars, but also the population. There are many cases where we would expect stellar parameters to be correlated between stars. Not accounting for such correlations cause us to overestimate uncertainty on individual stellar parameters. For example, we believe that Y belongs to a distribution governed by chemical enrichment of the interstellar medium. Stars formed in a similar place and time should share similar chemical abundances. We can incorporate prior beliefs like this using a hierarchical Bayesian model (HBM). In the next chapter, we introduce the concept of an HBM. We explore a simple HBM in the context of astronomy and show how it can improve our estimates of stellar parameters. Following that, we present a hierarchical model on a population of dwarf and subgiant stars observed by *Kepler* in Chapter 3.

Chapter 2

HIERARCHICAL BAYESIAN MODELS

In this chapter, I introduce the concept of a hierarchical Bayesian model in the context of determining stellar parameters. Using a simplified analogy for measuring distances to stars in an open cluster, I demonstrate the advantages and usage of a hierarchical model in Section 2.2. Finally, I identify parameters which could be treated hierarchically when modelling the physical properties of stars, preparing the reader for Chapter 3.

2.1 Introduction

When modelling a single star, we can use Bayes' theorem (Equation 1.4) to estimate the posterior probability of its parameters. We choose priors to represent our current knowledge. If we observe a star at random in the Milky Way, we expect its mass, age, and chemical composition to belong to some empirical or physical distribution. For example, the initial mass function (IMF; e.g. Chabrier, 2003) is used as a prior on mass by the BASTA stellar inference code (Aguirre Børsen-Koch et al., 2022). Another prior could restrict the age of the star to no older than the universe.

Let us consider modelling a large population of stars simultaneously. For example, we may want to create a catalogue of stellar parameters to use in exoplanet and galactic research. We could treat the parameters for each star independently, repeating the modelling procedure for each star in the sample. However, we know that stars belong to population distributions like

the IMF. In some cases, these population distributions are not well understood and assuming one exactly can introduce bias into our model. Instead, we could let the data inform such population-level priors.

Hierarchical (or multi-level) Bayesian models parametrise population-level prior distributions on individual stellar parameters. The parameters of these population distributions (or *hyperparameters*) should themselves have prior distributions as per the Bayesian formalisation. The hierarchical aspect comes from the distinction between hyperparameters, which take one value across a population, and parameters which vary from star-to-star. In the next section, we demonstrate a simple hierarchical model example. Then, we discuss a few parameters which could be given the hierarchical treatment in Section 2.3.

2.2 Stellar Distances in an Open Cluster Analogue

In this section, we demonstrate a hierarchical Bayesian model (HBM) with an analogy to measuring distances to stars in an open cluster. This example is similar to Leistedt and Hogg (2017), who present a hierarchical model of the colour-magnitude diagram to improve distances from *Gaia* (Gaia Collaboration et al., 2016). However, instead of considering the population distributions over magnitude and colour, we build a hierarchy over the stellar distances.

We created a dataset analogous to an open cluster of $N_{\text{stars}} = 20$ stars. We gave each i-th star a dimensionless distance from the observer (d_i) drawn randomly from a normal distribution with a mean of 10 and standard deviation of 0.1. Then, we converted these to dimensionless parallax using the relation $\varpi = 1/d$. We also gave each star an absolute visual magnitude (V_i) drawn randomly from a standard normal distribution. To get a quantity proportional to apparent magnitude (v_i) for each star, we used the relation $\mathrm{v}_i = \mathrm{V}_i + 5 \log_{10} d_i$. For the purpose of this example, we ignored additional real-world effects such as extinction and reddening.

We simulated noisy observations of ϖ_i and v_i by randomly drawing from a normal distribution centred on their true values with standard deviations of $\sigma_{\mathrm{v},i} = 0.1$ and $\sigma_{\varpi,i} = 0.01$ respectively. For real-world context, the uncertainties on ϖ from *Gaia* Data Release 3 (DR3;

Gaia Collaboration et al., 2022) are approximately 0.02 mas for stars with *Gaia* G-band magnitudes of less than 15. Therefore, if our choice of $\sigma_{\varpi,i}$ was representative of *Gaia* uncertainties, then the distance to our stellar cluster would be ~ 5 kpc. An example open cluster at this distance is NGC 6791, at 4 kpc (Brogaard et al., 2011). Typical cluster sizes are a few parsecs. On the same scale, the distance spread in our cluster analogue is an order of magnitude larger. However, we chose to exaggerate the distance spread to make it easier to measure and are not making a case for applying this example to the real-world.

In Section 2.2.1, we describe a simple Bayesian model for determining distances and absolute magnitudes of stars in the cluster. Then, we define the HBM in Section 2.2.2 which incorporates a population-level distribution over distance. We outline our method for inferring model parameters in Section 2.2.3 and then compare results from the models in Section 2.2.4. Finally, we explore how the HBM scales with the number of stars observed in Section 2.2.5.

2.2.1 Simple Model

We started with a simple model which treats each star independently. Using Bayes' theorem, we write the posterior probability density of the model parameters, d_i, V_i given the observed parameters ϖ_i, v_i as,

$$p(d_i, \mathrm{V}_i \mid \varpi_i, \mathrm{v}_i) \propto p(\varpi_i, \mathrm{v}_i \mid d_i, \mathrm{V}_i)\, p(d_i, \mathrm{V}_i). \tag{2.1}$$

The likelihood is given by $p(\varpi_i, \mathrm{v}_i \mid d_i, \mathrm{V}_i)$ and the prior probability of the parameters is $p(d_i, \mathrm{V}_i)$.

We modelled the likelihood as the product of normal distributions, assuming ϖ_i and v_i were observed independently,

$$p(\varpi_i, \mathrm{v}_i \mid d_i, \mathrm{V}_i) = \mathcal{N}(\varpi_i \mid d_i^{-1}, \sigma^2_{\varpi,i})\, \mathcal{N}(\mathrm{v}_i \mid \mathrm{V}_i + 5\log_{10} d_i\,, \sigma^2_{\mathrm{v},i}), \tag{2.2}$$

where $\mathcal{N}(x \mid \mu, \sigma^2)$ is a normal distribution over x with a mean of μ and variance of σ^2.

We assumed stars in the cluster were equally likely to be at a distance from 0 and 20.

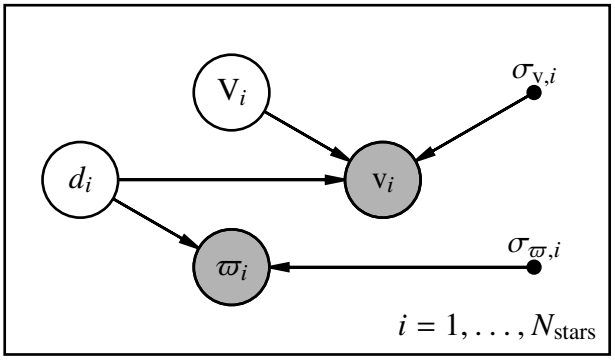


Figure 2.1: Probabilistic graphical model for the simple (non-hierarchical) model. Parameters are given by circular nodes and connected by arrows showing the direction of dependency. Observed parameters are shaded, and fixed parameters are given by filled dots. The box represents a set of parameters belonging to the i-th star. *This diagram was made using the* Python *package* `daft` *(Foreman-Mackey et al., 2021).*

Moreover, we assumed the absolute magnitudes were normally distributed centred on 0 and scaled by 10. Therefore, the prior probability of the model parameters was,

$$p(d_i, \mathrm{V}_i) = \mathcal{U}(d_i \mid 0, 20)\, \mathcal{N}(\mathrm{V}_i \mid 0, 100), \tag{2.3}$$

where $\mathcal{U}(x \mid a, b)$ is a uniform distribution over x from a to b. We chose wide, weakly-informative priors here because the focus of this example is on the hierarchical component introduced in the next section. However, the priors should be chosen more carefully in a real-world application.

In Figure 2.1, we show a probabilistic graphical model of the simple model. This shows the connections between parameters in the model. There is no hierarchy in this model because each parameter exists within the box, meaning no parameters are shared between stars. However, we know that the stellar distances are correlated. For example, their distances should be tightly related; it is unlikely that one star in the cluster is at a distance of 5 and another is at a distance of 15. If we can exploit this expectation, we could improve our prior and thus reduce our uncertainty in d_i and V_i.

2.2.2 Hierarchical Model

In this section, we present an HBM which includes expected correlation between distances to the stars in this open cluster analogue. We assumed the stars are all members of the same open cluster. Therefore, their distances can be modelled by some tight distribution. For this example, we assumed that each distance is drawn from a normal distribution characterised by hyperparameters μ_d and σ_d,

$$d_i \sim \mathcal{N}(\mu_d, \sigma_d^2). \tag{2.4}$$

Each stellar distance, d_i, is no longer treated independently by the model. Hence, we modified the posterior probability distribution to account for this correlation,

$$p(\mu_d, \sigma_d, \boldsymbol{d}, \mathbf{V} \mid \boldsymbol{\varpi}, \mathbf{v}) \propto p(\boldsymbol{\varpi}, \mathbf{v} \mid \boldsymbol{d}, \mathbf{V})\, p(\boldsymbol{d} \mid \mu_d, \sigma_d)\, p(\mu_d, \sigma_d, \mathbf{V}), \tag{2.5}$$

where $p(\boldsymbol{\varpi}, \mathbf{v} \mid \boldsymbol{d}, \mathbf{V})$ is the likelihood — a product over Equation 2.2,

$$p(\boldsymbol{\varpi}, \mathbf{v} \mid \boldsymbol{d}, \mathbf{V}) = \prod_{i=1}^{N_{\text{stars}}} p(\varpi_i, \mathrm{v}_i \mid d_i, \mathrm{V}_i). \tag{2.6}$$

Our posterior now depends on all stars in the population, so we used bold symbols to represent the set of individual stellar parameters. For example, $\boldsymbol{d} \equiv d_1, \ldots, d_{N_{\text{stars}}}$.

We modelled d_i as a random variable which depends on the hyperparameters. This is why our prior became the product of a conditional distribution, $p(\boldsymbol{d} \mid \mu_d, \sigma_d)$ and a prior on the remaining independent parameters, $p(\mu_d, \sigma_d, \mathbf{V})$. We write the conditional distribution for distance as a product of independent draws from the population-prior,

$$p(\boldsymbol{d} \mid \mu_d, \sigma_d) = \prod_{i=1}^{N_{\text{stars}}} \mathcal{N}(d_i \mid \mu_d, \sigma_d^2), \tag{2.7}$$

where the prior probability distribution for μ_d and σ_d is,

$$p(\mu_d, \sigma_d) = \mathcal{U}(\mu_d \mid 0, 20)\, \mathcal{N}(\ln \sigma_d \mid -\ln 10, 1). \tag{2.8}$$

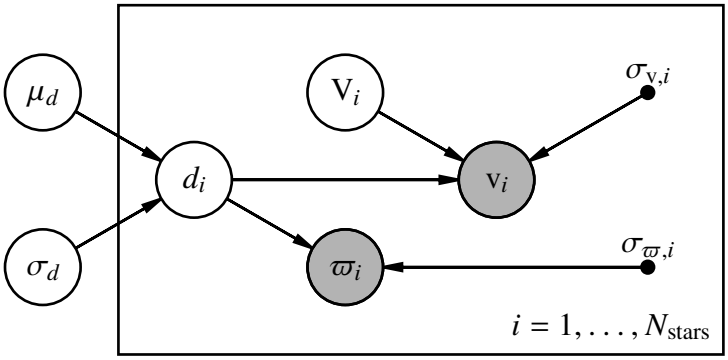


Figure 2.2: Probabilistic graphical model extension of Figure 2.1 but for the HBM. The hyperparameters exist outside the box to show that they are the same across the population of stars.

We chose to use a log-normal prior for σ_d to ensure that the parameter is positive. The standard deviation of the prior on σ_d approximately corresponds to a fractional uncertainty of 1. Finally, we inherited the prior on V_i from the simple model,

$$p(\mathbf{V}) = \prod_{i=1}^{N_{\mathrm{stars}}} \mathcal{N}(\mathrm{V}_i \mid 0, 100). \tag{2.9}$$

We show the probabilistic graphical model of the HBM in Figure 2.2. Here, we see how all individual stellar parameters depend on μ_d and σ_d. This is a simple extension of Figure 2.1, but with parameters existing outside the box to illustrate the model hierarchy. We can imagine extending this framework to multiple levels, or adding additional hyperparameters.

2.2.3 Inferring the Model Parameters

To infer the model parameters, we need to calculate the marginalised posterior distributions for each parameter. We could obtain these analytically by integrating the full posterior distribution over all model parameters except for the parameter of interest. Alternatively, we can approximate the marginalised posterior using a Markov Chain Monte Carlo (MCMC) sampling algorithm. We chose the latter approach because it is scalable to more complicated models where the marginalised posterior is not analytically solvable.

To sample from the posterior distribution for both models, we used the No U-Turn Sampler (NUTS; Hoffman and Gelman, 2014) as implemented in the PYTHON package `numpyro` (Bingham et al., 2019; Phan, Pradhan, and Jankowiak, 2019). The NUTS algorithm was chosen for its efficient exploration of parameter space when compared with other MCMC methods. We ran the sampler for 1000 steps following 500 warm-up steps (used to adapt the sampling procedure) and repeated for 10 MCMC chains. This resulted in 10 000 posterior samples obtained in few seconds on a modern laptop.

2.2.4 Comparing the Models

We compared the difference between model parameters and their true values in Figure 2.3. The HBM showed clear improvement over the simple model in all parameters except for apparent magnitude. Pooling the distances reduced their standard deviations by about one third and converged on their true values. This had the effect of reducing the observational noise on ϖ, as we can see the posterior predictions for parallax also approached their true values. Since apparent magnitude depended on both d and V, better distances also improved the absolute magnitudes.

In Figure 2.4, we compared the joint and marginalised posteriors of the distance to a few of the stars. The HBM found distances to higher precision, with typical standard deviations of $s_{d,i} \approx 0.35$ compared to $s_{d,i} \approx 1.2$ from the simple model. We also noticed a one-to-one correlation between distances present in the HBM but not in the simple model. This was a result of the correlation introduction by the population-prior on d_i. For each distance, we saw how a lower value in one corresponded to a lower value for the others. This represents our belief that cluster members should share a similar distance. On the other hand, the simple model suggested it was, just as likely to find two stars at distances of around 9 and 12 as it would for both to be at a distance of 11.

One consequence of the HBM was that it parametrised the population mean (μ_d) and standard deviation (σ_d) of the distances to stars in the cluster separately from the observed noise. For the simple model, we estimated μ_d and σ_d by taking the sample mean and standard

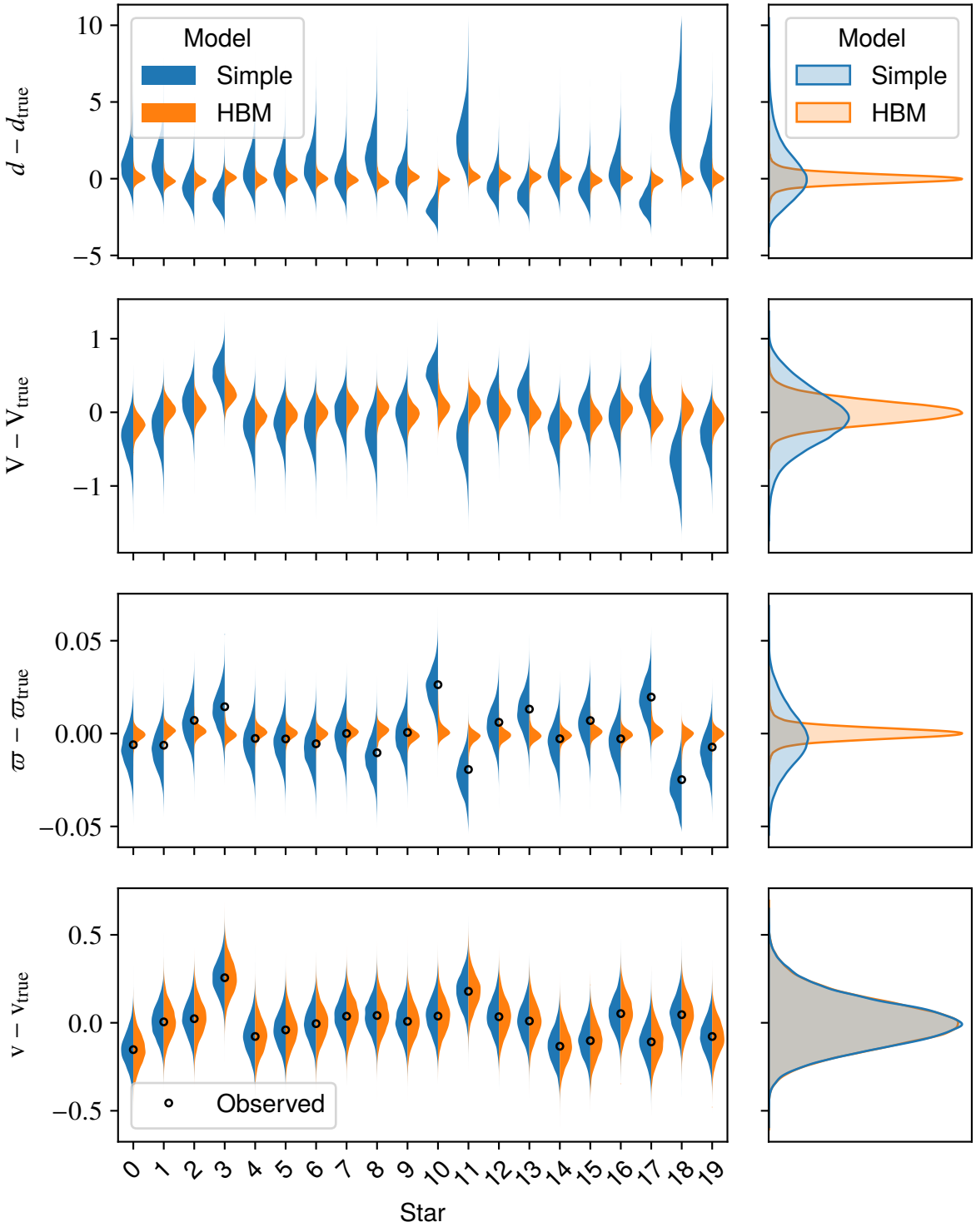


Figure 2.3: Plots comparing results from the simple model (*blue*) to the HBM (*orange*). *Left column:* Split violin-plots showing the difference between marginalised posterior distributions for the parameters and their true values. *Right column:* Kernel density estimates of the parameter-truth differences over all N_{stars}. Observed values of the parameters are given by empty black circles.

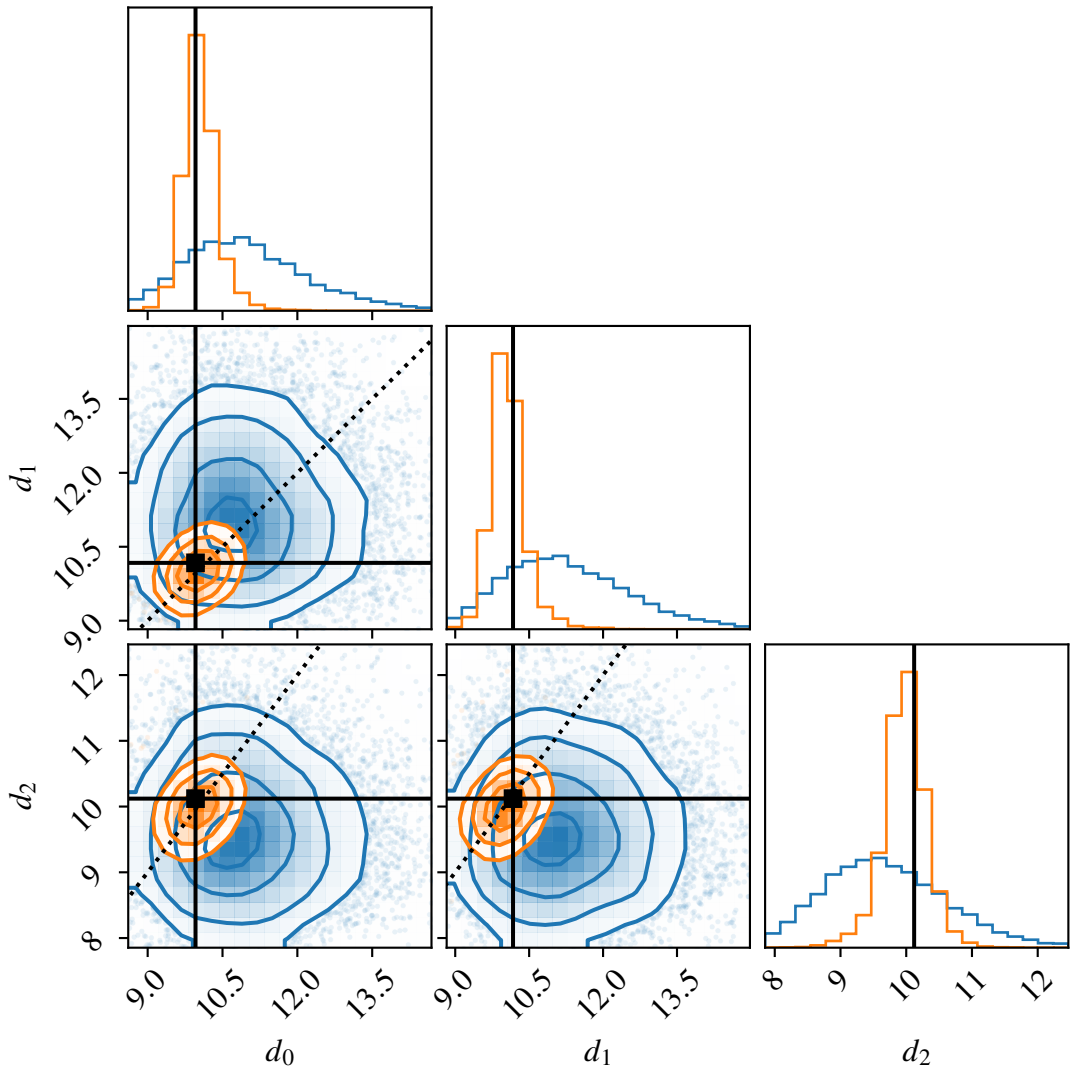


Figure 2.4: A corner plot showing the marginalised and joint posterior distributions of the distance to the first three stars. The simple model is shown in *blue* and the HMB is shown in *orange*. The true values are given by the *solid black* lines. The *dotted black* lines show the one-to-one relationships between distances. The range of the distributions are limited to 98 per cent of the simple model's marginalised posteriors. *This figure and all corner plots hereafter were created with the* Python *package* `corner` *(Foreman-Mackey, 2016).*

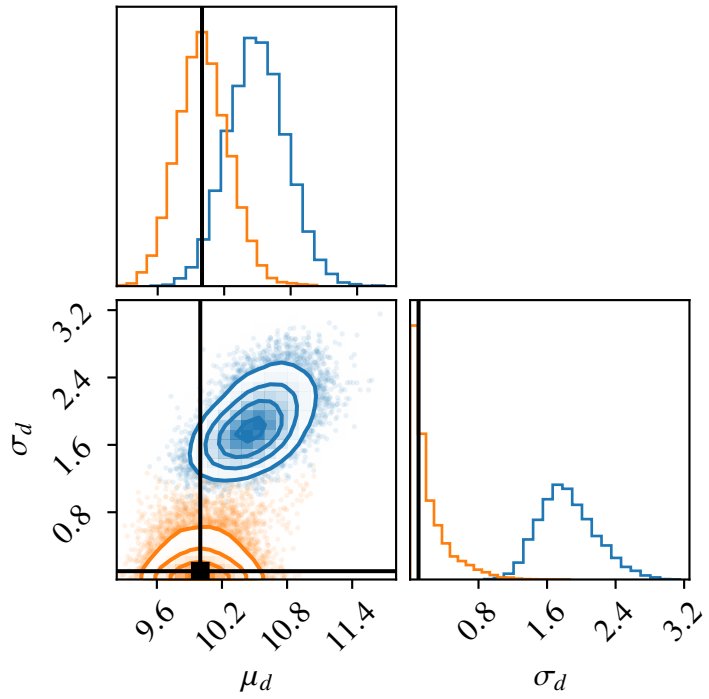


Figure 2.5: The joint and marginalised posterior distributions for the mean (μ_d) and standard deviation (σ_d) of stellar distances in the cluster. The simple model is shown in *blue* and the HMB is shown in *orange*. The true values are given by the *black* lines.

deviation of distances in the cluster for each posterior sample. We compared the resulting posterior distributions for μ_d and σ_d from the two models in Figure 2.5. The mean distance from the simple model was $\mu_d = 10.50 \pm 0.27$, whereas the HBM was more accurate with $\mu_d = 10.01 \pm 0.24$. The simple model overestimated the standard deviation of cluster distances with $\sigma_d = 1.815^{+0.360}_{-0.293}$, compared to the HBM's more accurate value of $\sigma_d = 0.131^{+0.280}_{-0.088}$. The simple model could not distinguish between the uncertainty on individual distances and the spread of the population.

2.2.5 Scaling with the Number of Stars

To test how the model scales with number of stars, we repeated the HBM method for different N_{stars}. We randomly generated true and observed parameters for 320 stars in the same way as described at the beginning of this section. The observable data were different to the last section, but drawn from the same true distributions. We then sampled from the HBM for the first 20, 80

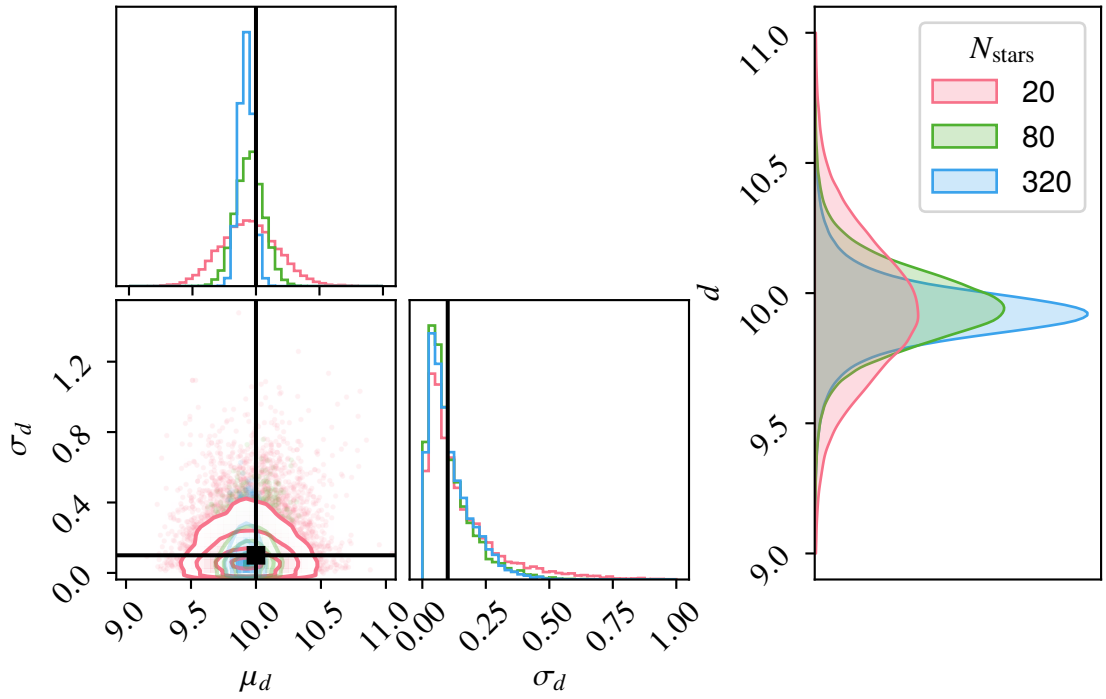


Figure 2.6: Joint and marginalised posterior distributions from the HBM for 20, 80 and 320 stars.

and 320 stars. These corresponded to ratios in $N_{\text{stars}}^{1/2}$ of $1 : 2 : 4$ respectively.

Figure 2.6 shows how the distance estimates improved with increasing number of stars observed, $N_{\text{stars}} = (20, 80, 320)$. Unsurprisingly, the standard deviation of μ_d decreased by a factor of $N_{\text{stars}}^{1/2}$ each time, with $s_{\mu_d} \approx (0.24, 0.12, 0.06)$. We also expected a similar reduction in uncertainty for individual distances, albeit limited by N_{stars} and the observational noise. Typical standard deviations for d_i were $s_{d,i} \approx (0.30, 0.17, 0.15)$. As expected, the distance uncertainties also halved from 20 to 80 stars. However, this effect lessened with 320 stars. This demonstrated the method is still limited by the error budget across all observed parameters. Despite this, the HBM proved to be a useful way of getting more information from a limited number of noisy observations.

2.3 Inferring Physical Properties of Stars

We have seen how a simple HBM can be used to improve the distances and absolute magnitudes of stars in a cluster analogue. This kind of model only represents part of the puzzle for determining ages, masses, and radii in a population of stars. In these Bayesian stellar models (discussed previously in Section 1.4) we have more parameters which could be considered hierarchically. In this section, we discuss a few possible population distributions which could form a hierarchical model of stellar physics.

Mass As mentioned in the introduction to this chapter, an IMF may be used as a prior on the mass of a star. If modelling stars with masses spanning a couple of orders of magnitude, it makes sense to draw their masses from an IMF. To make this hierarchical, we would parametrise the IMF as a conditional distribution which informs individual stellar masses. This way, the population of stars constrains the IMF while in-turn sharing information. In this thesis, we only consider stars in a narrow mass range (0.8 to 1.2 $M_\odot$) where the IMF does not change much.

Age We do not expect a star to be older than the universe. This makes the age of the universe a natural upper limit to a prior on stellar age. There are also populations of stars where we expect the age to be tightly related. For example, star systems like binaries and clusters are expected to have formed at a similar time. Similarly to the distances in Section 2.2, a hierarchical model in age would parametrise the mean and variance of ages in these systems and in-turn improve other correlated model parameters. This sort of analysis could be extended to find mixtures of age distributions in the galaxy which could indicate mergers (e.g. *Gaia*-Enceladus; Helmi et al., 2018).

Chemical Abundances Unlike mass and age, surface abundances ratios can be measured directly in stellar atmospheres with spectroscopy. However, not all abundances are easily measured. For example, helium ionises below the surface of the cool stars being studied in this work. Also, diffusion and settling of elements during stellar evolution means surface abundances

differ from formation to time of observation. Similarly to age, we can expect clusters of stars to share initial abundances. However, the distribution of abundances in the Milky Way may also be parametrised. As stars evolve they convert hydrogen into helium and heavier elements. Supernovae enrich the galaxy with these elements monotonically, which go on to constitute new stars. A hierarchical model could include this assumption to tie the abundance of helium to heavier elements.

Approximations of Stellar Physics A hierarchical model need not be limited to real physical parameters. In stellar modelling, we often approximate stellar physics with *pseudo*-physical parameters. For example, convective energy transfer is approximated by the mixing-length theory (MLT; Gough, 1977) parametrised by $\alpha_{\rm MLT}$. The value of $\alpha_{\rm MLT}$ which best models a star is likely to vary with other stellar parameters. The effect of $\alpha_{\rm MLT}$ on observables is small, making it difficult to constrain on a star-by-star basis. Hence, a hierarchical model would be able to condition individual values of $\alpha_{\rm MLT}$ on its distribution in the population. A distribution over pseudo-physical parameters is less intuitive, yet the HBM would allow us to build one empirically. In addition to the initial helium abundance, we model $\alpha_{\rm MLT}$ hierarchically in Chapter 3.

Chapter 3

HIERARCHICALLY MODELLING *KEPLER* DWARFS AND SUBGIANTS TO IMPROVE INFERENCE OF STELLAR PROPERTIES WITH ASTEROSEISMOLOGY

In this chapter, I present the entirety of Lyttle et al. (2021) with minor modification. I confirm that all writing is my own except for the description of stellar models and input physics in Section 3.3.1 which was written by Tanda Li[1]*. In this work, we created a hierarchical Bayesian model to pool information about the distribution of helium abundance (Y) and mixing-length-theory parameter ($\alpha_{\rm MLT}$) in a population of dwarf and subgiant solar-like oscillators.*

3.1 Introduction

The inference of stellar ages, masses, and radii has improved through the use of asteroseismology in recent years (e.g. see the review by Chaplin and Miglio, 2013). Measuring the oscillation modes in stars using photometric time series data, from missions such as *CoRoT* (Baglin et al., 2006), *Kepler* (Borucki et al., 2010), and *TESS* (Ricker et al., 2015) has given us new insights into the structure and evolution of stars (García and Ballot, 2019). Recent examples include a

[1]Institute for Frontiers in Astronomy and Astrophysics, Beijing Normal University, Beijing 102206, China

deeper understanding of stellar structure (Verma et al., 2017), chronology of a Milky Way merger (Chaplin et al., 2020), and classifying exoplanetary systems (Huber et al., 2019; Tayar et al., 2020). Several studies used grids of stellar models with constraints from asteroseismology to produce catalogues of precise stellar parameters (Pinsonneault et al., 2014; Silva Aguirre et al., 2017). However, with increasing precision on fundamental parameters inferred from stellar models with asteroseismology, we should take extra care to ensure that we are accounting for uncertainty in our choice of stellar physics.

Typically, stars are modelled on a star-by-star basis with estimates of stellar properties including some assumptions handled by simple scaling relations and solar calibrations. For instance, a helium (Y) to heavy-element (Z) enrichment ratio ($\Delta Y/\Delta Z$) and mixing-length theory parameter ($\alpha_{\rm MLT}$) are often assumed. However, there has been little effort to account for the population distribution of such quantities. Assuming values for Y and $\alpha_{\rm MLT}$ can result in inaccurate inference and systematics on the inferred stellar parameters (Valle et al., 2015). Independently modelling Y and $\alpha_{\rm MLT}$ can also be computationally demanding and requires high-precision observations in order to return meaningful stellar properties.

In this work, we apply a new method to determine stellar properties for a sample of *Kepler* dwarfs and subgiants using a hierarchical Bayesian model (HBM). With an HBM, we introduce population-level distributions for Y and $\alpha_{\rm MLT}$ to encode prior information throughout the sample. We will show that when an HBM is used, we can increase the precision of inferred masses, radii, and ages despite increasing the number of free parameters.

The use of HBMs has been demonstrated in other areas of stellar astrophysics to reduce individual parameter uncertainties by encoding prior information about the distribution of the parameter in a population. For example, HBMs have been used with data from *Gaia*, improving distance measures (Leistedt and Hogg, 2017; Anderson et al., 2018) and calibrating the red clump as a standard candle (Hawkins et al., 2017; Chan and Bovy, 2020) using asteroseismology (Hall et al., 2019). In other instances, HBMs have been used to infer binary-star and exoplanet eccentricities (Hogg, Myers, and Bovy, 2010), obliquity of transit systems (Morton and Winn, 2014), stellar inclination with asteroseismology (Campante et al., 2016; Kuszlewicz et al., 2019),

and the age-metallicity relation of the solar neighbourhood (Feuillet et al., 2016).

To describe the distribution of Y in this work, we assume a linear helium enrichment law characterised by freely varied population parameters: the gradient given by $\Delta Y/\Delta Z$, the primordial helium abundance at $Z = 0$ (Y_P) and an intrinsic spread in helium (σ_Y). There have been many studies to determine a linear enrichment law, from modelling eclipsing binaries (Ribas et al., 2000) to spectroscopy of galactic H II regions (Balser, 2006). Values of $\Delta Y/\Delta Z$ have also been determined for samples of main sequence (MS) stars (Casagrande et al., 2007), open clusters (Brogaard et al., 2012), and more recently with asteroseismology using individual oscillation frequencies (Silva Aguirre et al., 2017) and the glitch due to the second helium ionisation zone (Verma et al., 2019). In these studies the value of the enrichment ratio was typically inferred to be $1 \lesssim \Delta Y/\Delta Z \lesssim 3$. However, helium enrichment is unlikely to be exactly linear with metallicity and may depend on other chemical abundances (West and Heger, 2013) or the location of the star in the Milky Way (Frebel, 2010). Therefore, we account for some deviation from the linear law by introducing σ_Y. Moreover, our method may be adapted to different helium enrichment priors in future work.

The widely used mixing-length theory (MLT) of convection, parametrised by $\alpha_{\rm MLT}$, has been tested throughout the Hertzsprung-Russell (HR) diagram with 3D hydrodynamical simulations (Trampedach et al., 2014; Magic, Weiss, and Asplund, 2015) and asteroseismology (Tayar et al., 2017; Li et al., 2018; Viani et al., 2018) with values in the range $0.8 \lesssim \alpha_{\rm MLT}/\alpha_{\rm MLT,\odot} \lesssim 1.2$ where $\alpha_{\rm MLT,\odot}$ is the value calibrated to the Sun. However, in many grids of stellar models, a constant value calibrated to reproduce the Sun is assumed. This can lead to systematic uncertainties in stellar ages due to the effects of variable mixing depending on the mixing length. The MLT approximates convective mixing, and thus we expect the value of $\alpha_{\rm MLT}$ to vary from star-to-star due to various effects, from changes in chemical composition to other sources of mixing described poorly by MLT. The investigation of more complex $\alpha_{\rm MLT}$ distributions is beyond the scope of this work. Instead, we experiment with two prior assumptions for $\alpha_{\rm MLT}$ in the population. The first assumes $\alpha_{\rm MLT}$ is normally distributed in our sample with a spread (σ_α) to account for the aforementioned variation in $\alpha_{\rm MLT}$. The second assumes $\alpha_{\rm MLT}$ is constant

throughout the sample.

Our HBM requires a way to map from the stellar initial (or bulk) properties to predict observables. We can achieve this with a large grid of stellar evolutionary models. However, a discrete grid can produce inaccurate posterior distributions, limited to the grid resolution. Increasing the resolution is computationally demanding, especially when scaling to higher input dimensions. One solution is to interpolate the stellar models, as is common in the isochrone-fitting method (see e.g. Berger et al., 2020). However, interpolation can become computationally expensive at high input dimensions and grid size, and evaluating the likelihood using modern Bayesian sampling techniques is slow. Therefore, we use machine learning to map stellar model inputs to observables to provide a fast way to sample the HBM. In this work, we train an artificial neural network (ANN) on a large grid of stellar models. There have been similar applications of ANNs in asteroseismology (Bellinger et al., 2016; Verma et al., 2016; Hon, Stello, and Yu, 2017, 2018; Hendriks and Aerts, 2019) but not yet in the context of an HBM. Using the machine learning speed-up, we demonstrate a scalable method for obtaining fundamental stellar parameters.

In Section 3.2, we describe the data for the sample of 81 *Kepler* dwarfs and subgiants studied in this work. We then present the methods in Section 3.3 for which we produced a large grid of stellar evolutionary models to use as training data for an ANN. We use the ANN as an emulator in a series of statistical models described in Section 3.3.2 to model the sample and present our results in Section 3.4. Finally, in Section 3.5 we discuss the results from each model and compare them to results for the sample of stars in the literature.

3.2 Data

For this study, we selected the sample of 415 stars from the first APOKASC catalogue of dwarfs and subgiants (Serenelli et al., 2017, hereafter S17). This sample provides an extensive set of dwarfs and subgiant stars with asteroseismic detections observed by the *Kepler* mission. S17 used grid-based modelling to determine the ages (τ), masses (M), radii (R), and surface gravity

($\log g$) of stars in the sample, using global asteroseismic parameters, effective temperature ($T_{\rm eff}$), and metallicity ([M/H]) as inputs.

Using five independent pipelines, S17 determined values for global asteroseismic parameters — the large frequency separation $\Delta\nu$ and the frequency at maximum power, $\nu_{\rm max}$ with median uncertainties of 1.7 per cent and 4 per cent respectively. We chose to adopt the $\Delta\nu$ determined in their work as inputs for our method. They also used [M/H] published in Data Release 13 (DR13; Albareti et al., 2017) of the APOGEE stellar abundances pipeline (ASPCAP; García Pérez et al., 2016) with uncertainties of 0.1 dex. For their preferred set of results, they adopted $T_{\rm eff}$ from the Sloan Digital Sky Survey (SDSS) *griz*-band photometry (Pinsonneault et al., 2012) with a median uncertainty of 70 K.

We removed more evolved stars from the APOKASC sample by cutting those with $\log g < 3.8$ dex. We then kept stars within 1-σ of $-0.5 <$ [M/H] < 0.5 dex to remove metal-poor stars. Main sequence stars with $M \gtrsim 1.2\,{\rm M}_\odot$ are understood to have a convective, hydrogen-burning core, with some dependence on the chemical composition and choice of stellar physics (Appourchaux et al., 2015). Stellar models with a convective core require the treatment of extra stellar physics such as overshooting, which is beyond the scope of this work. Therefore, we keep only stars with masses within 1-σ of 0.8 to $1.2\,{\rm M}_\odot$ from the preferred set of results of S17.

Following cuts to the sample, we adopted updated ASPCAP spectroscopic metallicities, [M/H], from Data Release 14 (DR14; Blanton et al., 2017) which had a median uncertainty of 0.07 dex. We also chose to adopt $T_{\rm eff}$ from the same catalogue to be internally consistent. We note that our chosen effective temperature scale is offset from the photometric temperature scale of S17 by approximately -170 K, with a dispersion of ~ 120 K corresponding to typical uncertainties on the individual temperatures. The median uncertainty in our adopted ASPCAP $T_{\rm eff}$ was 125 K.

To provide a means of calculating luminosities, we used *Gaia* Data Release 2 (DR2) parallaxes (Gaia Collaboration et al., 2016, 2018). We cross-matched the remaining sample with the DR2 catalogue, taking the nearest neighbours within a 4 arcsec radius. Although DR2 parallaxes have improved upon the DR1 values at the time of S17, there was still evidence for a zero-point

Table 3.1: The observables and their respective uncertainties for 5 stars in the sample of 81 stars. The whole table is available as supplementary material for the original article (Lyttle et al., 2021).

Name	$T_{\rm eff}$ (K)	L ($L_\odot$)	$\Delta\nu$ (μHz)	$[{\rm M/H}]_{\rm surf}$ (dex)
KIC10079226	5929 ± 125	1.571 ± 0.049	116.04 ± 0.73	0.159 ± 0.074
KIC10215584	5667 ± 119	1.637 ± 0.063	115.16 ± 2.83	0.043 ± 0.069
KIC10319352	5456 ± 107	1.848 ± 0.056	78.75 ± 1.73	0.265 ± 0.065
KIC10322381	6147 ± 149	2.445 ± 0.079	86.64 ± 6.57	−0.317 ± 0.079
KIC10417911	5628 ± 110	3.408 ± 0.115	56.14 ± 2.10	0.336 ± 0.068

offset (Lindegren et al., 2018). We adopted a global offset of 0.05 mas, in the sense that DR2 parallaxes were underestimated, representative of values obtained in the literature using *Kepler* field stars (see e.g. Hall et al., 2019; Zinn et al., 2019a) and through other methods (see e.g. Riess et al., 2018; Chan and Bovy, 2020). We then cross-matched our sample with the Two-Micron All Sky Survey (2MASS; Skrutskie et al., 2006) to obtain K_S-band (2.16 μm) photometry.

We determined luminosities, L for the sample using the direct method of ISOCLASSIFY (Huber et al., 2017; Berger et al., 2020) with K_S-band photometry, *Gaia* DR2 parallaxes, ASPCAP [M/H] and $T_{\rm eff}$, and asteroseismic $\log g$ as inputs. This involved computing absolute K_S-band magnitudes using the *Gaia* DR2 parallaxes and extinctions determined by the 3D galactic reddening maps of Green et al. (2018). We determined absolute bolometric magnitudes by interpolating the MIST bolometric correction tables as a function of $T_{\rm eff}$, $\log g$ and [M/H] (Choi et al., 2016; Dotter, 2016). We adopted the uncertainty of 0.02 mag assumed by ISOCLASSIFY for both the extinctions and bolometric corrections, corresponding to typical systematic errors in extinction maps and bolometric fluxes (e.g. Zinn et al., 2019b; Tayar et al., 2020). We obtained luminosities for the sample with a median uncertainty of 3.4 per cent.

The final sample comprised 81 stars for which we had complete data for $T_{\rm eff}$, [M/H], $\Delta\nu$, and L to use as inputs for our stellar modelling method — see Table 3.1. In Fig. 3.1, we show the HR diagram for the sample plot in context with a series of stellar evolutionary tracks at solar metallicity.

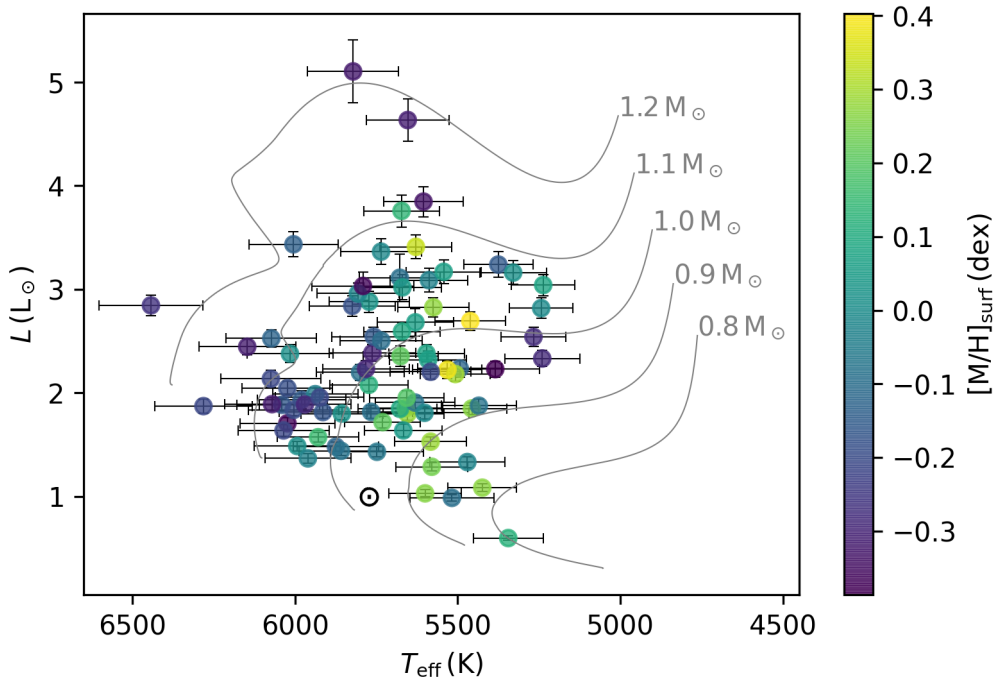


Figure 3.1: The luminosity, L against effective temperature, $T_{\rm eff}$ of the sample of 81 *Kepler* dwarfs and subgiants studied in this work. Each stars is coloured according to metallicity. The grey lines depict evolutionary tracks with $[{\rm M/H}]_{\rm init} = 0.0$ dex, $Y_{\rm init} = 0.28$ and $\alpha_{\rm MLT} = 1.9$ for different stellar masses. The current position of the Sun is shown by the $\odot$ symbol.

3.3 Methods

Firstly, we used a stellar evolutionary code to compute a grid of models to predict observable quantities (see Section 3.3.1). Subsequently, we trained an ANN on the grid of stellar models to map input parameters to output observables (see Appendix A.1 for further details). We then constructed three Bayesian models in Section 3.3.2 which each sampled the trained ANN to estimate stellar fundamental parameters as described in Section 3.3.4. Evaluation of the ANN gradient is required during training. Consequently, estimating the gradient of the model likelihood is fast and simple when the observables are generated by an ANN. Hence, we open up the possibility of using a Hamiltonian Monte Carlo (HMC) algorithm — for example, using the No-U-Turn Sampler (NUTS; Hoffman and Gelman, 2014) — which requires the gradient to sample the model posterior. Once we had tested the accuracy of the model on a sample of synthetic stars, we evaluated each model on the subset of the APOKASC catalogue selected in

Table 3.2: Stellar model grid parameters for train and test datasets. $N_{\rm track}$ are the numbers of stellar evolutionary tracks for each dimension of the grid, multiplied to a total of 17 220 tracks.

Stellar model grid			
Input Parameter	Range	Increment	$N_{\rm track}$
M ($M_\odot$)	0.80 – 1.20	0.01	41
$[{\rm M/H}]_{\rm init}$ (dex)	-0.5 – 0.2/0.25 – 0.5	0.1/0.05	14
$Y_{\rm init}$	0.24 – 0.32	0.02	5
$\alpha_{\rm mlt}$	1.5 – 2.5	0.2	6
Total			17 220

Section 3.2.

3.3.1 Grid of Stellar Models

We built a stellar model grid to use in training the ANN. The grid includes four independent model inputs: stellar mass (M), initial helium fraction ($Y_{\rm init}$), initial metallicity ($[{\rm M/H}]_{\rm init}$), and the mixing-length parameter ($\alpha_{\rm MLT}$). Ranges and grid steps of the four model inputs are summarised in Table 3.2. We increased the resolution at higher [M/H] to give a more consistent resolution in $Z_{\rm init}$ which we used as an ANN input in Appendix A.1. We computed each stellar evolutionary track from the Hayashi line to the base of red-giant branch defined here by $\log g$ = 3.6 dex. This amounted to approximately 1000 stellar models per track, totalling $\sim 10^7$ stellar models in the grid. We also computed an additional set of $\sim 10^6$ stellar models, with input values chosen randomly within the range of the regular grid, to use when testing the accuracy of the ANN.

Stellar Models and Input Physics

We used Modules for Experiments in Stellar Astrophysics (MESA, version 12115) to establish a grid of stellar models. MESA is an open-source stellar evolution package which is undergoing active development. Descriptions of input physics and numerical methods can be found in Paxton et al. (2011, 2013, 2015, 2018, 2019). We adopted the solar chemical mixture, $(Z/X)_\odot$ = 0.0181, provided by Asplund et al. (2009). The initial chemical composition was calculated

by:

$$\log(Z_{\rm init}/X_{\rm init}) = \log(Z/X)_{\odot} + [{\rm M/H}]_{\rm init}. \tag{3.1}$$

We used the MESA $\rho - T$ tables based on the 2005 update of OPAL EOS tables (Rogers and Nayfonov, 2002) and OPAL opacity supplemented by low-temperature opacity (Ferguson et al., 2005). The grey Eddington $T - \tau$ relation is used to determine boundary conditions for modelling the atmosphere. The MLT of convection was implemented, where $\alpha_{\rm MLT} = \ell_{\rm MLT}/H_p$ is the mixing-length parameter. We also applied the MESA predictive mixing scheme (Paxton et al., 2018, 2019) in the model computation.

Atomic diffusion of helium and heavy elements was also taken into account. MESA calculates particle diffusion and gravitational settling by solving Burger's equations using the method and diffusion coefficients of Thoul, Bahcall, and Loeb (1994). We considered eight elements (^{1}H, ^{3}He, ^{4}He, ^{12}C, ^{14}N, ^{16}O, ^{20}Ne, and ^{24}Mg) for diffusion calculations, and had the charge calculated by the MESA ionisation module, which estimates the typical ionic charge as a function of T, ρ, and free electrons per nucleon from Paquette et al. (1986).

Oscillation Models and Asteroseismic Quantities

Theoretical stellar oscillations were calculated with the GYRE code (version 5.1), which was developed by Townsend and Teitler (2013). We computed radial modes (for $\ell = 0$) for 42 radial orders by solving the adiabatic stellar pulsation equations with the structural models generated by MESA. We determined the asteroseismic large separation ($\Delta\nu$) for each model with theoretical radial modes to avoid the systematic offset of the scaling relation. We derived $\Delta\nu$ with the approach given by White et al. (2011), which is a weighted least-squares fit to the radial frequencies as a function of n.

We chose to ignore the well-known, yet poorly-characterised impact of modelled oscillation mode inaccuracies in the near-surface region of the star (Kjeldsen, Bedding, and Christensen-Dalsgaard, 2008; Ball and Gizon, 2014; Sonoi et al., 2015). This typically presents only a small effect compared to observational uncertainties when considering the average large frequency spacing, $\Delta\nu$. Additionally, there may be further inaccuracies in the modelled $\Delta\nu$ because of

variations in p mode frequencies with stellar activity (Chaplin et al., 2007; García et al., 2010; Kiefer et al., 2017). Therefore, a thorough treatment of systematic uncertainties in $\Delta\nu$ is instead left to future work.

3.3.2 Statistical Models

We devised three Bayesian models, each with varying levels of parameter sharing (pooling) between stars in the population. Initially, we tested the models and demonstrated reduction of statistical uncertainties in the stellar fundamental parameters by analysing a random sample of 100 synthetic stars modelled using MESA. Then, we applied the models to the sample of stars in Table 3.1 (with and without solar data for two of the models) and compared the results with that of S17.

Our first model was equivalent to modelling each star individually and featured no pooling; henceforth, we refer to it as the no-pooled (NP) model. We then derived two hierarchical Bayesian models (HBMs) which use population-level parameters to describe their distribution in the sample. Both of these models partially-pooled helium using a linear enrichment law. We drew the initial helium fraction for each star from a normal distribution with a mean described by the enrichment law and standard deviation representing its spread. Similarly, we partially-pooled the MLT parameter, $\alpha_{\rm MLT}$ in one model, whereas we maximally-pooled $\alpha_{\rm MLT}$ in the other, such that it assumes the same value for the entire sample. Hence, we refer to the former as the partial-pooled (PP) model and the latter as the max-pooled (MP) model.

No-Pooled (NP) Model

Firstly, we constructed a model comprising independent parameters corresponding to the ANN inputs $\boldsymbol{\theta}_i = \{f_{{\rm evol},i}, M_i, \alpha_{{\rm MLT},i}, Y_i, Z_i\}$ for a given star, i. The parameter $f_{{\rm evol},i}$ describes the evolutionary stage of the i-th star (see Equation A.4 in the Appendix). Using Bayes' theorem, the *posterior* probability density function (PDF) of the model parameters given a set of observed data, $\boldsymbol{d}_i$ is,

$$p(\boldsymbol{\theta}_i|\boldsymbol{d}_i) \propto p(\boldsymbol{\theta}_i)\, p(\boldsymbol{d}_i|\boldsymbol{\theta}_i)\,, \tag{3.2}$$

where $p(\boldsymbol{\theta}_i)$ is the *prior* PDF of the model parameters and $p(\boldsymbol{d}_i|\boldsymbol{\theta}_i)$ is the *likelihood* of observing the data given the model.

We chose weakly-informative, bounded priors for the independent parameters, restricting them to their respective ranges in the ANN training data. Although the neural network is able to make predictions outside the training data range, these have not been tested and may be unreliable. Therefore, we used a beta distribution with $\alpha = \beta = 1.2$ as the prior PDF on the independent parameters, transformed such that the probability is null outside the chosen range,

$$p(\boldsymbol{\theta}_i) \propto \prod_{k=1}^{N_\theta} \mathcal{B}\left(\tilde{\theta}_{k,i}|1.2, 1.2\right), \tag{3.3}$$

where the beta distribution is defined as,

$$\mathcal{B}(x|\alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 u^{\alpha-1}(1-u)^{\beta-1}\mathrm{d}u}. \tag{3.4}$$

and,

$$\tilde{\theta}_{k,i} = \frac{\theta_{k,i} - \theta_{k,\min}}{\theta_{k,\max} - \theta_{k,\min}}, \tag{3.5}$$

is the transformed parameter where $\theta_{k,\min}$ and $\theta_{k,\max}$ are the upper and lower bounds for each parameter. See Fig. A.4 in the Appendix for an example of the beta distribution. We preferred this over a uniform distribution because our choice of MCMC sampler was less efficient close to the distribution bounds.

Using notation which represents a given random variable $x \sim q$ as equivalent to being drawn from a probability distribution $p(x) \propto q(x)$ where $q(x)$ is a non-normalised probability density

function, we may write the priors for $\boldsymbol{\theta}_i$ as,

$$
\begin{aligned}
f_{\mathrm{evol},i} &\sim 0.01 + 1.99 \cdot \mathcal{B}(1.2, 1.2), \\
M_i &\sim 0.8 + 0.4 \cdot \mathcal{B}(1.2, 1.2), \\
\alpha_{\mathrm{MLT},i} &\sim 1.5 + \mathcal{B}(1.2, 1.2), \\
Y_{\mathrm{init},i} &\sim 0.24 + 0.08 \cdot \mathcal{B}(1.2, 1.2), \\
Z_{\mathrm{init},i} &\sim 0.005 + 0.035 \cdot \mathcal{B}(1.2, 1.2),
\end{aligned}
$$

where each beta distribution is scaled to cover the boundaries of the grid of stellar models computed in Section 3.3.1.

We made predictions for each star using the trained ANN, $\{\log(\tau)_i, T_{\mathrm{eff},i}, R_i, \Delta\nu_i, [\mathrm{M/H}]_{\mathrm{surf},i}\} = \boldsymbol{f}_{\mathrm{ANN}}(\boldsymbol{\theta}_i)$, from which we derived the luminosity, L_i using the Stefan-Boltzmann law. Out of the model parameters, those which may be observed are denoted by $\boldsymbol{\mu}_i = \boldsymbol{f}(\boldsymbol{\theta}_i)$. Therefore, we write the likelihood that we observe any $\boldsymbol{d}_i$ with known uncertainty, $\boldsymbol{\sigma}_i$ given our model as,

$$
p(\boldsymbol{d}_i|\boldsymbol{\theta}_i) = \prod_{k=1}^{N_{\mathrm{obs}}} \frac{1}{\sigma_{k,i}\sqrt{2\pi}} \exp\left[-\frac{(d_{k,i} - \mu_{k,i})^2}{2\sigma_{k,i}^2}\right], \tag{3.6}
$$

where N_{obs} is the number of observed variables. We chose to use observed T_{eff}, L, $\Delta\nu$, and [M/H] collated for our sample as described in Section 3.2.

It follows that the posterior PDF for a population of N_{stars} stars for the NP model is,

$$
p(\boldsymbol{\Theta}|\boldsymbol{D}) = \prod_{i=1}^{N_{\mathrm{stars}}} p(\boldsymbol{\theta}_i|\boldsymbol{d}_i), \tag{3.7}
$$

where $\boldsymbol{\Theta}$ is the matrix of model parameters and $\boldsymbol{D}$ is the matrix of observables. A probabilistic graphical model (PGM) of the NP model can be seen inside the grey box of Fig. 3.2, without the arrow connecting Z_{init} to Y_{init}. We ignore the nodes outside the box because these correspond to the PP model described next.

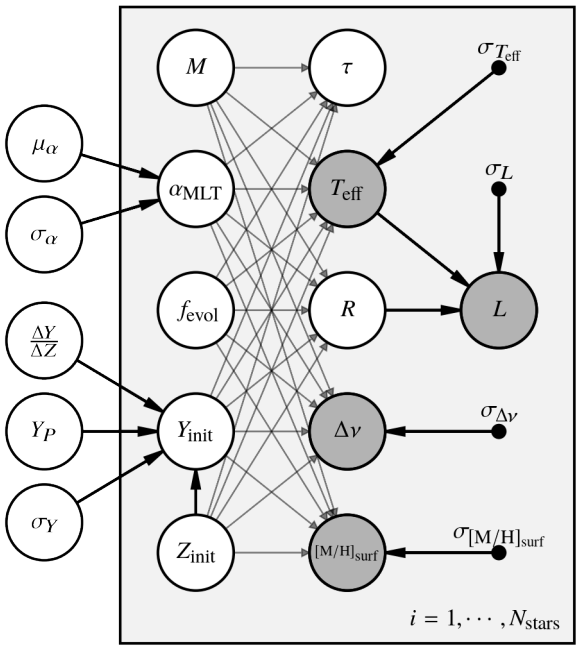


Figure 3.2: A probabilistic graphical model (PGM) of the partially-pooled (PP) hierarchical model. Nodes outside the grey rectangle represent the hyperparameters in the model. Nodes inside the grey rectangle represent individual stellar parameters. Dark grey nodes represent observables which each have their respective observational uncertainties given by the solid black nodes. The direction of the arrows represent the dependencies in the generative model. *This diagram was made using the* Python *package* `daft` *(Foreman-Mackey et al., 2021).*

Partial-Pooled (PP) Model

Sharing, or pooling parameters between stars in a population can improve the uncertainties on stellar fundamentals by encoding our prior knowledge of their distribution in a population. We constructed a hierarchical model, which builds upon the NP model by introducing population-level *hyperparameters*. Specifically, we chose to describe initial helium and $\alpha_{\rm MLT}$ by partially-pooling them.

We constructed the PP model such that each of the initial helium, $\boldsymbol{Y}_{\rm init}$ and MLT parameter, $\boldsymbol{\alpha}_{\rm MLT}$ are drawn from a common distribution characterised by the set of hyperparameters, $\boldsymbol{\phi}$. Thus, Bayes' theorem becomes,

$$p(\boldsymbol{\phi}, \boldsymbol{\Theta}|\boldsymbol{D}) \propto p(\boldsymbol{\phi})\, p(\boldsymbol{Y}_{\rm init}, \boldsymbol{\alpha}_{\rm MLT}|\boldsymbol{\phi})\, p(\boldsymbol{f}_{\rm evol}, \boldsymbol{M}, \boldsymbol{Z})\, p(\boldsymbol{D}|\boldsymbol{\Theta}), \tag{3.8}$$

where $\boldsymbol{\Theta}$ is the same as in the NP model, i.e. each object-level parameter, $\boldsymbol{\theta}_j = \{\theta_{j,i}\}_{i=1}^{N_{\rm stars}}$, and $\boldsymbol{\phi} = \{\Delta Y/\Delta Z, Y_P, \sigma_Y, \mu_\alpha, \sigma_\alpha\}$. The hyperparameters for $\boldsymbol{Y}_{\rm init}$ comprise the helium enrichment ratio ($\Delta Y/\Delta Z$), primordial helium abundance fraction (Y_P), and the spread in helium (σ_Y). The remaining hyperparameters for $\boldsymbol{\alpha}_{\rm MLT}$ comprise the mean (μ_α) and spread, (σ_α).

We assumed the initial helium and the mixing-length parameter are each drawn from a normal distribution characterised by a population mean and standard deviation. The probability of $\boldsymbol{Y}_{\rm init}$ and $\boldsymbol{\alpha}_{\rm MLT}$ given $\boldsymbol{\phi}$ is,

$$p(\boldsymbol{Y}_{\rm init}, \boldsymbol{\alpha}_{\rm MLT}|\boldsymbol{\phi}) = p(\boldsymbol{Y}_{\rm init}|\boldsymbol{\mu}_Y, \sigma_Y)\, p(\boldsymbol{\alpha}_{\rm MLT}|\mu_\alpha, \sigma_\alpha), \tag{3.9}$$

where $\boldsymbol{\mu}_Y$ and is the mean initial helium fraction as described by the linear helium enrichment law,

$$\boldsymbol{\mu}_Y = Y_P + \frac{\Delta Y}{\Delta Z}\boldsymbol{Z}_{\rm init}. \tag{3.10}$$

Therefore, we may write the prior PDF of initial helium given its population-level hyperparameters as,

$$p(\boldsymbol{Y}_{\rm init}|\boldsymbol{Z}_{\rm init}, \Delta Y/\Delta Z, Y_P, \sigma_Y) = \prod_{i=1}^{N_{\rm stars}} \mathcal{N}(Y_{{\rm init},i}|\mu_{Y,i}, \sigma_Y). \tag{3.11}$$

Similarly, for the second component of Equation 3.9, we chose to partially-pool $\alpha_{\rm MLT}$. We assume that convection in stars of a similar mass, evolutionary stage and area of the HR diagram may be approximated using a similar value of $\alpha_{\rm MLT}$, but the accuracy of the MLT may vary from star-to-star. Given the small range of our sample, any such variation will be absorbed by the spread parameter, σ_α. Therefore, we decided to describe the prior on $\boldsymbol{\alpha}_{\rm MLT}$ as,

$$p(\boldsymbol{\alpha}_{\rm MLT}|\mu_\alpha, \sigma_\alpha) = \prod_{i=1}^{N_{\rm stars}} \mathcal{N}(\alpha_{{\rm MLT},i}|\mu_\alpha, \sigma_\alpha). \tag{3.12}$$

We gave all the hyperparameters weakly informative priors, except for Y_P for which we adopt a recent measurement of the primordial helium abundance from big bang nucleosynthesis (BBN) as the mean (Pitrou et al., 2018), with a standard deviation representative of the range of values in the literature (Aver, Olive, and Skillman, 2015; A. Peimbert, M. Peimbert, and Luridiana, 2016;

Cooke and Fumagalli, 2018). Hence, we assumed priors on the hyperparameters as follows,

$$\begin{aligned}
\Delta Y/\Delta Z &\sim 4.0 \cdot \mathcal{B}(1.2, 1.2), \\
Y_P &\sim \mathcal{N}(0.247, 0.001), \\
\sigma_Y &\sim \ln \mathcal{N}(0.01, 1.0), \\
\mu_\alpha &\sim 1.5 + \mathcal{B}(1.2, 1.2), \\
\sigma_\alpha &\sim \ln \mathcal{N}(0.1, 1.0),
\end{aligned}$$

where, $x \sim \ln \mathcal{N}(m, \sigma)$ represents a random variable drawn from the log-normal distribution,

$$\ln \mathcal{N}(x|m, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left[-\frac{\ln(x/m)^2}{2\sigma^2}\right]. \tag{3.13}$$

We produced a PGM for the model, depicted in Fig. 3.2. The hyperparameters are shown outside the grey box containing the individual stellar parameters to represent the hierarchical aspect of the model.

Max-Pooled (MP) Model

We built another hierarchical model similar to the PP model except that $\alpha_{\rm MLT}$ is max-pooled (MP). In this model, we assumed that $\alpha_{\rm MLT}$ must be the same value for every star in the sample, but still allowed it to freely vary on a population-level. Thus, the hyperparameters are now, $\boldsymbol{\phi} = \{\Delta Y/\Delta Z, Y_P, \sigma_Y, \alpha_{\rm MLT}\}$. The posterior distribution of the model takes the same form as in Equation 3.8 except that the MLT parameter for the i-th star is,

$$\alpha_{{\rm MLT},i} = \alpha_{\rm MLT}, \tag{3.14}$$

where,

$$\alpha_{\rm MLT} \sim 1.5 + \mathcal{B}(1.2, 1.2), \tag{3.15}$$

Table 3.3: Solar input data. The references correspond to the central values and the uncertainties are chosen to either be representative of the ANN accuracy or the spread of values in the literature (see text for details).

Input	Value	Reference
M ($M_\odot$)	1.000 ± 0.001	—
τ (Gyr)	4.6 ± 0.1	Connelly et al. (2012)
$T_{\rm eff}$ (K)	5777 ± 20	Scott et al. (2015)
R ($R_\odot$)	1.000 ± 0.001	—
L ($L_\odot$)	1.00 ± 0.01	—
$\Delta\nu$ (μHz)	135.1 ± 0.2	Huber et al. (2011)
$[{\rm M/H}]_{\rm surf}$ (dex)	0.00 ± 0.01	Asplund et al. (2009)

chosen such that $\alpha_{\rm MLT}$ is confined to the boundaries of the grid of stellar models ($1.5 < \alpha_{\rm MLT} < 2.5$).

3.3.3 The Sun as a Star

Pooling parameters in an HBM allows us to use the Sun as a calibrator uniquely. Rather than calibrating our model physics to the Sun and then assuming the calibrated parameters across our sample, we can include the Sun as a part of the same population as our sample of stars. If we assume $Y_{\rm init}$ and $\alpha_{\rm MLT}$ for the Sun are a part of the same prior distribution as for the rest of the sample, then we can simply add solar observables to our model inputs.

For both the PP and MP models, we iterated with and without data for the Sun included in the population, referred to as PPS and MPS respectively. We adopted the solar data in Table 3.3 with uncertainties conservatively limited to the accuracy of the ANN for R, L and representative of variation in the literature for $T_{\rm eff}$. We also adopted $\Delta\nu = 135.1 \pm 0.2\,\mu$Hz with a central value from Huber et al. (2011) and a standard deviation representative of variations in measurements of the solar $\Delta\nu$ (Broomhall et al., 2011).

3.3.4 Sampling

We obtained results for each of the models described above by sampling their posterior distributions using a Markov chain Monte Carlo (MCMC) algorithm. In particular, we used the

NUTS algorithm implemented in TensorFlow Probability (TFP; Abadi et al., 2016; Dillon et al., 2017)[2]. For each model, we produced 20 000 samples split across 10 MCMC chains and computed summary statistics for the marginalised posteriors of each parameter in the model. We removed stars with problems during tuning using the Gelman-Rubin diagnostic ($\hat{r}$; Gelman and Rubin, 1992). We used results from each model once $\hat{r} < 1.04$ for all parameters, indicating good model convergence.

Initially, we created a random synthetic population of stars using MESA to test the ability of the method to recover stellar properties according to our choice of model physics and population priors. We tested the NP, PP, and MP models. Since our sample was fictitious, it would not have been appropriate to include real solar data. We summarise the results for the synthetic stellar parameters and hyperparameters in Appendix A.3. We found that the models were able to recover the true synthetic properties accurately, with increased precision when pooling parameters.

Once we had tested the method on synthetic stars, we obtained results for the sample of 81 dwarfs and subgiants described in Section 3.2. Here, we included the PPS and MPS to test the effects of adding the Sun as a star in our population. For the purpose of comparison, we fit the hyperparameters of the PP model ($\Delta Y/\Delta Z, Y_P, \sigma_Y, \mu_\alpha, \sigma_\alpha$) to the results from the NP model.

Since our initial sample was chosen based on masses from S17, we expected some stars to lie outside (or near the boundary) of the observational parameter space provided by our grid of stellar models. We used an initial run of the NP model to catch and remove these stars. During the initial run, we dropped 16 of the 81 stars from the sample. Of the removed stars, we found the posteriors in M for 6 skewed towards the prior upper mass limit of 1.2 $M_\odot$. The remaining 10 removed stars suffered poor convergence during sampling ($\hat{r} >> 1.04$) which could be because of poor step-size tuning and sampling at the prior boundary.

Out of the remaining 65 stars with results from the NP model, 2 stars were dropped from the PP model. A consequence of partially pooling parameters is that a population spread, σ allows for individual parameters to vary outside the prior range given to the population mean, μ. In this case, individual stellar $\alpha_{\rm MLT}$ was allowed to vary outside the range for which the ANN was

[2]We interacted with TFP using the now deprecated PyMC4 package, developed as a successor to PyMC3 (Salvatier, Wiecki, and Fonnesbeck, 2016).

valid if σ_α was large. The 2 removed stars happened to have high likelihoods outside the valid $\alpha_{\rm MLT}$ range. We found that removing the same 2 stars from the other models made negligible difference to the results, so we leave a solution to this problem to future work. Naturally, we did not see the same issue in the MP model, so we proceeded with modelling the same 65 stars as with the NP model.

3.4 Results

In this section, we present the results for each of the NP, PP, and MP models with the sample of 81 APOKASC dwarfs and subgiants as inputs. We also present the results for the PPS and MPS models which include the Sun as a star in the population. Firstly, we show the reduction in age, mass and radius uncertainty with the addition of pooling in Section 3.4.1. We then show the results for model hyperparameters in Section 3.4.2 where we infer the initial helium abundance and mixing-length parameter distribution in the sample.

3.4.1 Stellar Parameter Results

In Table 3.4, we present results for the 65 APOKASC stars from the NP model. Running the NP model with synthetic stars resulted in unreliable uncertainties (see Appendix A.3). This was because the boundary of the priors in $Y_{\rm init}$ and $\alpha_{\rm MLT}$ truncated the posterior distribution leading to underestimated uncertainties and skewing their posterior means towards the centre of their priors. Therefore, we present the NP results only for comparison purposes, but we exclude them from further discussion. In Tables 3.5 and 3.6 we present the results for the 63 stars from the PP and PPS model respectively. In Tables 3.7 and 3.8 we also present results for the 65 stars from the MP and MPS models respectively. We note that for the MP models, there is no column for $\alpha_{\rm MLT}$ because this parameter is the same across the population and hence is given in Section 3.4.2.

Table 3.4: The median of the marginalised posterior samples for selected parameters output by the NP model, with their respective upper and lower 68 per cent credible intervals. The full table is available as supplementary material in the original article (Lyttle et al., 2021).

Name	$f_{\rm evol}$	M ($\mathrm{M}_\odot$)	$\alpha_{\rm MLT}$	$Y_{\rm init}$	$Z_{\rm init}$	τ (Gyr)
KIC10079226	$0.44^{+0.16}_{-0.20}$	$1.14^{+0.04}_{-0.04}$	$2.07^{+0.26}_{-0.30}$	$0.28^{+0.02}_{-0.02}$	$0.021^{+0.003}_{-0.003}$	$2.5^{+1.2}_{-1.3}$
KIC10215584	$0.50^{+0.21}_{-0.21}$	$1.13^{+0.04}_{-0.05}$	$1.92^{+0.33}_{-0.26}$	$0.27^{+0.02}_{-0.02}$	$0.018^{+0.002}_{-0.002}$	$2.9^{+1.6}_{-1.3}$
KIC10319352	$1.51^{+0.15}_{-0.28}$	$1.09^{+0.05}_{-0.05}$	$1.87^{+0.33}_{-0.23}$	$0.28^{+0.03}_{-0.02}$	$0.028^{+0.004}_{-0.003}$	$9.6^{+1.7}_{-1.5}$
KIC10322381	$0.98^{+0.23}_{-0.28}$	$1.07^{+0.07}_{-0.07}$	$2.04^{+0.29}_{-0.31}$	$0.28^{+0.02}_{-0.02}$	$0.010^{+0.002}_{-0.002}$	$4.7^{+1.5}_{-1.7}$
KIC10732098	$1.60^{+0.14}_{-0.19}$	$1.12^{+0.04}_{-0.05}$	$1.86^{+0.33}_{-0.23}$	$0.28^{+0.02}_{-0.02}$	$0.017^{+0.002}_{-0.002}$	$6.7^{+0.8}_{-0.8}$

Table 3.5: The same as Table 3.4, but for the PP model.

Name	$f_{\rm evol}$	M ($\mathrm{M}_\odot$)	$\alpha_{\rm MLT}$	$Y_{\rm init}$	$Z_{\rm init}$	τ (Gyr)	$T_{\rm eff}$ (K)	R ($\mathrm{R}_\odot$)	$\Delta\nu$ (μHz)	$[\mathrm{M/H}]_{\rm surf}$ (dex)
KIC10079226	$0.22^{+0.10}_{-0.09}$	$1.16^{+0.02}_{-0.03}$	$1.75^{+0.11}_{-0.09}$	$0.28^{+0.01}_{-0.01}$	$0.020^{+0.003}_{-0.002}$	$1.2^{+0.6}_{-0.5}$	5962^{+44}_{-42}	$1.17^{+0.01}_{-0.01}$	$115.9^{+0.7}_{-0.7}$	$0.15^{+0.07}_{-0.07}$
KIC10215584	$0.37^{+0.15}_{-0.13}$	$1.14^{+0.03}_{-0.03}$	$1.74^{+0.10}_{-0.09}$	$0.27^{+0.01}_{-0.01}$	$0.018^{+0.002}_{-0.002}$	$2.1^{+1.0}_{-0.8}$	5941^{+57}_{-56}	$1.18^{+0.02}_{-0.02}$	$112.5^{+2.6}_{-2.7}$	$0.07^{+0.07}_{-0.07}$
KIC10319352	$1.41^{+0.11}_{-0.27}$	$1.08^{+0.03}_{-0.03}$	$1.73^{+0.10}_{-0.09}$	$0.29^{+0.02}_{-0.01}$	$0.028^{+0.004}_{-0.004}$	$8.6^{+1.1}_{-1.0}$	5512^{+45}_{-46}	$1.49^{+0.02}_{-0.02}$	$78.6^{+1.6}_{-1.6}$	$0.28^{+0.06}_{-0.07}$
KIC10322381	$0.78^{+0.23}_{-0.19}$	$1.14^{+0.03}_{-0.06}$	$1.75^{+0.10}_{-0.09}$	$0.26^{+0.01}_{-0.01}$	$0.011^{+0.002}_{-0.002}$	$3.6^{+1.7}_{-1.1}$	6081^{+95}_{-92}	$1.41^{+0.05}_{-0.05}$	$86.2^{+4.8}_{-5.2}$	$-0.31^{+0.07}_{-0.07}$
KIC10732098	$1.50^{+0.13}_{-0.14}$	$1.14^{+0.03}_{-0.04}$	$1.74^{+0.10}_{-0.09}$	$0.28^{+0.01}_{-0.01}$	$0.018^{+0.002}_{-0.002}$	$6.4^{+0.6}_{-0.6}$	5701^{+59}_{-58}	$1.78^{+0.03}_{-0.03}$	$62.2^{+1.7}_{-1.7}$	$0.06^{+0.06}_{-0.06}$

Table 3.6: The same as Table 3.4, but for the PPS model.

Name	$f_{\rm evol}$	M ($\mathrm{M}_\odot$)	$\alpha_{\rm MLT}$	$Y_{\rm init}$	$Z_{\rm init}$	τ (Gyr)	$T_{\rm eff}$ (K)	R ($\mathrm{R}_\odot$)	$\Delta\nu$ (μHz)	$[\mathrm{M/H}]_{\rm surf}$ (dex)
KIC10079226	$0.35^{+0.11}_{-0.12}$	$1.17^{+0.02}_{-0.03}$	$1.95^{+0.14}_{-0.15}$	$0.27^{+0.01}_{-0.01}$	$0.020^{+0.003}_{-0.002}$	$2.1^{+0.8}_{-0.8}$	5962^{+44}_{-43}	$1.17^{+0.01}_{-0.01}$	$116.0^{+0.7}_{-0.7}$	$0.15^{+0.06}_{-0.07}$
KIC10215584	$0.47^{+0.16}_{-0.16}$	$1.14^{+0.03}_{-0.03}$	$1.90^{+0.15}_{-0.17}$	$0.27^{+0.01}_{-0.01}$	$0.018^{+0.002}_{-0.002}$	$2.7^{+1.2}_{-1.1}$	5943^{+56}_{-58}	$1.18^{+0.02}_{-0.02}$	$112.6^{+2.6}_{-2.6}$	$0.07^{+0.06}_{-0.07}$
KIC10319352	$1.51^{+0.10}_{-0.22}$	$1.09^{+0.03}_{-0.03}$	$1.88^{+0.16}_{-0.16}$	$0.28^{+0.01}_{-0.01}$	$0.028^{+0.004}_{-0.004}$	$9.6^{+1.1}_{-1.2}$	5507^{+47}_{-48}	$1.49^{+0.02}_{-0.02}$	$78.6^{+1.6}_{-1.6}$	$0.28^{+0.06}_{-0.06}$
KIC10322381	$0.89^{+0.21}_{-0.22}$	$1.12^{+0.05}_{-0.06}$	$1.93^{+0.15}_{-0.16}$	$0.26^{+0.01}_{-0.01}$	$0.010^{+0.002}_{-0.002}$	$4.3^{+1.7}_{-1.2}$	6093^{+92}_{-89}	$1.41^{+0.04}_{-0.04}$	$86.1^{+5.0}_{-4.9}$	$-0.31^{+0.07}_{-0.08}$
KIC10732098	$1.60^{+0.11}_{-0.14}$	$1.14^{+0.03}_{-0.04}$	$1.90^{+0.15}_{-0.17}$	$0.27^{+0.01}_{-0.01}$	$0.017^{+0.002}_{-0.002}$	$6.9^{+0.6}_{-0.6}$	5704^{+62}_{-61}	$1.78^{+0.04}_{-0.03}$	$62.2^{+1.8}_{-1.7}$	$0.06^{+0.06}_{-0.06}$

Table 3.7: The same as Table 3.4, but for the MP model.

Name	$f_{\rm evol}$	M ($M_\odot$)	$Y_{\rm init}$	$Z_{\rm init}$	τ (Gyr)	$T_{\rm eff}$ (K)	R ($R_\odot$)	$\Delta\nu$ (μHz)	$[M/H]_{\rm surf}$ (dex)
KIC10079226	$0.20^{+0.08}_{-0.08}$	$1.17^{+0.02}_{-0.03}$	$0.28^{+0.01}_{-0.01}$	$0.019^{+0.003}_{-0.002}$	$1.1^{+0.5}_{-0.4}$	5961^{+42}_{-41}	$1.17^{+0.01}_{-0.01}$	$115.9^{+0.7}_{-0.7}$	$0.15^{+0.06}_{-0.07}$
KIC10215584	$0.36^{+0.14}_{-0.13}$	$1.14^{+0.03}_{-0.03}$	$0.27^{+0.01}_{-0.01}$	$0.018^{+0.002}_{-0.002}$	$2.0^{+0.9}_{-0.8}$	5941^{+57}_{-57}	$1.18^{+0.02}_{-0.02}$	$112.5^{+2.6}_{-2.7}$	$0.07^{+0.06}_{-0.07}$
KIC10319352	$1.41^{+0.10}_{-0.25}$	$1.08^{+0.03}_{-0.03}$	$0.29^{+0.02}_{-0.01}$	$0.028^{+0.004}_{-0.004}$	$8.6^{+1.0}_{-0.9}$	5512^{+44}_{-45}	$1.49^{+0.02}_{-0.02}$	$78.6^{+1.7}_{-1.6}$	$0.28^{+0.06}_{-0.07}$
KIC10322381	$0.77^{+0.23}_{-0.19}$	$1.14^{+0.04}_{-0.06}$	$0.27^{+0.01}_{-0.01}$	$0.011^{+0.002}_{-0.002}$	$3.5^{+1.6}_{-1.0}$	6076^{+96}_{-91}	$1.41^{+0.05}_{-0.05}$	$86.1^{+4.7}_{-5.3}$	$-0.32^{+0.07}_{-0.07}$
KIC10732098	$1.50^{+0.13}_{-0.13}$	$1.14^{+0.03}_{-0.04}$	$0.28^{+0.01}_{-0.01}$	$0.018^{+0.002}_{-0.002}$	$6.4^{+0.6}_{-0.6}$	5702^{+56}_{-58}	$1.78^{+0.03}_{-0.03}$	$62.2^{+1.7}_{-1.7}$	$0.06^{+0.06}_{-0.06}$

Table 3.8: The same as Table 3.4, but for the MPS model.

Name	$f_{\rm evol}$	M ($M_\odot$)	$Y_{\rm init}$	$Z_{\rm init}$	τ (Gyr)	$T_{\rm eff}$ (K)	R ($R_\odot$)	$\Delta\nu$ (μHz)	$[M/H]_{\rm surf}$ (dex)
KIC10079226	$0.44^{+0.07}_{-0.06}$	$1.16^{+0.02}_{-0.03}$	$0.26^{+0.01}_{-0.01}$	$0.021^{+0.003}_{-0.002}$	$2.7^{+0.5}_{-0.4}$	5965^{+40}_{-40}	$1.17^{+0.01}_{-0.01}$	$116.0^{+0.7}_{-0.7}$	$0.15^{+0.06}_{-0.06}$
KIC10215584	$0.59^{+0.11}_{-0.13}$	$1.13^{+0.03}_{-0.03}$	$0.26^{+0.01}_{-0.01}$	$0.018^{+0.002}_{-0.002}$	$3.6^{+0.9}_{-0.9}$	5952^{+55}_{-56}	$1.18^{+0.02}_{-0.02}$	$112.7^{+2.7}_{-2.7}$	$0.08^{+0.06}_{-0.07}$
KIC10319352	$1.61^{+0.04}_{-0.06}$	$1.08^{+0.03}_{-0.03}$	$0.27^{+0.01}_{-0.01}$	$0.028^{+0.004}_{-0.003}$	$10.8^{+0.7}_{-0.8}$	5516^{+46}_{-47}	$1.49^{+0.02}_{-0.02}$	$78.6^{+1.7}_{-1.6}$	$0.28^{+0.07}_{-0.06}$
KIC10322381	$0.98^{+0.19}_{-0.20}$	$1.10^{+0.06}_{-0.05}$	$0.26^{+0.01}_{-0.01}$	$0.010^{+0.002}_{-0.001}$	$5.1^{+1.3}_{-1.5}$	6106^{+94}_{-80}	$1.40^{+0.04}_{-0.04}$	$85.8^{+5.6}_{-4.3}$	$-0.30^{+0.08}_{-0.08}$
KIC10732098	$1.69^{+0.06}_{-0.09}$	$1.14^{+0.03}_{-0.04}$	$0.26^{+0.01}_{-0.01}$	$0.017^{+0.002}_{-0.002}$	$7.4^{+0.5}_{-0.5}$	5715^{+61}_{-61}	$1.77^{+0.04}_{-0.03}$	$62.3^{+1.8}_{-1.8}$	$0.07^{+0.06}_{-0.07}$

Table 3.9: Hyperparameter results for each model with the omission of Y_P.

Model	$\Delta Y/\Delta Z$	σ_Y	μ_α	σ_α	$\alpha_{\rm mlt}$
NP	$1.69^{+0.21}_{-0.21}$	$0.0074^{+0.0026}_{-0.0022}$	$1.954^{+0.040}_{-0.041}$	$0.065^{+0.030}_{-0.024}$	—
PP	$1.60^{+0.45}_{-0.42}$	$0.0051^{+0.0045}_{-0.0027}$	$1.742^{+0.081}_{-0.070}$	$0.056^{+0.051}_{-0.030}$	—
PPS	$1.05^{+0.28}_{-0.25}$	$0.0045^{+0.0038}_{-0.0023}$	$1.900^{+0.095}_{-0.088}$	$0.133^{+0.057}_{-0.047}$	—
MP	$1.60^{+0.45}_{-0.42}$	$0.0051^{+0.0044}_{-0.0027}$	—	—	$1.728^{+0.077}_{-0.066}$
MPS	$0.76^{+0.24}_{-0.27}$	$0.0049^{+0.0039}_{-0.0025}$	—	—	$2.088^{+0.031}_{-0.029}$

3.4.2 Population Parameter Results

We obtained values for the hyperparameters for each of the models and present them in Table 3.9 along with their upper and lower 68 per cent credible regions. We omit the results for Y_P because its posterior is the same as the prior, $Y_P = 0.247 \pm 0.001$ for all the models. We fit the same hyperparameters from the PP model to the NP model results for $Y_{\rm init}$, $Z_{\rm init}$, and $\alpha_{\rm MLT}$ for the purpose of comparison. However, the NP model results suffer from boundary effects which makes the resulting fit unreliable, pushing the population mean to the centre of the priors and underestimating the uncertainties. We leave the NP results here only for completeness.

Fig. 3.3 shows the joint and marginal distributions (corner plot) output by the PP and PPS model. We see an anti-correlation between $\Delta Y/\Delta Z$ and μ_α, expected due to the degeneracy between the two parameters in the stellar evolutionary models. In Fig. 3.4, we also show the corner plot for the MP and MPS model output. Similarly, we see an anti-correlation between $\Delta Y/\Delta Z$ and $\alpha_{\rm MLT}$.

We present the helium enrichment relation resulting from the PPS model in Fig. 3.5. In this figure, we plot the individual results for $Y_{\rm init}$ and $Z_{\rm init}$ for each of the stars from the NP and PPS models. This is an example of shrinkage in the HBM; the estimates for individual stellar parameters move towards the mean of the population when they are pooled.

3.5 Discussion

So far, we have shown that we can add parameters to stellar models without sacrificing statistical uncertainties through the application of an HBM. We freed the $Y_{\rm init}$ and $\alpha_{\rm MLT}$ using pooling to

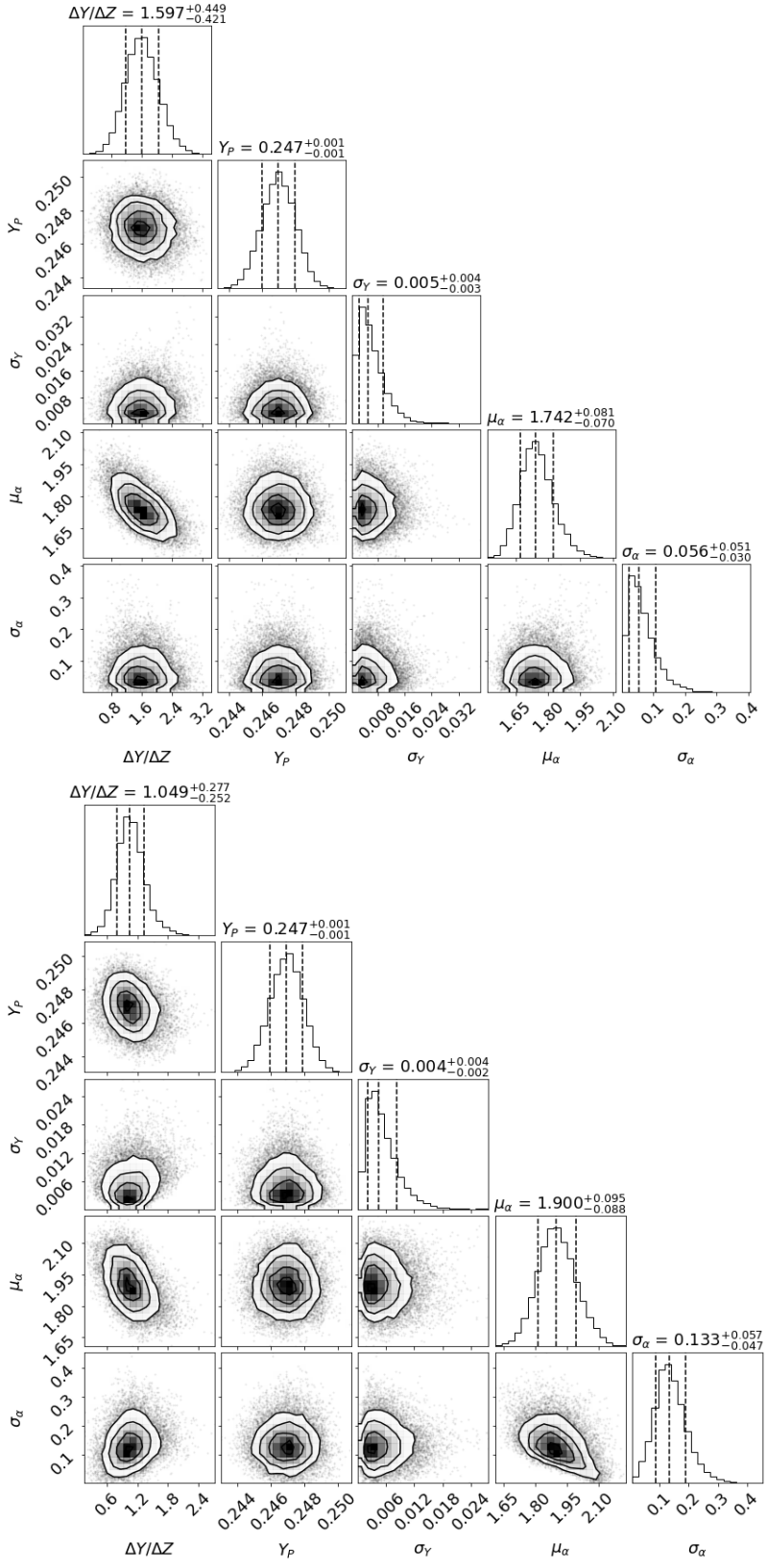


Figure 3.3: Corner plots showing the joint and marginalised sampled posterior distributions for the hyperparameters for the PP (top) and PPS (bottom) models. The vertical dashed lines give the 16th, 50th and 84th percentiles.

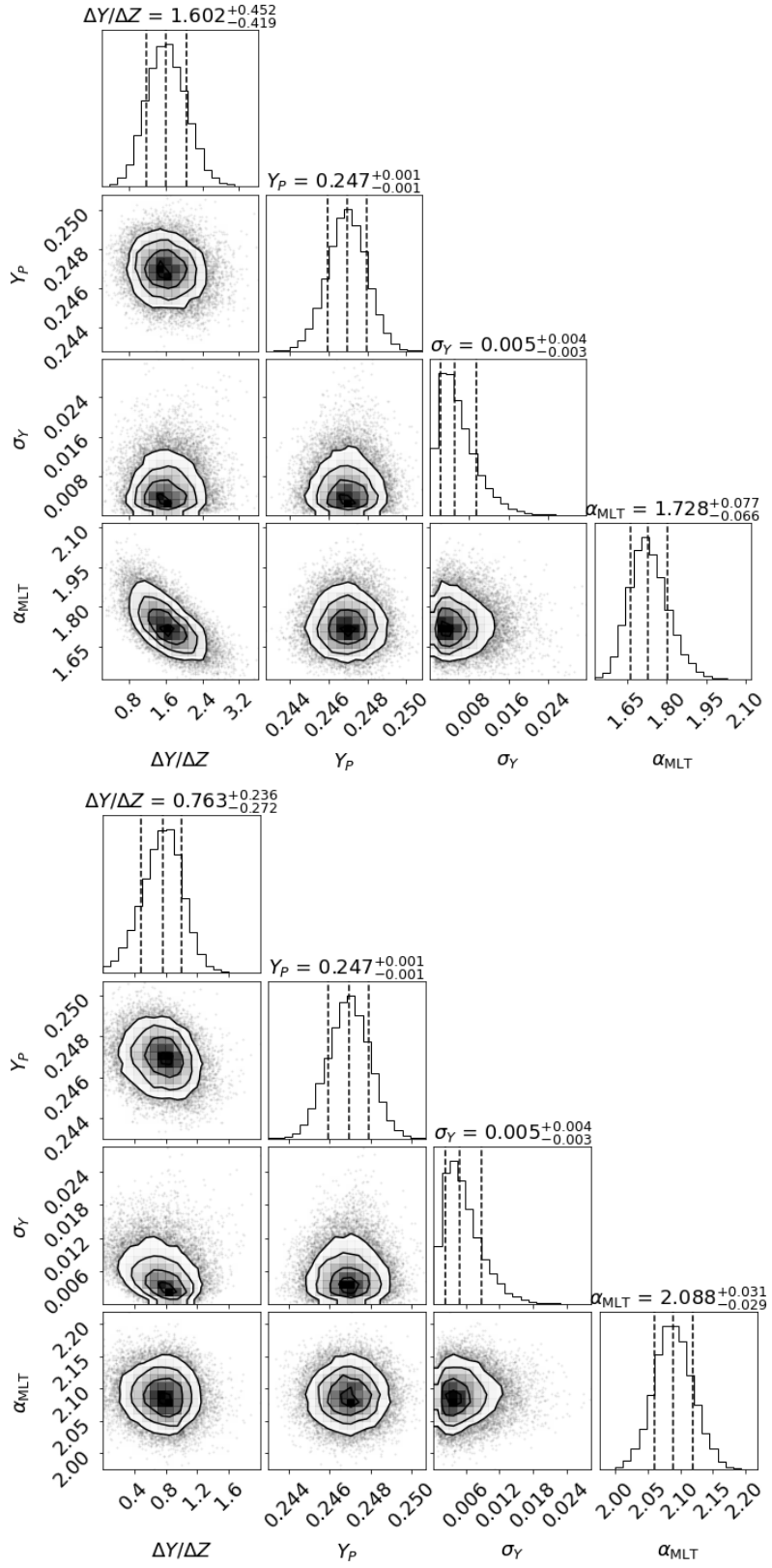


Figure 3.4: The same as Fig. 3.3 but for the MP (top) and MPS (bottom) models.

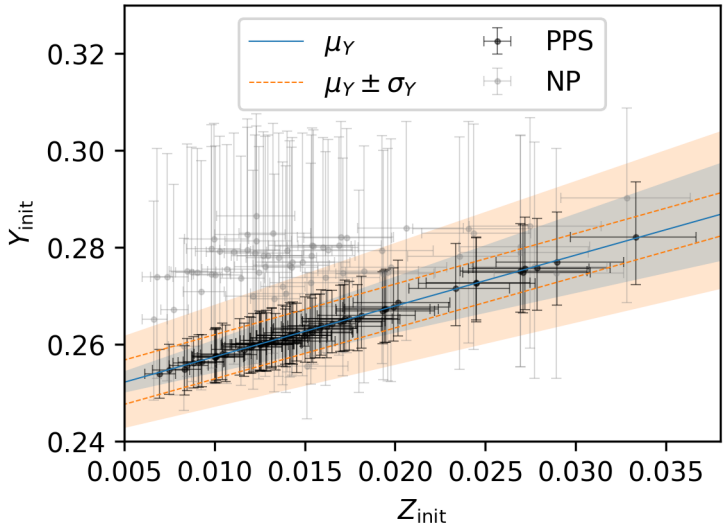


Figure 3.5: The results for initial helium fraction ($Y_{\rm init}$) against initial heavy-element fraction ($Z_{\rm init}$) for each star from the PPS model are shown by the black markers. The mean helium enrichment, $\mu_Y = Y_P + (\Delta Y/\Delta Z)Z_{\rm init}$ with its 68 per cent credible interval are shown in blue by a solid line and shaded region respectively. The population spread, $\mu_Y \pm \sigma_Y$ and its 68 per cent credible interval are shown in orange by a dashed line and shaded region respectively. The individual results from the NP model are shown by the light grey markers.

encode our prior knowledge of their distribution in the population. We also tested the impact of including the Sun as a star in our population. We first discuss the impact of pooling and our choice of population priors for $Y_{\rm init}$ and $\alpha_{\rm MLT}$ in Section 3.5.1 and 3.5.2. To assess the accuracy of our model with respect to the literature, we compare our results to those of S17 in Section 3.5.3. We found good agreement between this work and their results, despite some differences in observables and stellar model physics which we discuss further. Then, we discuss sources of systematic uncertainties in Section 3.5.4. Although we have accounted for uncertainties in 3.5.1 and 3.5.2 in our model, there are still differences between stellar modelling codes and other model physics which should be considered. Finally, in Section 3.5.5, we highlight a possible outlier in our dataset.

3.5.1 Helium Enrichment

We found the value for the helium enrichment ratio, $\Delta Y/\Delta Z$ to be the same in both the PP and MP models, $\Delta Y/\Delta Z = 1.6^{+0.5}_{-0.4}$. This is consistent with values of ~ 1.4 in the literature (when

heavy element diffusion is included) albeit obtained through different methods: e.g. measuring the metallicity and helium abundance of an open cluster and comparing with the primordial helium abundance (Brogaard et al., 2012), and fitting to helium abundances determined for *Kepler* field stars using asteroseismology (Verma et al., 2019).

When we added the Sun to the pooled models, PPS and MPS, obtained $\Delta Y/\Delta Z$ up to 2-σ lower than the models without the Sun. In both models, the resulting $\Delta Y/\Delta Z$ of approximately 0.8 to 1.0 was consistent with the initial helium fraction expected from solar models with our choice of Asplund et al. (2009) abundances (Serenelli and Basu, 2010). However, such solar models have been shown to not recover helioseismic measurements of helium in the Sun (Basu and Antia, 2004; Serenelli et al., 2009; Villante et al., 2014). Solar models with the older Grevesse and Sauval (1998) abundances typically yield higher helium fractions more in-line with helioseismology. The $\Delta Y/\Delta Z$ from the PP and MP models are higher than those including the Sun. We could extend our model to include asteroseismic indicators of helium to improve the uncertainties on Y and test whether this difference becomes more significant.

Our models assumed a prior of $Y_P = 0.247 \pm 0.001$ for the primordial helium fraction which dominated its posterior. This was a sensible assumption to make when using a linear enrichment law, because measurements of the primordial helium correspond to the abundance at the epoch of BBN according to current cosmological theory (Cyburt et al., 2016). However, if we used a less informative prior for Y_P we might yield more uncertain results, or even a different value for Y_P. In previous work fitting a linear enrichment law, some results for Y_P suggested a value below the BBN value (Casagrande et al., 2007; Silva Aguirre et al., 2017). It is more probable that the assumption of a linear enrichment law is inaccurate than a sample of stars could contradict independent Y_P from cosmology. We justify our prior on Y_P as in-line with the assumption of a linear enrichment law, but highlight the need to investigate other ways of describing helium in a population of stars.

3.5.2 Mixing-Length Theory

To a greater degree than chemical composition, the best-fitting $\alpha_{\rm MLT}$ depends on the choice of model physics and stellar modelling code. The MLT is an approximation of convection which is often calibrated to the Sun and then assumed for all stars in a model. However, studies of 3D hydrodynamical simulations suggest that the degree to which $\alpha_{\rm MLT}$ approximates convection varies across the HR diagram (Magic, Weiss, and Asplund, 2015) and this is confirmed when modelling stars with asteroseismology (Tayar et al., 2017).

The PP model (without the Sun) favoured a mean mixing-length parameter of $\mu_\alpha \simeq 1.7$. Whereas, the PPS model yielded a higher value of $\mu_\alpha \simeq 1.9$ by ~ 2-σ. We found this was attributed to the addition of the Sun. The individual solar results for the PPS model yielded a value of $\alpha_{\rm MLT\odot} = 2.12 \pm 0.03$ which was considerably higher than the $\alpha_{\rm MLT}$ obtained for the other stars in the sample (see Appendix A.4). This result agrees with the 3D simulations (e.g. Trampedach et al., 2014), which predict lower $\alpha_{\rm MLT}$ for stars with lower $T_{\rm eff}$ and higher $\log g$. However, the solar value also exceeds the reference solar calibrated values of ≈ 1.92 for the same stellar evolution code (Paxton et al., 2011). This is caused by the differences in adopted solar mixture, atmospheric boundary conditions and the treatment of convective mixing between this work and typical reference values.

Despite the difference in μ_α, the resulting spread in mixing-length for the PPS model $\sigma_\alpha \approx 0.13$ was double that of the PP model to cope with the high solar value. This implies that a large population spread in $\alpha_{\rm MLT}$ could explain the difference we see. In other words, if we assume that the best-fitting $\alpha_{\rm MLT}$ is normally distributed in our population, then the Sun lies within 2-σ of the mean, among 95 per cent of all stars in the population.

There are a few prior studies which look at the spread in $\alpha_{\rm MLT}$ for a population of stars, typically by fitting $\alpha_{\rm MLT}$ as a function of [M/H], $T_{\rm eff}$, and $\log g$ (e.g. Bonaca et al., 2012; Viani et al., 2018). For example, results from Viani et al. (2018) for stellar models including diffusion, predict $\alpha_{\rm MLT}$ in the range 1.5 to 2.3 across our sample. This dispersion would be more compatible with the larger spread obtained by our PPS model. However, in future work

we should further investigate how $\alpha_{\rm MLT}$ varies with stellar parameters, as our assumption of a normal distribution may not be accurate.

We found a greater difference in $\alpha_{\rm MLT}$ between the models with and without the Sun when we max-pooled $\alpha_{\rm MLT}$. The MP models yielded a global $\alpha_{\rm MLT}$ in-line with μ_α from the PP model. However, when we added the Sun, the model yielded $\alpha_{\rm MLT} \approx 2.1$ which is in common with the solar results (see Appendix A.4). This had a similar affect as assuming a solar calibrated value, because the model favoured fitting to data with the best observational precision. The change in $\alpha_{\rm MLT}$ between the MP and MPS models resulted in a mean difference of ~ 20 per cent between the individual stellar ages. This is an example of how adopting a solar calibrated value can bias stellar ages. We argue that carefully including the Sun as a part of the population with an intrinsic spread is a better way to calibrate the stellar models than assuming as solar $\alpha_{\rm MLT}$ across the sample.

In all observables except for L, the Sun is near the centre of our distribution of stars. However, we found no relationship between L and $\alpha_{\rm MLT}$ in both our NP and PP models. A possible explanation for the difference in $\alpha_{\rm MLT}$ with and without the Sun could be some systematic offset in our observational data for the sample. Here, we point to our choice of spectroscopic $T_{\rm eff}$ which typically underestimates $T_{\rm eff}$ compared to photometric scales, as noted in S17. We ran a solar model with an additional parameter, $\Delta T_{\rm eff} = T_{\rm eff,obs} - T_{\rm eff}$ which represents a bias in the observed effective temperature. The estimated covariance between $\Delta T_{\rm eff}$ and $\alpha_{\rm MLT}$ was 0.452 K[3] (with a correlation of 0.517). Therefore, underestimating $T_{\rm eff}$ by about 100 K could underestimate $\alpha_{\rm MLT}$ by about 0.1. If we extend this result to the other stars in the sample, the lower $\alpha_{\rm MLT}$ obtained without including the Sun as a star could be caused by underestimating $T_{\rm eff}$ relative to the Sun. Alternatively, the $\alpha_{\rm MLT}$ of the Sun could have been higher than the rest of the sample to compensate for neglecting additional sources of mixing required to reproduce the higher precision solar observables. A deeper quantification of systematic uncertainties is left to future work.

[3] $\alpha_{\rm MLT}$ is dimensionless, hence the units of covariance are K.

3.5.3 Comparison with APOKASC Results

Before we compare our results to S17, we should highlight some key differences between our data and methodology. The results from S17 were determined using a grid-based-modelling technique, which estimates the likelihood across a dense grid of stellar models. They used results from several pipelines to estimate the systematic uncertainties. For the central values of their results, they used the Bayesian stellar algorithm (BASTA; Silva Aguirre et al., 2015) using a grid computed with GARSTEC (Weiss and Schlattl, 2008). Their choice of stellar physics was similar to this work, except for two major differences.

Firstly, the results of S17 were determined using stellar models calculated without heavy-element diffusion. The inclusion of diffusion when modelling the Sun has been commonplace over the last few decades, with good agreement between models and helioseismic observations (Christensen-Dalsgaard, Proffitt, and Thompson, 1993; Bahcall, Pinsonneault, and Wasserburg, 1995). More recent work explored the diffusion in cluster stars (Korn et al., 2007; Önehag, Gustafsson, and Korn, 2014) and another demonstrated the impact of including diffusion on stellar ages (Dotter et al., 2017). Our stellar models were computed with heavy-element diffusion. Recently, work by Nsamba et al. (2018) on a similar sample of stars showed, on average, models without diffusion compared to those including diffusion can lead to underestimated radii and masses, and overestimated ages by 1, 3, and 16 per cent respectively.

Secondly, our choice of Asplund et al. (2009) solar chemical mixture differs from the Grevesse and Sauval (1998) mixtures adopted by S17. The former leads to a solar heavy-element to hydrogen ratio of $(Z/X)_\odot = 0.0181$, and the latter, $(Z/X)_\odot = 0.0230$. Typically, Grevesse and Sauval (1998) abundances are favoured in asteroseismic modelling because they are better able to reproduce measurements of helium in the Sun from helioseismology (Serenelli et al., 2009). An effect of using the Asplund et al. (2009) abundances, is that it favours lower $Z_{\rm init}$ for a given $[{\rm M/H}]_{\rm surf}$. As a result, models using Grevesse and Sauval (1998) abundances on average underestimate radii and mass compared to those without by about 1 and 0.5 per cent respectively (Nsamba et al., 2018).

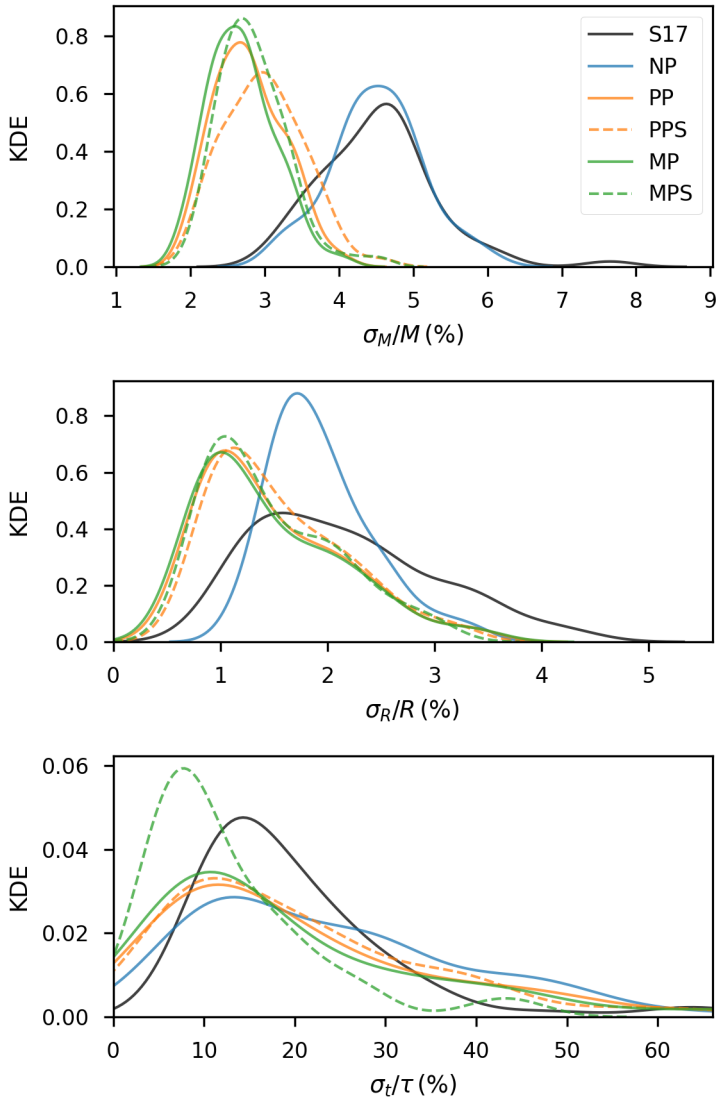


Figure 3.6: Kernel density estimates (KDEs) of the distribution of statistical uncertainties in the results from each model compared with those of S17 for the sample of APOKASC dwarfs and subgiants.

Although updated, much of our observable data is comparable to that of S17, except for $T_{\rm eff}$. The preferred results from S17 were determined using a photometric $T_{\rm eff}$ scale which we found to be on average $\sim$ 170 K greater than our spectroscopic scale from DR14. In S17, they saw a similar offset between the DR13 $T_{\rm eff}$ available at the time. They found a median difference in mass, radius, and age of approximately −6, −2, and +35 per cent respectively with results from the photometric $T_{\rm eff}$ scale subtracted from the spectroscopic scale.

In Fig. 3.6 we compare our statistical uncertainties for M, R, and τ with those for the

equivalent stars from S17. We found that the NP model yielded comparable uncertainties to S17 but note that these are likely underestimated due the influence of the prior boundaries for $Y_{\rm init}$ and $\alpha_{\rm MLT}$. We expected larger uncertainties because we included additional free parameters ($Y_{\rm init}$ and $\alpha_{\rm MLT}$) over the work of S17. However, when we treat these parameters hierarchically, we saw a reduction in uncertainties from all the pooled models. This is because our prior assumptions about the population allows for the sharing of information between the stars. This uncertainty reduction scales with the number of stars in our sample, demonstrated by the results for the synthetic stars in Fig. A.5. Thus, hierarchically modelling our population resulted in improved statistical uncertainties in stellar fundamental parameters.

In the following subsection, we compare the results between our PPS model with that of S17 for mass, radius, and age with reference to Fig. 3.7. We preferred the PPS model for comparison because it utilised the high-precision data available for the Sun as a star to help calibrate the sample, while partially pooling both $Y_{\rm init}$ and $\alpha_{\rm MLT}$ to allow for small variations within the population.

Mass, Radius, and Age

In the top panel of Fig. 3.7, we compare the masses obtained by the PPS model with S17 and found a dispersion of around 2 per cent. Our masses were on average 1 per cent above the results from S17. Although we might expect the lower $T_{\rm eff}$ scale in this work to underestimate the mass, we attribute this overall effect to our choice of stellar model physics. As previously discussed, the use of Asplund et al. (2009) solar abundances and heavy-element diffusion has the cumulative effect of overestimating stellar masses compared to the physics adopted by S17. We also found that the results from all the pooled models returned similar masses, with or without the Sun.

In the central panel of Fig. 3.7, we show that our radii are similar to S17 with a spread of 1 per cent. We also found radii on average 1 per cent greater than the APOKASC results. Similarly to with mass, this contradicts what would be expected from a lower $T_{\rm eff}$ scale and could also be explained by model physics. Our radii also varied little between models with and

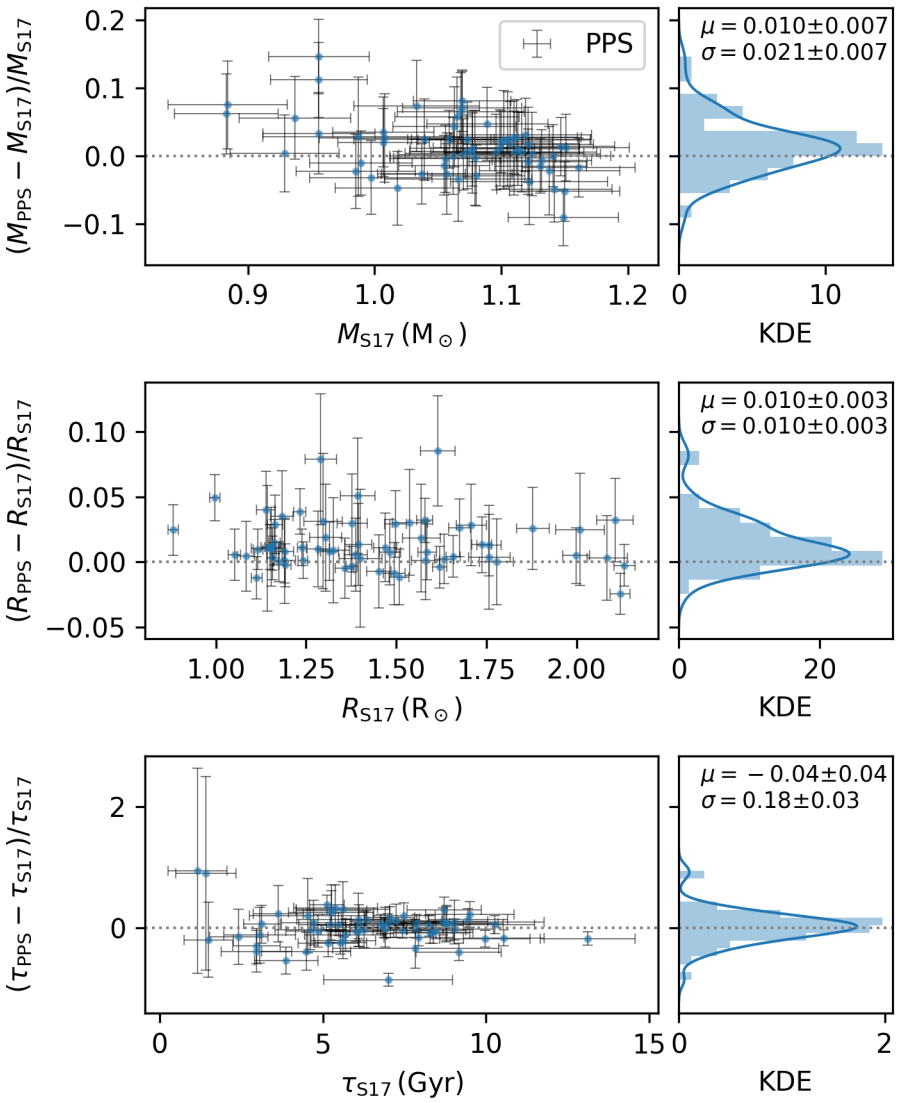


Figure 3.7: The mean and standard deviation in age, mass and radius results from the PPS model compared with the results (using the photometric temperature scale) from S17.

without the Sun.

Our ages were also consistent with those from S17. The bottom panel of Fig. 3.7 shows the spread in the relative age differences to be about 18 per cent, slightly underestimated by 4 per cent. We would expect the lower $T_{\rm eff}$ scale to overestimate the ages as found in S17, but instead they are comparable. However, as discussed previously, including diffusion has been shown to reduce age estimates compared to those without. Since we have included diffusion, this could explain the similar ages despite the difference in $T_{\rm eff}$ scale.

Including the Sun in our pooled models affected the resulting ages more than mass and

radius. Including the Sun typically overestimated the ages compared to models without the Sun. This is expected given the higher $\alpha_{\rm MLT}$ for the models including solar data, because a larger mixing-length leads to more efficient nuclear burning and more time spent during the MS phase.

3.5.4 Systematic Uncertainties

We have already accounted for systematics due to the choice of helium enrichment and mixing-length parameter by marginalising over their uncertainties assuming their population distributions. However, there are other model physics which we have not freely varied, including diffusion and choice of solar mixture. Although our method can be adapted to different stellar evolutionary codes and choice of physics, an in-depth analysis of systematic uncertainties is left to future work.

In previous work studying stars in the APOKASC sample, several pipelines used a range of stellar evolutionary codes and model physics are employed to evaluate systematic uncertainties from the models (Serenelli et al., 2017; Silva Aguirre et al., 2017). Using a hierarchical model in this work enabled us to reduce median statistical uncertainties to 2.5 per cent in mass, 1.2 per cent in radius, and 12 per cent in age. However, the systematic uncertainty analysis of S17 found median systematic uncertainties of 3, 1, and 13 per cent in mass, radius, and age respectively. These are comparable, highlighting the importance of including sources of systematic uncertainty in our model.

Other systematics could arise from observational data. For example, we chose the ASPCAP DR14 $T_{\rm eff}$ scale which was systematically lower than the photometric scale of choice in S17. However, our method was still able to recover similar masses, radii, and ages. This could be explained by our choice of stellar model physics, as discussed previously.

3.5.5 Outliers

We identified KIC 9025370 as a possible outlier. Consistent across all our models, its modelled effective temperature, $T_{\rm eff} = 5934 \pm 50\,{\rm K}$ was about 4-σ greater than its observed $T_{\rm eff}$, and its modelled L was about 2-σ dimmer than its observed luminosity. Only $\Delta\nu$ and $[{\rm M/H}]_{\rm surf}$ were

consistent between modelled and observed values. The difference was also apparent in our comparison of ages with S17 where we obtained an age of $1.5^{+0.7}_{-0.6}$ Gyr compared to their value of $7.0^{+2.0}_{-1.6}$ Gyr.

KIC 9025370 turned out to be a double-lined spectroscopic binary (Nissen et al., 2017), discovered after S17 and hence included in the original sample. The higher observed luminosity from ISOCLASSIFY and inconsistent spectroscopic $T_{\rm eff}$ compared with our model posteriors were compatible with a spectroscopic binary. We calculated a photometric $T_{\rm eff}$ using the IRFM method (Casagrande et al., 2010) with the available 2MASS photometry for the target and obtained $T_{\rm eff} = 5983 \pm 120$ K, more consistent with our modelled effective temperature and inconsistent with its spectroscopic $T_{\rm eff}$. Thus, our inferred $T_{\rm eff}$ was within the dispersion between different observed $T_{\rm eff}$ scales. Running the model without KIC 9025370 did not affect the resulting inferred hyperparameters, demonstrating the robustness of our model. Therefore, we present KIC 9025370 in our results but suggest that further investigation should be carried out.

3.6 Conclusion

We have shown that modelling $Y_{\rm init}$ and $\alpha_{\rm MLT}$ to improve inference of fundamental parameters can be done through the use of an HBM, whilst still improving statistical uncertainties. Our results were in good agreement with S17 with small changes in mass and radii expected from our choice of model physics and updated observables. Taking our partially-pooled model including the Sun (PPS) as our preferred set of results, we obtained median statistical uncertainties on M, R, and τ of 2.5, 1.2, and 12 per cent respectively. Furthermore, we demonstrated that the uncertainties reduced with increasing sample size in a population of synthetic stars, giving scope to further improve our inference on larger sample sizes from *TESS*.

We found that the gradient, $\Delta Y/\Delta Z$ of the linear helium enrichment law ranged from 0.8 to 1.6 depending on the level of parameter pooling and the inclusion of the Sun in our sample, with $\Delta Y/\Delta Z = 1.1^{+0.3}_{-0.3}$ from our preferred PPS model. Consistent across our models was the

spread in initial helium about the enrichment law, $\sigma_Y = 0.005^{+0.004}_{-0.003}$. The mean $\alpha_{\rm MLT}$ in the population was $\mu_\alpha = 1.90^{+0.10}_{-0.09}$ for the PPS model, with values from 1.7 to 2.1 depending on the level of pooling and whether or not solar data was included. We also found the spread in $\alpha_{\rm MLT}$ doubled to $\sigma_\alpha = 0.13^{+0.06}_{-0.05}$ to account for the addition of the Sun in our sample. We conclude that there are still discrepancies between the best-fitting $\alpha_{\rm MLT}$ in our population and that of the Sun which need to be investigated further. Perhaps, the addition of asteroseismic signatures of helium abundance (see e.g. Verma et al., 2017, and Chapter 4) would improve our constraints on $Y_{\rm init}$ and thus reduce star-by-star uncertainties in $\alpha_{\rm MLT}$.

Using HBMs has allowed us to introduce more free parameters without inflating statistical uncertainties. We used an ANN to approximate stellar models, a method which can be extended to higher input dimensions with little impact on training and evaluation time. Our model also scales well with the number of stars, making use of GPU parallel processing when sampling the posterior.

As shown in tests with synthetic stars (Appendix A.3) and apparent in Fig. 3.6, increasing the number of stars decreases the statistical uncertainties when parameters are pooled. The theoretical limit to this improvement is $\sqrt{N_1/N_2}$ for two populations of size N_1 and N_2. For example, if we increase our sample to 300 stars, we would expect the uncertainties to reduce by up to a factor of 2. Naturally, the uncertainty is still limited by observational precision. However, hierarchical modelling as demonstrated in this work, allows us to get the most out of our data and paves the way for a data-driven analysis of model systematics.

Including all-sky data from *TESS* and in anticipation of *PLATO* (Rauer et al., 2014) we can expect our sample size of asteroseismic dwarfs and subgiants only to increase. There is also scope to extend our grid of models to include red giants, for which there are vast catalogues of stars already studied with *Kepler* (Pinsonneault et al., 2018).

Chapter 4

ACOUSTIC GLITCHES IN SOLAR-LIKE OSCILLATORS AS A SIGNATURE OF HELIUM ABUNDANCE

In this chapter, I introduce the concept of an acoustic glitch in the structure of a star producing a measurable signal in its observable oscillation modes. In Section 4.2, I start with a simple one-dimensional example of a glitch. Then, I review the theory of glitch signatures due to helium ionisation and the base of the convective zone in Section 4.3.

4.1 Introduction

An acoustic glitch is a rapid variation of the sound speed inside a medium. The presence of a glitch induces a periodic deviation from the regular separation of p mode frequencies. Acoustic glitches in Sun-like stars arise from sharp variations in their structure, such as the base of the convection zone (BCZ) and the first and second ionisation of helium (He I and He II). There have been several attempts to measure this effect in the Sun and other stars over the past few decades.

Early work identified glitches in solar p modes by analysing their second differences, $\Delta_2 \nu_{nl} \equiv \nu_{n-1\,l} - 2\nu_{nl} + \nu_{n+1\,l}$. Second and higher-order differences remove some slowly varying components of the mode frequencies and amplify faster varying components. Assuming a

sharp localised discontinuity in sound speed at He II ionisation and the BCZ, Basu, Antia, and Narasimha (1994) and Basu (1997) modelled glitch signatures in the Sun to constrain the extent of convective overshoot at the BCZ. Monteiro and Thompson (1998) further developed a model of the He II ionisation glitch signature by accounting for its finite width in the star, later applying it to study the helium ionisation zone of the Sun (Monteiro and Thompson, 2005). Around the same time, Basu et al. (2004) showed that the amplitude of the He II glitch signature correlated with the fractional helium abundance (Y) near the surface of Sun-like stars. A few years later, Houdek and Gough (2007) proposed a closer physical approximation of the He II glitch. They derived the glitch signature which was later used in several studies of solar-like oscillators.

Since helium ionises at $T \gtrsim 10^4$ K, below the atmosphere of cool stars ($T_{\rm eff} \sim 5 \times 10^3$ K), we can probe helium abundance with asteroseismology where we cannot with spectroscopy. Glitches could be further studied in stars other than the Sun with the advent of space-based missions. Miglio et al. (2010) and Mazumdar et al. (2012b) were among the first to measure glitch signatures in other stars after *CoRoT* (Baglin et al., 2006) provided evidence of solar-like oscillations in red giants. Then, studies of glitches in main sequence stars observed by *Kepler* (Borucki et al., 2010) compared using modes frequencies directly with using second differences (Mazumdar et al., 2012a, 2014; Verma et al., 2017). More recently, Verma et al. (2019) used measurements of the glitch to estimate the helium abundance for the LEGACY sample of stars Lund et al. (2017).

The reason that acoustic glitches cause a periodic signal in the mode frequencies is not obvious. Hence, we go through a simple, one-dimensional example in Section 4.2. Then, in Section 4.3, we demonstrate how glitches in stellar structure due to helium ionisation and the BCZ affect the mode frequencies of solar-like oscillators. Starting with the variational principle (Chandrasekhar, 1964), we will show how to get to the He II glitch signature formula from Houdek and Gough (2007).

4.2 A One-Dimensional Example of a Glitch

A rapid variation in the structure of a medium induces a periodic perturbation ($\delta\omega$) to the eigenfrequencies. To demonstrate this, we will explore a simple one-dimensional example (similar to that of Verner, 2005). Consider a medium bound from $x = 0$ to $x = L$ in which pressure waves can propagate at constant speed, c. The longitudinal displacement of the wave (ξ) obeys the wave equation,

$$\frac{\partial^2 \xi(x,t)}{\partial t^2} = c^2 \frac{\partial^2 \xi(x,t)}{\partial x^2}, \tag{4.1}$$

at a given position (x) and time (t). We may write a general solution to the wave as a sum of right- and left-travelling waves. In terms of the angular frequency (ω), wave number (k), and complex coefficients (A, B),

$$\xi(x,t) = A\mathrm{e}^{i(\omega t - kx)} + B\mathrm{e}^{i(\omega t + kx)}, \tag{4.2}$$

where ω and k satisfy $\omega = ck$. Solving for the boundary condition, $\xi(0,t) = 0$, we find $B = -A$. Substituting Euler's formula, $A = (r/2)\mathrm{e}^{i\phi}$, we can represent the physical component of the wave by its real solution,

$$\mathrm{Re}\left[\xi(x,t)\right] = r \sin kx \sin(\omega t + \phi), \tag{4.3}$$

where r and ϕ are the amplitude and phase respectively. Solutions for ω which satisfy $\xi(L,t) = 0$ may then be found,

$$\omega_n = c\frac{n\pi}{L}, \tag{4.4}$$

where n is a non-zero integer (the $n = 0$ solution would give $\xi = 0$ everywhere).

Now, let us suppose there is a small structural perturbation (or glitch) in the medium at position x_g with half-width δx. Figure 4.1 shows this system divided into 3 regions, with region 2 containing the glitch. The speed of sound is $c + \delta c$ in region 2, where we let the corresponding wave number be $k + \delta k$. We want to find the frequencies which correspond to standing waves in this system and compare them to that of the homogeneous medium above. We will show that

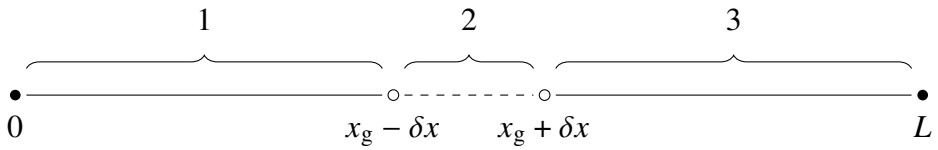


Figure 4.1: A diagram showing a one-dimensional medium split into three regions: (1) Fixed at $x = 0$ with a constant speed of sound c; (2) A small structural perturbation centred at $x = x_g$ with width $2\delta x$ and constant speed of sound $c + \delta c$; (3) Fixed at $x = L$ with a constant speed of sound c.

the resulting perturbation to the eigenfrequencies ($\delta\omega$) is periodic, with an amplitude and period that relates to the properties of the glitch.

Firstly, we propose general solutions to the wave for each region by considering reflection and transmission at each boundary. Initially ignoring the wave superposed by the boundary at $x = L$,

$$\xi_1(x,t) = \mathrm{e}^{i(\omega t - kx)} + A\mathrm{e}^{i(\omega t + kx)}, \tag{4.5}$$

$$\xi_2(x,t) = B\mathrm{e}^{i(\omega t - (k+\delta k)x)} + C\mathrm{e}^{i(\omega t + (k+\delta k)x)}, \tag{4.6}$$

$$\xi_3(x,t) = D\mathrm{e}^{i(\omega t - kx)}, \tag{4.7}$$

where complex coefficients A and C represent reflections, and B and D represent transmissions at $x_g \pm \delta x$ respectively. Later, we will substitute the left-travelling wave ($-\xi\{-k, -\delta k\}$) after determining the values of the coefficients.

The internal boundary conditions for this system are given by enforcing spacial continuity and smoothness at $x_g \pm \delta x$,

$$\begin{aligned}
\xi_1(x_g - \delta x, t) &= \xi_2(x_g - \delta x, t), \\
\xi_2(x_g + \delta x, t) &= \xi_3(x_g + \delta x, t), \\
\frac{\partial \xi_1}{\partial x}(x_g - \delta x, t) &= \frac{\partial \xi_2}{\partial x}(x_g - \delta x, t), \\
\frac{\partial \xi_2}{\partial x}(x_g + \delta x, t) &= \frac{\partial \xi_3}{\partial x}(x_g + \delta x, t).
\end{aligned} \tag{4.8}$$

Solving these simultaneously with the Python package `sympy` (Meurer et al., 2017) gives the

following equations for the complex coefficients[1],

$$A = \delta k(2k + \delta k)(1 - \mathrm{e}^{4i\delta x(k+\delta k)})\mathrm{e}^{-2ik(x_\mathrm{g}-\delta x)}\alpha^{-1}, \tag{4.9}$$

$$B = 2k(2k + \delta k)\mathrm{e}^{4i\delta x(k+\delta k)}\mathrm{e}^{i\delta k(x_\mathrm{g}-\delta x)}\alpha^{-1}, \tag{4.10}$$

$$C = 2k\delta k\mathrm{e}^{-i(x_\mathrm{g}-\delta x)(2k+\delta k)}\alpha^{-1}, \tag{4.11}$$

$$D = 4k(k + \delta k)\mathrm{e}^{2i\delta x(2k+\delta k)}\alpha^{-1}, \tag{4.12}$$

where,

$$\alpha = (2k + \delta k)^2\mathrm{e}^{4i\delta x(k+\delta k)} - \delta k^2. \tag{4.13}$$

Now we have solutions for the coefficients, we can superpose the right-travelling wave, $\xi\{k, \delta k\} \rightarrow -\xi\{-k, -\delta k\}$ to get the full solution for the wave function. Substituting $k \rightarrow -k$ and $\delta k \rightarrow -\delta k$ into the coefficients yields their complex conjugates, $(\overline{A}, \overline{B}, \overline{C}, \overline{D})$. This allows us to rewrite the wave functions into a more flexible form. For example, substituting Euler's formula, $A = (r_A/2)\mathrm{e}^{i\phi_A}$ and $\overline{A} = (r_A/2)\mathrm{e}^{-i\phi_A}$, Equation 4.5 now becomes,

$$\xi_1(x, t) = \mathrm{e}^{i\omega t}\left[\frac{r_A}{2}\left(\mathrm{e}^{i(kx+\phi_A)} - \mathrm{e}^{-i(kx+\phi_A)}\right) - \left(\mathrm{e}^{ikx} - \mathrm{e}^{-ikx}\right)\right], \tag{4.14}$$

where its real component in trigonometric form is,

$$\mathrm{Re}\left[\xi_1(x, t)\right] = \sin\omega t\left[2\sin kx - r_A\sin(kx + \phi_A)\right]. \tag{4.15}$$

However, this equation does not satisfy the outer boundary condition that the displacement is always zero at $x = 0$; i.e. $\xi_1(0, t) = -r_A\sin\omega t\sin(\phi_A) \neq 0$, everywhere. To fix this, we introduce a small phase displacement ϵ caused by the glitch and let $x \rightarrow x + \epsilon$.

Superposing the right-travelling wave and substituting Euler's formula into Equations 4.6

[1]The code for these derivations is available at https://github.com/alexlyttle/thesis/tree/glitch/notebooks

and 4.7, we write the real components of the wave functions in trigonometric form as,

$$\mathrm{Re}[\xi_1(x,t)] = \sin\omega t\,\{2\sin[k(x+\epsilon)] - r_A\sin[k(x+\epsilon)-\phi_A]\}\,, \tag{4.16}$$

$$\mathrm{Re}[\xi_2(x,t)] = \sin\omega t\,\{r_B\sin[(k+\delta k)(x+\epsilon)-\phi_B] - r_C\sin[(k+\delta k)(x+\epsilon)-\phi_C]\}\,, \tag{4.17}$$

$$\mathrm{Re}[\xi_3(x,t)] = \sin\omega t\,\{r_D\sin[k(x+\epsilon)-\phi_D]\}\,. \tag{4.18}$$

We can see that Equation 4.16 can now satisfy the condition that $\xi_1(0,t) = 0$. Imposing this boundary condition, we solve Equation 4.16 for ϵ at $x = 0$,

$$\begin{aligned}\epsilon &= \frac{1}{k}\tan^{-1}\left(\frac{(r_A/2)\sin(\phi_A)}{1-(r_A/2)\cos(\phi_A)}\right),\\ &= \frac{1}{k}\tan^{-1}\left(\frac{\mathrm{Im}[A]}{1-\mathrm{Re}[A]}\right).\end{aligned} \tag{4.19}$$

Finally, we use the boundary condition $\xi_3(L,t) = 0$ to solve for ω. Letting Equation 4.18 equal zero at $x = L$, we rewrite it in terms of the real and imaginary components of D,

$$\begin{aligned}\sin\omega t\,\{r_D\sin[k(L+\epsilon)-\phi_D]\} &= 0, \quad (\div\sin\omega t)\\ r_D\cos\phi_D\sin[k(L+\epsilon)] - r_D\sin\phi_D\cos[k(L+\epsilon)] &= 0,\\ \mathrm{Re}[D]\sin[k(L+\epsilon)] - \mathrm{Im}[D]\cos[k(L+\epsilon)] &= 0.\end{aligned} \tag{4.20}$$

The glitch affects the amplitude and phase of the wave, because if we set $\epsilon = 0$ and $D = 1$ we recover the homogeneous frequency solutions (see Equation 4.4). Unfortunately, solving Equation 4.20 for ω is not possible analytically. However, we can find eigenfrequencies (or oscillation modes) numerically.

Let us use ω'_n to denote the solutions to Equation 4.20, where n is a positive integer. Using dimensionless units of length and time, we set $c = 1$, $L = 1$, and test several values of x_g, δx, and δc. We find ω'_n by solving Equation 4.20 using Newton's method for $n = 1,\ldots,50$ modes. Initial guesses for ω'_n are obtained from the homogeneous medium solutions (ω_n) in Equation 4.4. We show the difference between the glitch solutions and those from the homogeneous

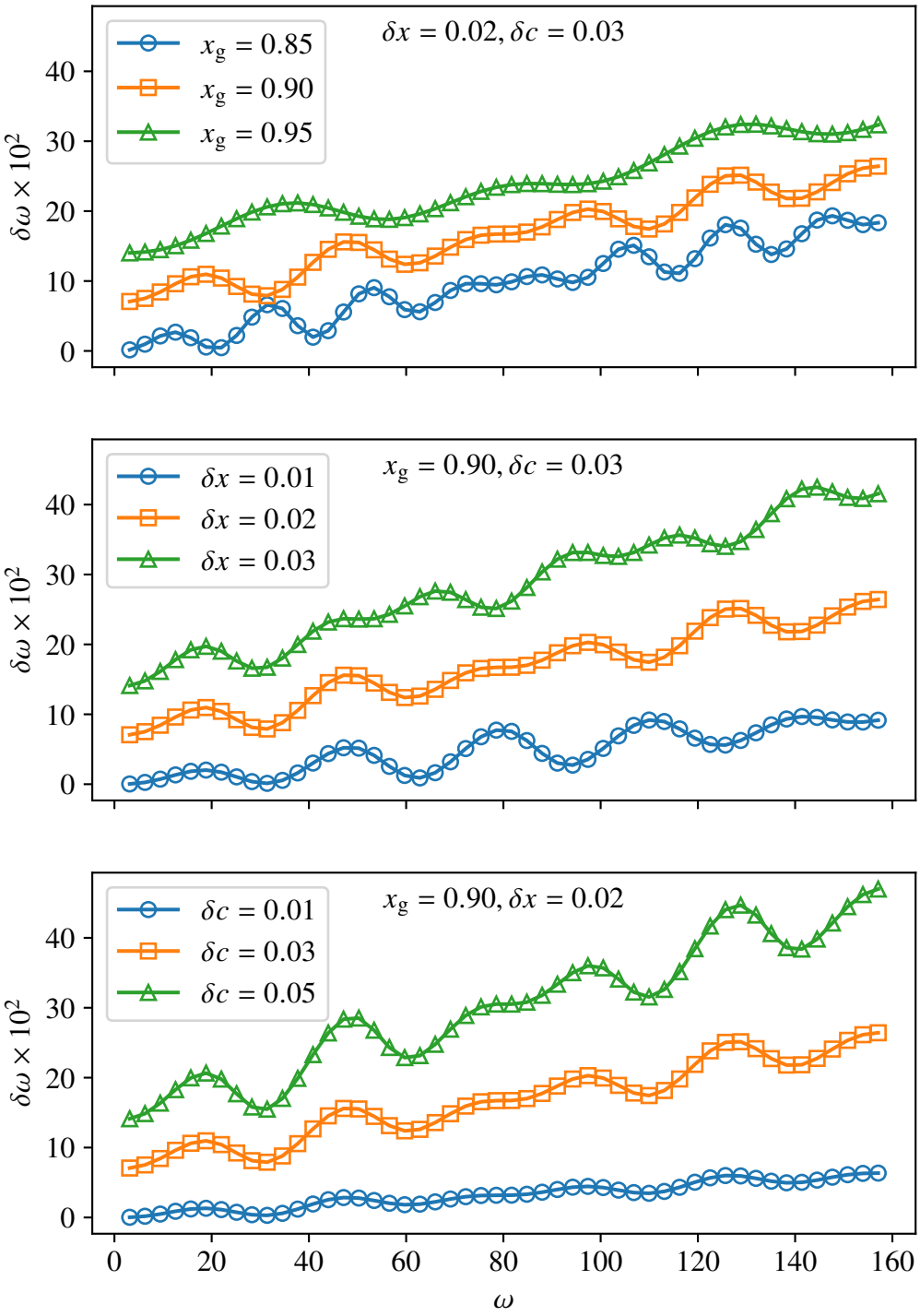


Figure 4.2: The change in mode frequency induced by a change in sound speed of δc from $x_g - \delta x$ to $x_g + \delta x$ in a one-dimensional medium, bound such that $x \in [0, 1]$ (see Figure 4.1). Outside the perturbation the speed of sound, $c = 1$. The frequency changes are offset by 0.00, 0.07, and 0.14 for the circle, square, and triangle points respectively. Points are joined by straight lines to guide the eye, but these lines do not represent real solutions.

medium, $\delta\omega_n = \omega'_n - \omega_n$, in Figure 4.2. We can see a periodic component to $\delta\omega$ induced by the glitch. As the wave nodes pass in and out of the region with changing n, the sensitivity of the wave to the glitch varies periodically. The overall sensitivity to the glitch depends on how much the wave changes inside the glitch, hence why low n modes have smaller $\delta\omega$.

The functional form of $\delta\omega$ appears to have a linear component and a short periodicity modulated by a longer periodicity. As the location of the glitch (x_g) gets smaller, the short period of $\delta\omega$ decreases. If we imagine the spacial distribution of nodes in the system as a function of n, the density of nodes is larger towards the centre of the system. The periodicity arises from the nodes passing in and out of the glitch region with changing n. Therefore, where the density of wave nodes is higher, we expect the short period of $\delta\omega$ to decrease. Similarly, as the half-width of the glitch (δx) increases, the longer period increases.

Furthermore, an increasing change in sound speed (δc) increases the amplitude of $\delta\omega$. This result is intuitive, as we expect a larger change in c to correspond to a large change in ω. Finally, increasing both δx and δc increases the slope of $\delta\omega$. The sensitivity of a mode to the glitch increases with n and depends on how much the wave changes in the glitch region. Modes of smaller n become more sensitive to the glitch region as δx and δc increase, thus increasing the linear slope of $\delta\omega$.

The small phase offset ϵ in Equation 4.19 is required for the wave function to satisfy the boundary conditions at $x = 0$. However, adding ϵ shifts the effective location of x_g — it changes the scale of the x-axis by a factor of $(1 + \epsilon)$. We plot ϵ against ω in Figure 4.3 and show that its magnitude is $\sim 10^{-4}$, much smaller than the location and size of region 2. The periodicity caused by the glitch also shows up in Figure 4.3, with its properties affected similarly to Figure 4.2. This is because the more a mode is affected by the glitch, the greater the phase offset required to satisfy the boundary conditions.

Finding an approximate solution for $\delta\omega$ is beyond the scope of this example. However, we can show that by modelling $\delta\omega$, we can recover information about the structural glitch. Let us

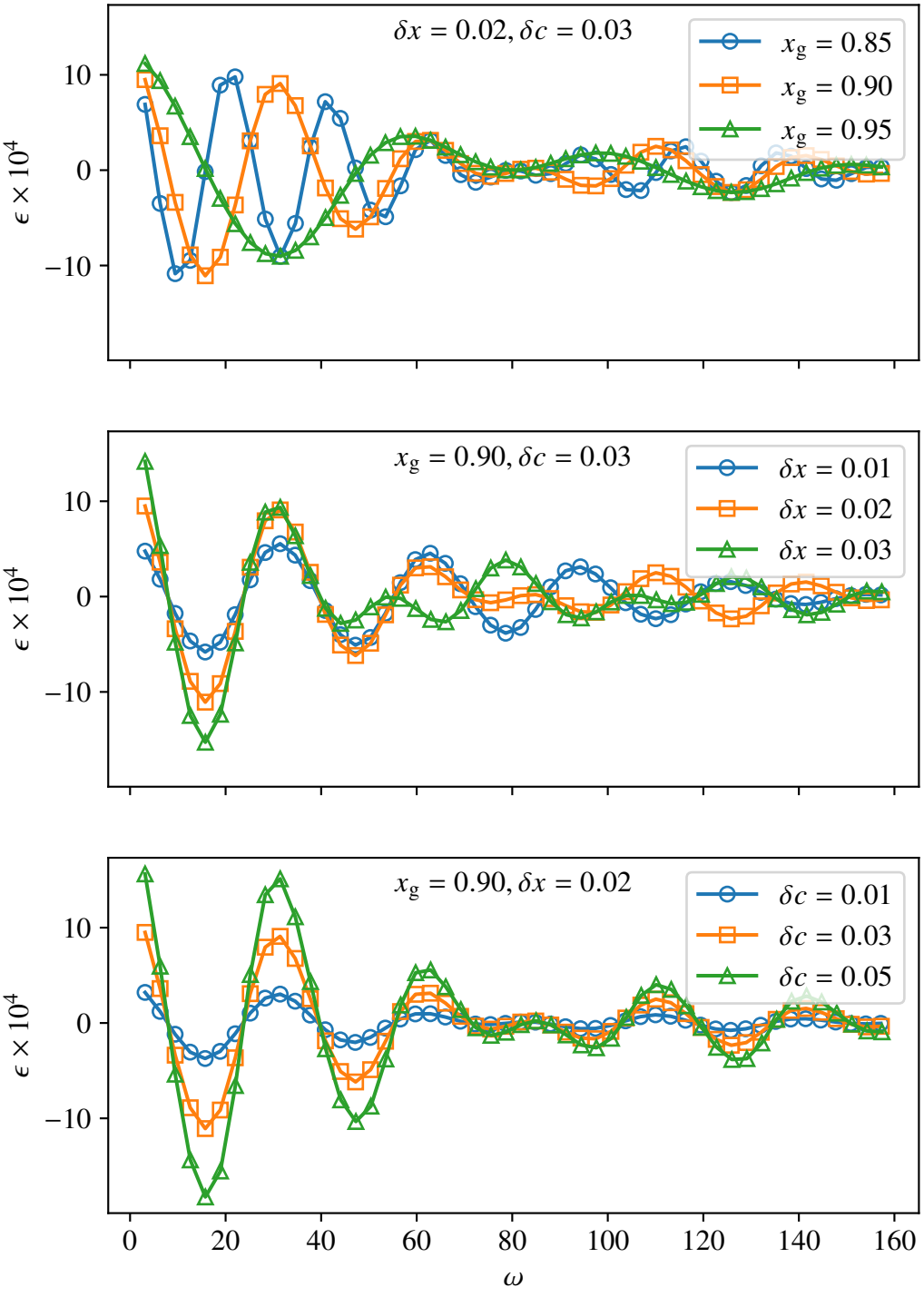


Figure 4.3: The same as Figure 4.2 but showing the phase offset ϵ required to satisfy the boundary conditions.

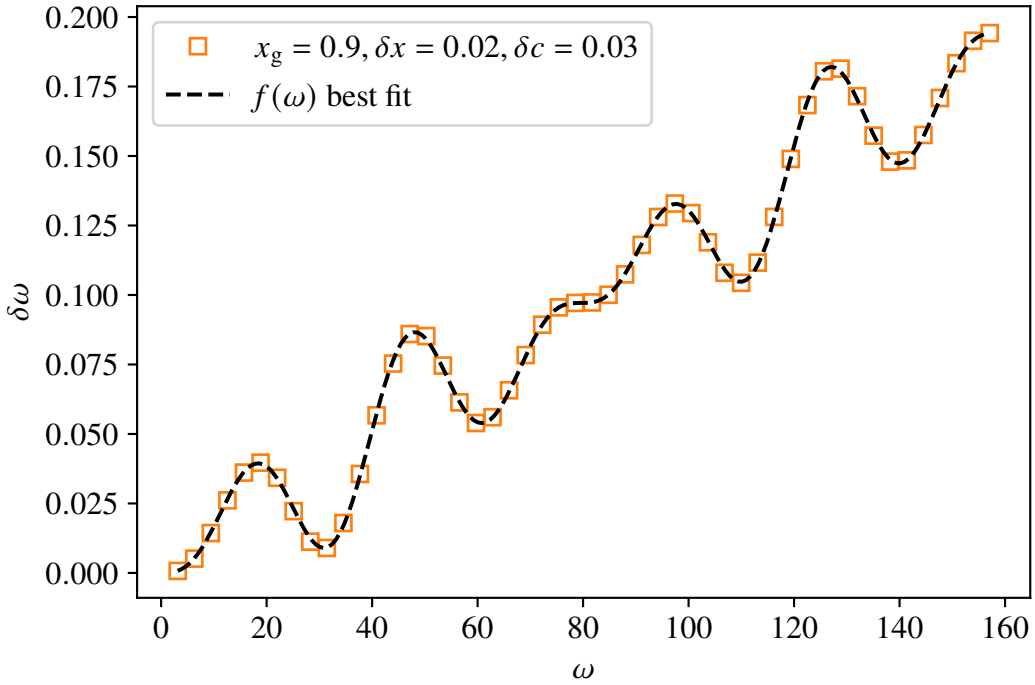


Figure 4.4: Best fit of Equation 4.21 to $\delta\omega_n$ for $n = 1, \ldots, 50$, $L = 1$, and $c = 1$.

build a model $\delta\omega = f(\omega)$. Looking at Figure 4.2, we propose a form for f,

$$f(\omega) = a_1\omega - a_2 \sin(\tau_1\omega)\cos(\tau_2\omega), \tag{4.21}$$

where a_1 and a_2 are coefficients which are both functions of δx and δc. Parameters τ_1 and τ_2 are the 'frequencies' of the periodic component to $\delta\omega$. Given Figure 4.2, we expect τ_1 and τ_2 to be related to δx and x_g respectively.

We fit Equation 4.21 to $\delta\omega_n$ obtained from a glitch located at $x_g = 0.9$, with half-width $\delta x = 0.02$, and change in sound speed $\delta c = 0.03$. The best fitting line is shown in Figure 4.4. We found $\tau_1 \approx 0.0389$ and $\tau_2 \approx 0.199$. The former corresponds to $\tau_1 \simeq 2\delta\tau$, where $\delta\tau$ is half the acoustic width of the glitch — the time at which sound takes to transverse the region. Since the speed of sound $c \approx 1$ throughout the medium, the true value of $\delta\tau \approx 0.02$. The latter 'frequency' corresponds to $\tau_2 \simeq 2\tau_g$, where τ_g is the sound travel time from the nearest edge to the centre of the glitch (in this case $\tau_g \approx 0.1$). Referring back to Figure 4.2, we can see how these relations should hold for different values of x_g, δx, and δc providing the glitch region is

small.

It would be tempting to test this further, but we have shown that the glitch signature in ω'_n relates to the properties of the glitch. Although the structure of a star is more complicated than this example, we can extend this principle to find glitch signatures in the mode frequencies of solar-like oscillators. Therefore, we will build upon this analogy in the next section where we review some examples of acoustic glitches in stars.

4.3 Acoustic Glitch Signatures in Stellar Oscillations

In the previous section, we considered a glitch in a homogeneous medium, where the speed of sound is constant everywhere else. In a star, the adiabatic sound speed is not constant. It depends on the density (ρ) and pressure (P),

$$c^2 = \gamma \frac{P}{\rho}, \tag{4.22}$$

where γ is the first adiabatic exponent,

$$\gamma = \left(\frac{\partial \ln P}{\partial \ln \rho} \right)_S , \tag{4.23}$$

at constant entropy, S. This is the first of three adiabatic exponents ($\Gamma_1, \Gamma_2, \Gamma_3$) introduced by Chandrasekhar (1939) to describe the non-ideal gas inside a star. However, in this thesis we do not use the other two and hence refer the first as $\gamma \equiv \Gamma_1$.

For the most part, γ, P, and ρ change smoothly with radius inside a star. However, a rapid structural glitch in these quantities would lead to a sudden change in sound speed. In the previous section, we showed how such a perturbation can lead to a periodicity in the eigenfrequencies for an otherwise homogeneous medium. Characterising this signal allowed us to measure the properties of the glitch. If similar glitches were present in a star, then we might be able to do the same. In this section, we explore the origins of glitches inside a solar-like star. Then, we see what effect these have on the eigenfrequencies, a quantity we can measure through asteroseismology.

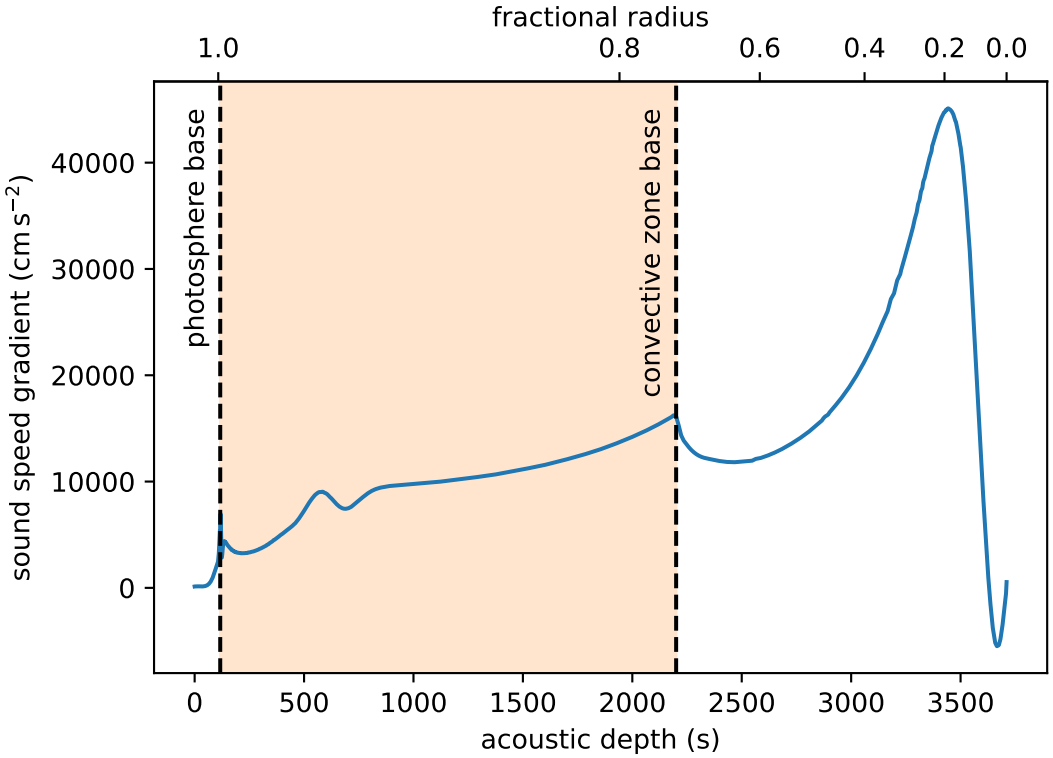


Figure 4.5: The sound speed gradient ($\mathrm{d}c/\mathrm{d}\tau$) of model R plot against the acoustic depth (τ). The fractional radius to the photosphere base is given on the top axis. The convective envelope is shaded and the bases of the photosphere and convective zone are marked with dashed lines.

Firstly, let us consider the sound speed profile of a Sun-like star. Particularly, we want to see how the sound speed changes on the timescale of a pressure wave travelling through the star. As discovered in Section 4.2, a convenient timescale to work with is the acoustic depth, τ. This is not to be confused with the symbol for stellar age in Chapter 3. Here, we define τ as the time taken for a pressure wave to travel from the surface (R) to some radius (r) in a star under the assumption of spherical symmetry,

$$\tau(r) = \tau_0 - \int_0^r \frac{\mathrm{d}r'}{c(r')}, \tag{4.24}$$

where τ_0 is the acoustic radius of the star. We recall from Section 1.2 that $\nu_0 \equiv (2\tau_0)^{-1}$ may be approximated by the large frequency separation ($\Delta\nu$) in the asymptotic limit that $l/n \to 0$.

In Figure 4.5, we show the sound speed gradient with respect to τ for a Sun-like stellar model (model R; see Section 4.3.1). We see how the speed of sound changes smoothly throughout

the star. In the convective zone, there is a noticeable wiggle around 700 s and a sharp change in direction at its base. The first is caused by the ionisation of helium, which we will explore further in Section 4.3.2. The second is due to a discontinuity in the temperature gradient as the dominant energy transport mechanism goes from convective to radiative. This will be discussed in Section 4.3.3.

4.3.1 Sun-Like Model Star

We computed a representative Sun-like model star (hereafter 'model R') using MESA (version 12115; Paxton et al., 2011, 2013, 2015, 2018, 2019; Jermyn et al., 2023). The model was computed with a mass of $1.0\,M_\odot$ to a central fractional hydrogen abundance of $X_c = 0.6$. We used initial fractional helium and heavy-element abundances of $Y_{\rm init} = 0.28$ and $Z_{\rm init} = 0.02$ respectively. The mixing-length was parametrised by $\alpha_{\rm MLT} = 1.9$ and a turbulent pressure factor of 1. We evolved the star using element diffusion with MESA's default coefficients (Stanton and Murillo, 2016). In addition, we used MESA's default Grevesse and Sauval (1998) solar chemical composition and opacity tables. We also output pulsation profile data in FGONG format to later use when computing oscillation modes in Chapter 5. Our resulting model R had an age of 4.073 Gyr, $T_{\rm eff} \approx 5682\,$K, and $R \approx 1.014\,R_\odot$. Since the model was evolved with element diffusion, the surface fractional helium and heavy-element were approximately $Y_{\rm surf} \approx 0.2515$ and $Z_{\rm surf} \approx 0.01810$ respectively.

4.3.2 Helium Ionisation Glitch

In this section, we will first show how the sound speed inside a solar-like star is affected by the ionisation of hydrogen and helium. Then, we will see that an increase in helium abundance (Y) increases the effect of helium ionisation on the speed of sound. Starting with the variational principle, we derive the commonly used equation for the glitch signature ($\delta\nu$) induced by the second ionisation of helium from Houdek and Gough (2007).

In the outer convective envelope of Sun-like stars, the temperature (10^4 to 10^5 K) and density (10^{-5} to $10^{-3}\,$g cm^{-3}) is suitable for ionising hydrogen and helium (Eggleton, Faulkner, and

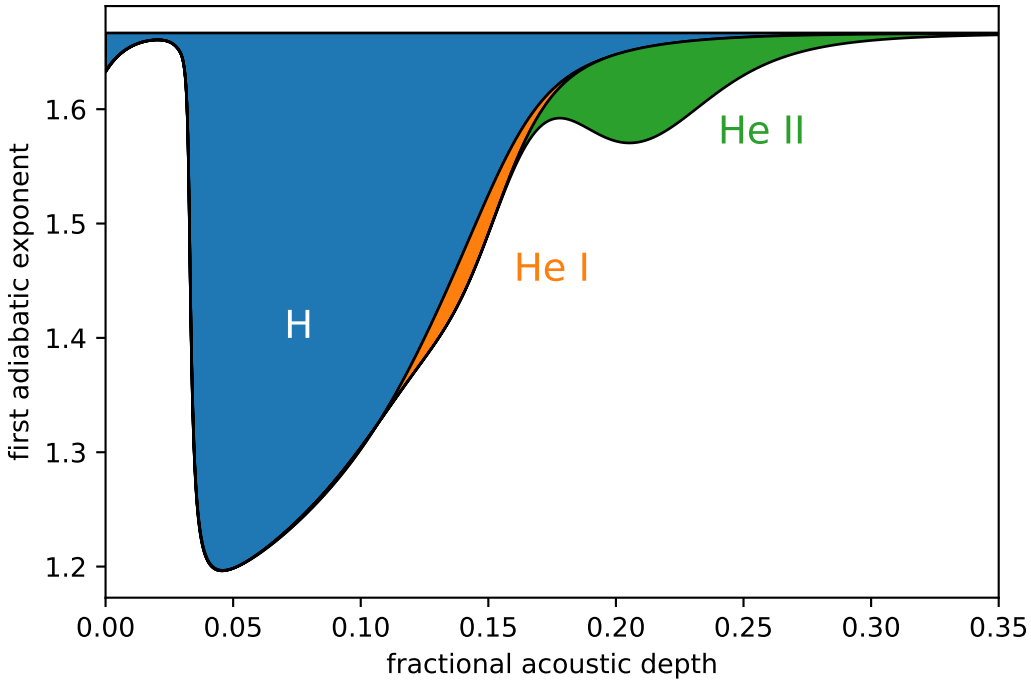


Figure 4.6: The depressions in the first adiabatic exponent (γ) caused by the ionisation of hydrogen (H), and the first and second ionisations of helium (He I and He II). The horizontal axis is the fractional acoustic depth from the surface of the star, τ/τ_0. The shading highlights relative contributions to the depression in γ from each ionisation process.

Flannery, 1973). As a given chemical species in the star ionises, the number of particles and hence chemical potential of the species changes. This induces a gradient in the thermal free energy of the gas which relates to the pressure and entropy of the gas. Thus, we expect ionisation to cause a change in the pressure-density gradient at constant entropy (γ). Since γ relates to the sound speed from Equation 4.22, any sudden change in γ will induce a change in c.

In Figure 4.6, we show γ for model R against fractional acoustic depth, shaded by regions of ionisation. For an ideal monatomic gas, $\gamma = 5/3$, but we see that $\gamma < 5/3$ in regions where helium and hydrogen ionise. Close to the surface of the star, hydrogen ionisation has the largest effect on γ because it makes up the majority of stellar matter. The first (He I) and second (He II) ionisations of helium occur deeper in the star. We can see that the second ionisation of helium has a greater effect on γ than the first. He II ionisation causes a rapid change in γ over a few per cent in τ (equivalent to ~ 100 s in model R).

The Effect of Helium Abundance on γ

We were able to isolate the contributions to γ from ionisation using a recent derivation by Houdayer et al. (2021). They obtained an approximate formula for γ as a function of temperature (T) and density, valid for the outer convective zone of Sun-like stars. Using their formula, we plot γ for a range of T and ρ in Figure 4.7. The first panel illustrates the three ionisation regions of hydrogen and helium. The value of γ decreases when the ionisation reaction is occurring. The magnitude of this effect depends on the intersection of the temperature-density profile of the star (over-plot for model R). We can imagine a hotter star shifting this line such that ionisation occurs closer to the surface. Along the model R profile, we see that an increase in helium abundance by 0.1 decreases γ by up to ~ 0.03 in the He II ionisation region. This shows that helium abundance correlates with the depth of the glitch in γ. However, there is clearly still some dependence on T and ρ which are governed by mass, chemical composition, and evolutionary phase.

To show the effect of small changes in Y on the sound speed, we plot the sound speed gradient in Figure 4.8. The three models shown were evolved to the same $X_c = 0.6$ as model R with Y_{init} of 0.26, 0.28, and, 0.3. The dominant effect of helium abundance appears as a Gaussian-like depression in γ around the second ionisation of helium. We see how larger helium abundance increases the width and depression in γ, which is reflected in the sound speed gradient. A larger Y also leads to a relative reduction in hydrogen abundance (X) which shrinks the width of the hydrogen ionisation region.

The Effect of a Change in γ on p Mode Frequencies

To see how a change in γ affects the mode frequencies, we explore the sensitivity of a mode to rapid structural changes in the star. Starting with the variational principle, we can approximate the characteristic frequencies of a spherically symmetric, slowly rotating star by (Chandrasekhar, 1964),

$$\omega^2 \simeq \frac{\mathcal{E}}{\mathcal{I}} \tag{4.25}$$

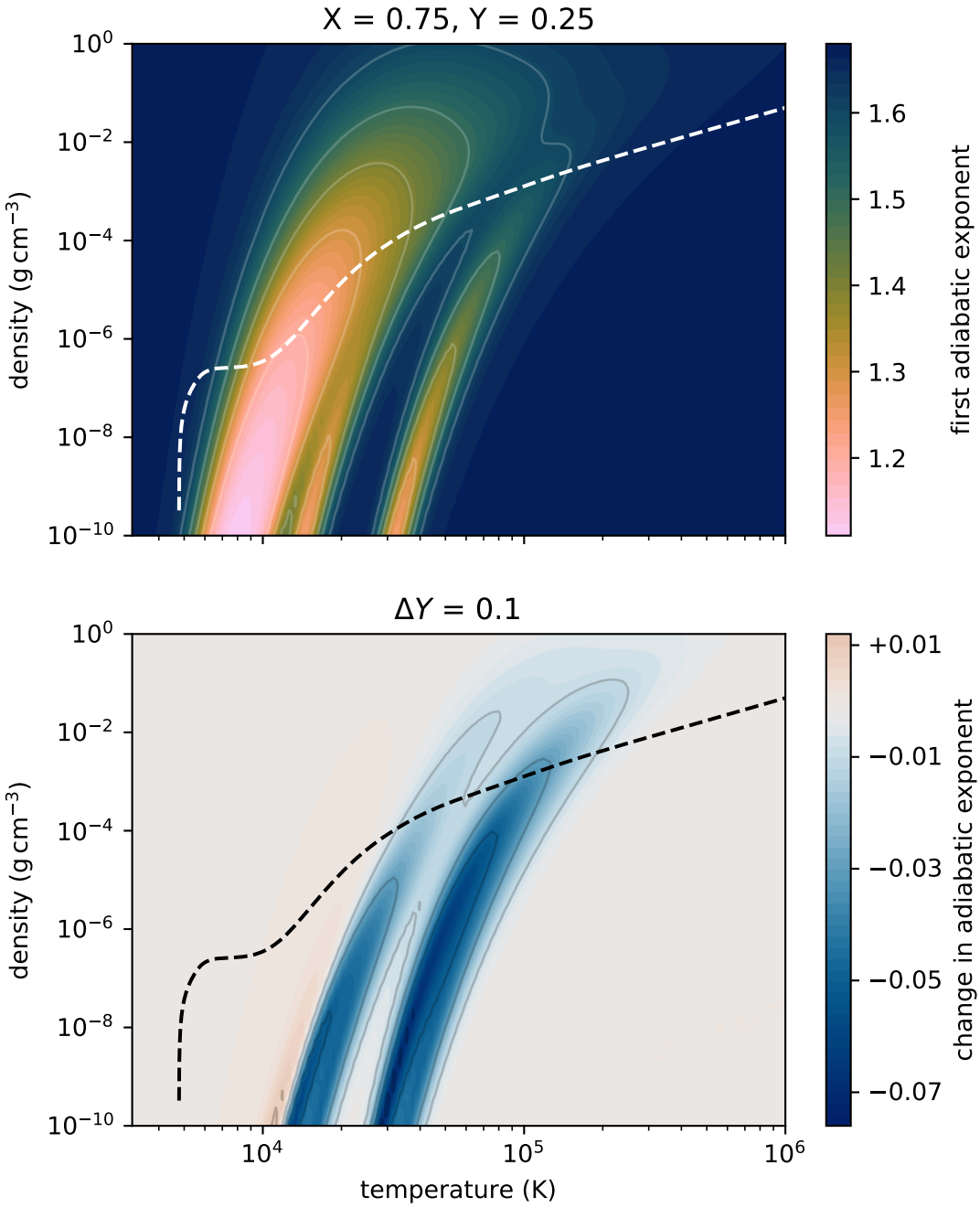


Figure 4.7: Temperature-density profile for a Sun-like star. *Top:* The first adiabatic exponent γ as a function of temperature and density, calculated using Eq. 53 of Houdayer et al. (2021) for a helium mass fraction, $Y = 0.25$. *Bottom:* The change in γ induced by a change in helium abundance, $\Delta Y = 0.1$. In both panels, the dashed line shows the temperature-density profile of model R. *This figure is a recreation of Fig. 5 in Houdayer et al. (2021).*

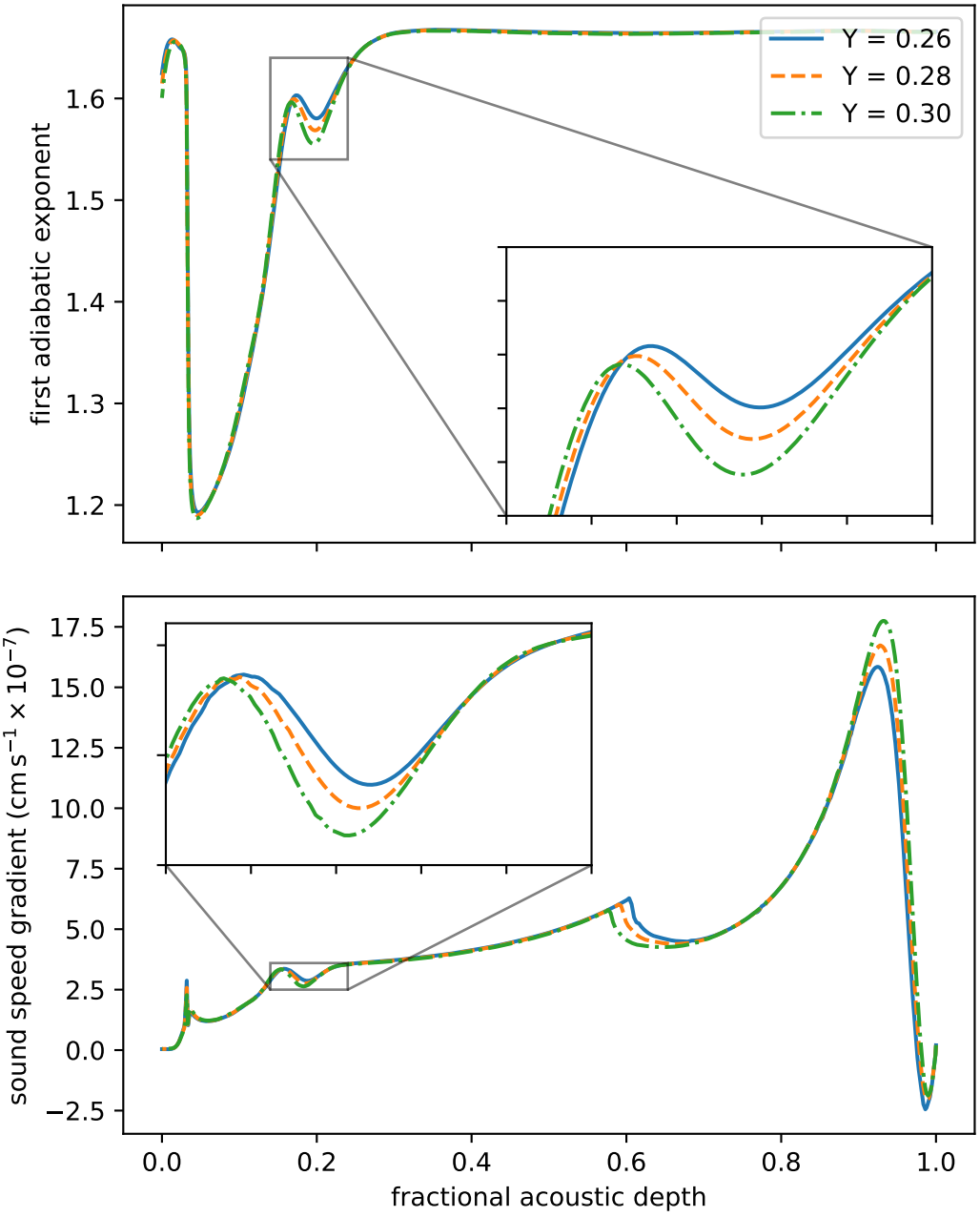


Figure 4.8: The effect on the sound speed profile for three solar-like stars with initial helium mass fractions of 0.26 (*solid*), 0.28 (*dashed*), and 0.3 (*dot-dashed*). The models were each evolved to a central helium mass fraction of 0.6. *Top:* The first adiabatic exponent γ. *Bottom:* The sound speed gradient $\mathrm{d}c/\mathrm{d}t$ where $t = \tau/\tau_0$ is the fractional acoustic depth.

which is a ratio of $\mathcal{E}$ (proportional to the mode's energy) to its inertia ($\mathcal{I}$). The mode inertia is given by,

$$\mathcal{I} = \int_0^R \boldsymbol{\xi} \cdot \boldsymbol{\xi}\, \rho r^2 \, \mathrm{d}r, \tag{4.26}$$

where the amplitude and direction of a pressure wave at a given point in a star is given by the so-called Lagrangian perturbation vector $\boldsymbol{\xi}$.

A structural glitch in the star would cause a small change in frequency ($\delta\omega$). Differentiating both sides of Equation 4.25 with respect to ω, we get,

$$\begin{aligned} 2\omega &= \frac{1}{\mathcal{I}}\left(\frac{\delta\mathcal{E}}{\delta\omega} - \omega^2 \frac{\delta\mathcal{I}}{\delta\omega}\right), \qquad \left[\times \frac{\delta\omega}{2\omega^2}\right] \\ \frac{\delta\omega}{\omega} &= \frac{1}{2\,\mathcal{I}}\left(\frac{\delta\mathcal{E}}{\omega^2} - \delta\mathcal{I}\right). \end{aligned} \tag{4.27}$$

Small changes $\delta\mathcal{E}$ and $\delta\mathcal{I}$ depend on changes to the state variables. By perturbing and substituting appropriate forms of $\mathcal{E}$ and $\mathcal{I}$, it is possible to rewrite this as a function of the so-called structural kernels $\mathcal{K}_{a,b}$. For example (Christensen-Dalsgaard, 2014),

$$\frac{\delta\omega}{\omega} = \int_0^R \left(\mathcal{K}_{c^2,\rho}\frac{\delta c^2}{c^2} + \mathcal{K}_{\rho,c^2}\frac{\delta\rho}{\rho}\right)\mathrm{d}r. \tag{4.28}$$

where $\mathcal{K}_{a,b}$ gives the relative effect on ω at a given r due to a perturbation in a state variable a at fixed b. The kernels, defined fully in Gough and Thompson (1991), can show the sensitivity of the wave to a change in state in the star. We plot $\mathcal{K}_{c^2,\rho}$ and $\mathcal{K}_{\rho,c^2}$ in Figure 4.9 for a few radial oscillation modes. Both decay through the star, meaning that the modes are less sensitive to structural changes deeper in the star.

We can evaluate Equation 4.28 to find the effect on ω due to a change in γ from helium ionisation. The first term can be rewritten in terms of a change in γ by perturbing Equation 4.22 and substituting into Equation 4.28. It happens that $\mathcal{K}_{\gamma,\rho}\delta\gamma/\gamma \equiv \mathcal{K}_{c^2,\rho}\delta c^2/c^2$. If we assume helium ionisation has a negligible effect on density, then $\delta\rho/\rho \approx 0$ in the ionisation region.

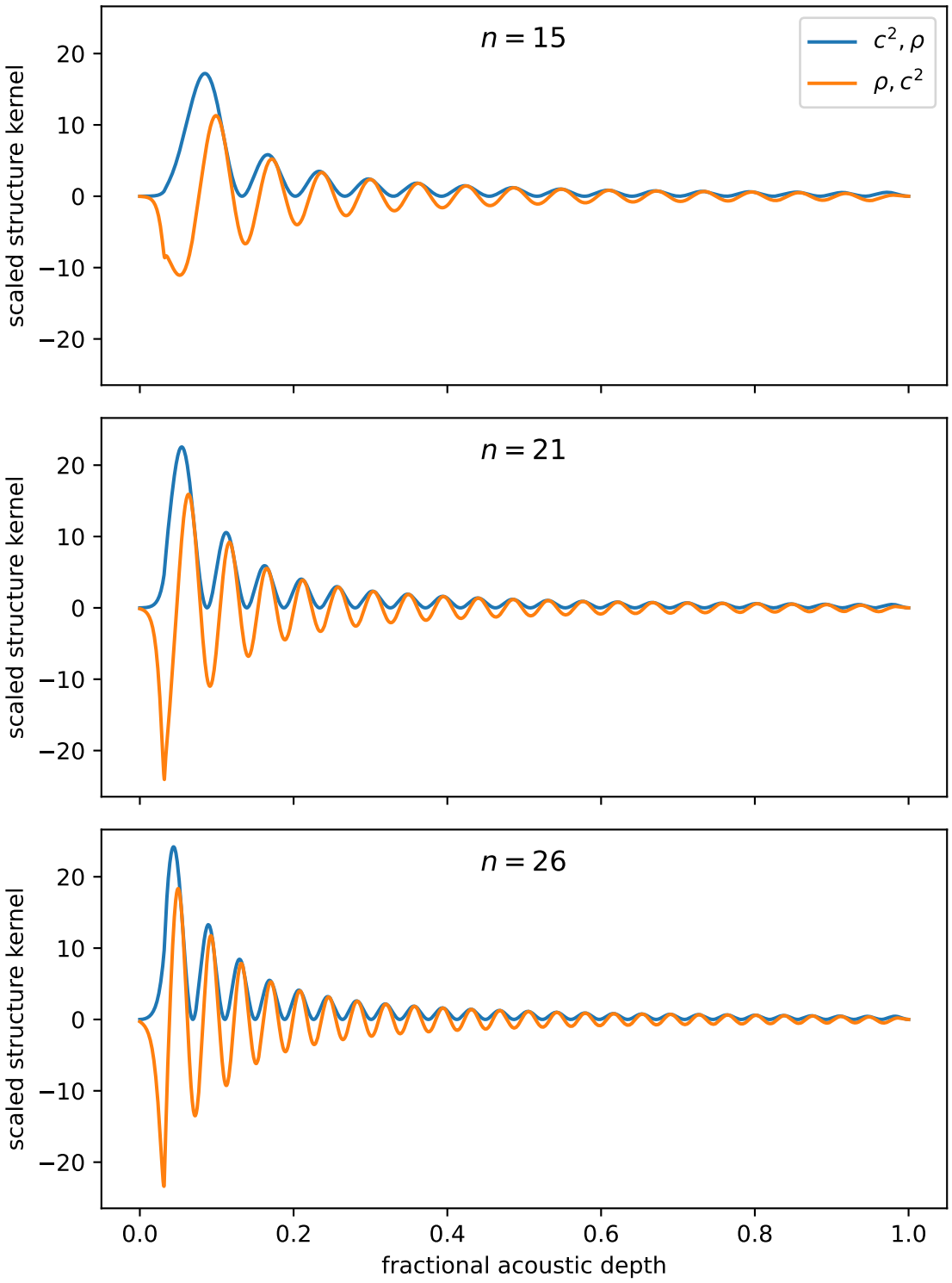


Figure 4.9: Structure kernels ($\mathcal{K}_{c^2,\rho}, \mathcal{K}_{\rho,c^2}$) scaled by the stellar radius as a function of fractional acoustic depth. The kernels were computed for three radial oscillation modes ($n = 15, 21, 26$) using oscillation calculations for stellar model R.

Therefore, we may rewrite Equation 4.28 due to a change in γ as,

$$\left.\frac{\delta\omega}{\omega}\right|_{\gamma} \simeq \int_0^R \mathcal{K}_{\gamma,\rho}\frac{\delta\gamma}{\gamma}\mathrm{d}r, \tag{4.29}$$

where $\mathcal{K}_{\gamma,\rho}$ is shown in Gough (1993) to satisfy,

$$\omega^2\mathcal{I}\mathcal{K}_{\gamma,\rho} = \frac{1}{2}\gamma P(\mathrm{div}\boldsymbol{\xi})^2 r^2. \tag{4.30}$$

In order to solve Equation 4.29, we first expand the mode inertia, $\mathcal{I}$. The Lagrangian perturbation vector can be written in terms of a radial and horizontal component, $\boldsymbol{\xi} = \xi_r\hat{r} + \boldsymbol{\xi}_h$. For low angular degree modes, the horizontal component $\boldsymbol{\xi}_h$ is negligible. Therefore, we can rewrite the mode inertia as,

$$\mathcal{I} \simeq \int_0^R \xi_r^2\rho r^2\,\mathrm{d}r. \tag{4.31}$$

Then, we use a suitable approximation of ξ_r from Gough (1993),

$$\xi_r \simeq \frac{\mathcal{A}}{r}\sqrt{\frac{K}{\rho}}\cos\psi, \tag{4.32}$$

where $\mathcal{A}$ is a proportionality constant. The radial wave number $K \simeq \omega/c$ for high-order acoustic modes. Assuming τ is not close to the upper turning point (where the wave is reflected near the stellar surface), the phase term $\psi \simeq \omega\tau + \epsilon$. Here, the small offset ϵ is a slowly varying function of τ. Substituting Equation 4.32 into the mode inertia gives,

$$\begin{aligned}\mathcal{I} &\simeq \mathcal{A}^2\int_0^R K\cos^2\psi\,\mathrm{d}r,\\ &= \frac{1}{2}\omega\mathcal{A}^2\int_0^R(1+\cos 2\psi)\frac{\mathrm{d}r}{c}.\end{aligned} \tag{4.33}$$

Finally, changing to an integral over the acoustic depth (τ) we can evaluate the integral for high-order modes where $\omega_n \ll \tau_0^{-1}$,

$$\mathcal{I} \simeq \frac{1}{2}\omega\mathcal{A}^2\int_0^{\tau_0}[1+\cos 2(\omega\tau+\epsilon)]\,\mathrm{d}\tau \simeq \frac{1}{2}\omega\mathcal{A}^2\tau_0. \tag{4.34}$$

Secondly, we substitute an approximation for $(\mathrm{div}\boldsymbol{\xi})^2$ from Gough (1993),

$$(\mathrm{div}\boldsymbol{\xi})^2 \simeq \frac{\mathcal{A}^2\omega^3}{\gamma P c r^2}\sin^2\psi, \tag{4.35}$$

along with $\mathcal{K}_{\gamma,\rho}$ from Equation 4.30 into Equation 4.29 to get,

$$\begin{aligned}\left.\frac{\delta\omega}{\omega}\right|_\gamma &\simeq \frac{1}{2\omega^2\mathcal{I}}\int_0^R \delta\gamma P(\mathrm{div}\boldsymbol{\xi})^2 r^2\,\mathrm{d}r,\\ &\simeq \frac{\omega\mathcal{A}^2}{2\mathcal{I}}\int_0^R \frac{\delta\gamma}{\gamma}\sin^2\psi\frac{\mathrm{d}r}{c},\\ &= \frac{\omega\mathcal{A}^2}{4\mathcal{I}}\int_0^R \frac{\delta\gamma}{\gamma}(1-\cos 2\psi)\frac{\mathrm{d}r}{c}.\end{aligned} \tag{4.36}$$

We can see that $\delta\omega$ is split into a smooth and a periodic component. The smooth component may be treated later, but for now we focus on the oscillating part of $\delta\omega$. Substituting $\mathcal{I}$ from Equation 4.34 and changing to an integral over acoustic depth we get,

$$\left.\frac{\delta\omega}{\omega}\right|_{\gamma,\mathrm{osc}} \simeq -\frac{1}{2\tau_0}\int_0^{\tau_0}\frac{\delta\gamma}{\gamma}\cos 2(\omega\tau+\epsilon)\,\mathrm{d}\tau. \tag{4.37}$$

A Functional Form of the Helium Glitch Signature

We have shown that a change in γ can induce a periodicity in ω. The functional from of this periodicity depends on $\delta\gamma/\gamma$. As shown in Figures 4.6 and 4.7, the dominant change in γ due to a change in helium abundance is from the second ionisation of helium. There have been different attempts to approximate $\delta\gamma/\gamma\,|_{\mathrm{II}}$ in the literature, for example using a Dirac delta function or a triangular function (Monteiro, Christensen-Dalsgaard, and Thompson, 1994; Monteiro and Thompson, 2005). In recent years, work modelling the glitch has used the formulation from Houdek and Gough (2007) where the change in γ due to He II ionisation is modelled with a Gaussian shape,

$$\left.\frac{\delta\gamma}{\gamma}\right|_{\mathrm{II}} \simeq -\frac{\Gamma_{\mathrm{II}}}{\Delta_{\mathrm{II}}\sqrt{2\pi}}\,\mathrm{e}^{-\frac{1}{2}(\tau-\tau_{\mathrm{II}})^2/\Delta_{\mathrm{II}}^2}, \tag{4.38}$$

where Γ_{II} is the area, Δ_{II} is the characteristic width, and τ_{II} is the centre of the ionisation region.

We will verify the result from Houdek and Gough (2007) for $\delta\omega$ due to He II ionisation. Substituting Equation 4.38 into Equation 4.37 with a change of variables to $x = (\tau - \tau_{\rm II})/\Delta_{\rm II}$, gives,

$$\left.\frac{\delta\omega}{\omega}\right|_{\rm II,osc} \simeq \frac{\Gamma_{\rm II}}{2\sqrt{2\pi}\,\tau_0} \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2} \cos 2(\Delta_{\rm II}\omega x + \widetilde{\epsilon_{\rm II}})\,\mathrm{d}x \tag{4.39}$$

where $\widetilde{\epsilon_{\rm II}} = \omega\tau_{\rm II} + \epsilon_{\rm II}$ and the phase $\epsilon = \epsilon_{\rm II}$ is assumed constant across the glitch region. We can solve the above integral analytically using differentiation under the integral sign by introducing an arbitrary variable a,

$$I(a) = \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2} \cos 2(\Delta_{\rm II}\omega x a + \widetilde{\epsilon_{\rm II}})\,\mathrm{d}x \tag{4.40}$$

Differentiating this with respect to a, and then integrating by parts gets,

$$\begin{aligned}
I'(a) &= -2\Delta_{\rm II}\omega \int_{-\infty}^{\infty} x\,\mathrm{e}^{-x^2/2} \sin 2(\Delta_{\rm II}\omega x a + \widetilde{\epsilon_{\rm II}})\,\mathrm{d}x, \\
&= -2\Delta_{\rm II}\omega \int_{-\infty}^{\infty} \sin 2(\Delta_{\rm II}\omega x a + \widetilde{\epsilon_{\rm II}})\,\mathrm{d}\mathrm{e}^{-x^2/2}, \\
&= 2\Delta_{\rm II}\omega \left\{ \left[\mathrm{e}^{-x^2/2} \sin 2(\Delta_{\rm II}\omega x a + \widetilde{\epsilon_{\rm II}}) \right]_{-\infty}^{\infty} - 2\Delta_{\rm II}\omega a \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2} \cos 2(\Delta_{\rm II}\omega x a + \widetilde{\epsilon_{\rm II}})\,\mathrm{d}x \right\}, \\
&= -4\Delta_{\rm II}^2\omega^2 a\, I(a),
\end{aligned} \tag{4.41}$$

which is a differential equation with the solution,

$$\begin{aligned}
I(a) = I(0)\mathrm{e}^{-2\Delta_{\rm II}^2\omega^2 a^2}; \qquad I(0) &= \int_{-\infty}^{\infty} \mathrm{e}^{-x^2/2} \cos 2\widetilde{\epsilon_{\rm II}}\,\mathrm{d}x, \\
&= \sqrt{2\pi}\,\cos 2\widetilde{\epsilon_{\rm II}},
\end{aligned} \tag{4.42}$$

Therefore, if we set $a = 1$ (to get back to the integral in Equation 4.39) we get the following for the periodic glitch signature (cf. Houdek and Gough, 2007, Eq. 15),

$$\left.\frac{\delta\omega}{\omega}\right|_{\rm II,osc} \simeq \frac{\Gamma_{\rm II}}{2\tau_0} \mathrm{e}^{-2\Delta_{\rm II}^2\omega^2} \cos 2(\tau_{\rm II}\omega + \epsilon_{\rm II}), \tag{4.43}$$

We can rearrange this to a more familiar form in terms of cyclic frequency (e.g. Verma et al.,

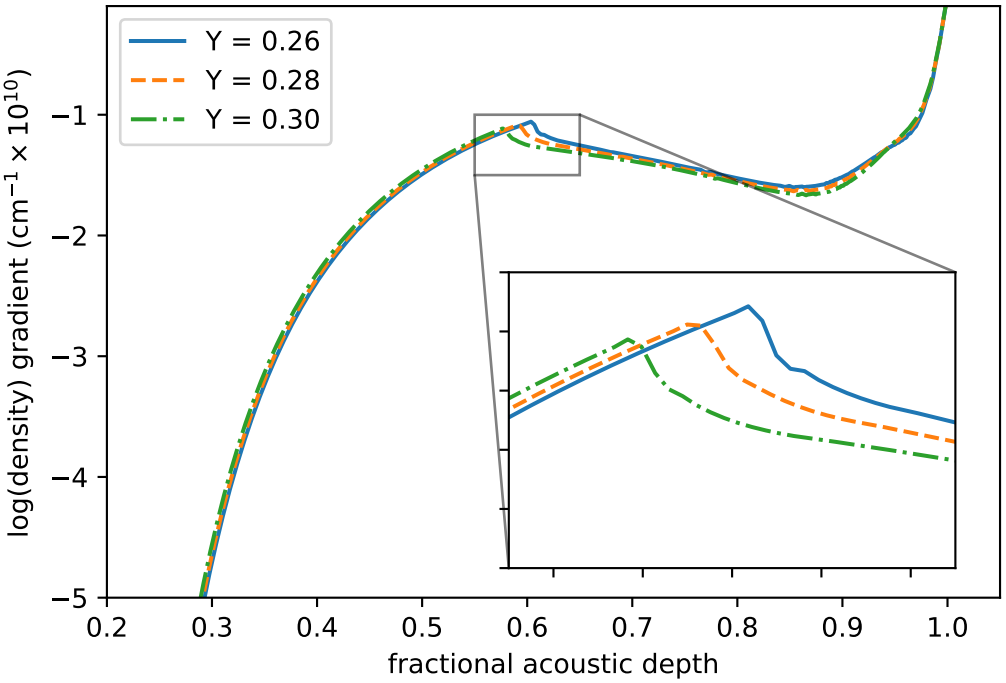


Figure 4.10: Discontinuity in the density gradient at the base of the convective zone for three solar-like stars with initial helium mass fractions of 0.26 (*solid*), 0.28 (*dashed*), and 0.3 (*dot-dashed*).

2014, 2017), by substituting $\omega = 2\pi\nu$,

$$\left.\frac{\delta\nu}{\nu}\right|_{\mathrm{II,osc}} \simeq \nu_0 \Gamma_{\mathrm{II}}\, \mathrm{e}^{-8\pi^2\Delta_{\mathrm{II}}^2\nu^2}\, \sin(4\pi\tau_{\mathrm{II}}\nu + \phi_{\mathrm{II}}), \tag{4.44}$$

where $\phi_{\mathrm{II}} = 2(\epsilon_{\mathrm{II}} + \pi/4)$, and $\nu_0 = (2\tau_0)^{-1}$ is the inverse acoustic diameter of the star.

4.3.3 Base of the Convective Zone Glitch

The sensitivity of p modes to glitches gets smaller further into the star. However, we should not neglect the effect of the near-discontinuity at the base of the convective zone (BCZ). This arises from a jump in the second derivative of temperature at the BCZ which translated to discontinuities in the second derivative of density and sound speed. In Figure 4.10 we can see a fast change in direction of the density gradient around $\tau/\tau_0 \approx 0.6$ (or at about 2200 s for model R). We also see that the relative location of the BCZ changes a little with initial Y.

We could take a similar approach to the previous section, now considering the second kernel

$\mathcal{K}_{\rho,c^2}$ since the discontinuity affects ρ. However, Houdek and Gough (2007) take another approach by considering the discontinuity as it appears in the acoustic cut-off frequency, ω_{ac}. The details of this derivation are beyond the scope of this work. Instead, we quote their result for the change in frequency induced by the BCZ at an acoustic depth of $\tau_{\rm cz}$ (cf. Houdek and Gough, 2007, Eq. 17),

$$\left.\frac{\delta\omega}{\omega}\right|_{\rm cz,osc} \simeq \frac{c_{\rm cz}^2\Delta_{\rm cz}}{8\tau_0\omega^3}\left(1+1/4\tau_0^2\omega^2\right)^{-1/2}\cos\left[2(\tau_{\rm cz}\omega+\epsilon_{\rm cz})+\tan^{-1}(2\tau_0\omega)\right] \tag{4.45}$$

where $\epsilon_{\rm cz}$ is some phase which depends slowly on ω, $c_{\rm cz}$ is the speed of sound and,

$$\Delta_{\rm cz} = \left[\frac{{\rm d}^2\ln\rho}{{\rm d}r^2}\right]_{-r_{\rm cz}}^{+r_{\rm cz}}, \tag{4.46}$$

is the change in second density derivative at the BCZ ($r = r_{\rm cz}$).

For high-order modes, $\tau_0\omega \gg 1$. Hence, $\tan^{-1}(2\tau_0\omega) \simeq \pi/2$, and $(1+1/4\tau_0^2\omega^2)^{-1/2} \simeq 1$. Therefore, we can simplify this and write it in a more familiar form in terms of cyclic frequency,

$$\left.\frac{\delta\nu}{\nu}\right|_{\rm cz,osc} \simeq \frac{c_{\rm cz}^2\Delta_{\rm cz}}{32\pi^3}\frac{\nu_0}{\nu^3}\sin(4\pi\tau_{\rm cz}\nu+\phi_{\rm cz}) \tag{4.47}$$

where $\phi_{\rm cz}$ is some approximately constant phase term.

Now we have equations for the effect of He II ionisation and the BCZ for a given mode frequency ν, we can use them to model the glitch signatures as they appear in observed mode frequencies, ν_{nl}. Naturally, these equations are not exact. Several assumptions were made during their derivations. Therefore, if we want to model the glitch signature in ν_{nl}, we need some way of accounting for the uncertainty in the above equations. In the next chapter, we apply this model to characterise the glitch for some example stars. Although such models have been applied to other Sun-like stars before (e.g. Mazumdar et al., 2014; Verma et al., 2014), we propose a new statistical model to account for the uncertainty in the functional form of ν_{nl} with n.

Chapter 5

MODELLING ACOUSTIC GLITCH SIGNATURES IN STELLAR OSCILLATIONS WITH A GAUSSIAN PROCESS

In this chapter, I apply a new method for modelling acoustic glitch signatures in the radial mode frequencies of solar-like oscillators. I compare my method to another using a model star with different levels of noise. Then, I apply both methods to the star 16 Cyg A. I show that my method can be used to find the strength and location of glitches caused by the second ionisation of helium and the base of the convective zone. I also demonstrate that my method improves treatment of the smoothly varying component of the mode frequencies.

5.1 Introduction

Let us consider a non-rotating star which oscillates at frequencies ν_{nl}, where n and l are the radial order and angular degree of the modes. We may model the modes as the sum of a smoothly-varying component, $\tilde{f}(n,l)$, and quickly-varying, small change in frequency, $\delta\nu_{nl}$, arising from glitches in the stellar structure,

$$f(n,l) = \tilde{f}(n,l) + \delta\nu_{nl}, \tag{5.1}$$

where $\delta\nu_{nl}$ is some function of frequency which may be evaluated at $\tilde{f}(n,l)$.

In principle, $\delta\nu$ could arise from any glitches expected in a particular star. Here, we consider glitches in main sequence solar-like oscillators. In Chapter 4, we derived approximations for $\delta\nu_{nl}$ arising from acoustic glitches caused by the second ionisation of helium and the BCZ. We choose to ignore contributions to $\delta\nu_{nl}$ from the first ionisation of helium. Consequently, we let $\delta\nu_{nl} = \delta\nu_{\rm He} + \delta\nu_{\rm cz}$ where each component depends on parameters relating to properties of the glitches (cf. Equations 4.44 and 4.47),

$$\delta\nu_{\rm He} = \alpha_{\rm He}\nu_0\nu\, \mathrm{e}^{-\beta_{\rm He}\nu^2} \sin(4\pi\tau_{\rm He}\nu + \phi_{\rm He}), \tag{5.2}$$

$$\delta\nu_{\rm cz} = \alpha_{\rm cz}\nu_0\nu^{-2} \sin(4\pi\tau_{\rm cz}\nu + \phi_{\rm cz}). \tag{5.3}$$

The parameters $\alpha_{\rm He} \simeq \Gamma_{\rm II}$ and $\beta_{\rm He} \propto \Delta_{\rm II}^2$ relate to the area and variance of the Gaussian-like depression in γ caused by the second ionisation of helium (see Equation 4.38). The amplitude parameter for the BCZ glitch is proportional to the difference in the second density derivative at the base of the convective zone ($\alpha_{\rm cz} \propto \Delta_{\rm cz}$) with units of frequency squared. The approximate acoustic depths of second helium ionisation and the BCZ are given by $\tau_{\rm He}$ and $\tau_{\rm cz}$ respectively, and $\phi_{\rm He}$ and $\phi_{\rm cz}$ are arbitrary phase constants.

Providing $\tilde{f}(n,l)$ is a good approximation of the mode frequencies, we can calculate $\delta\nu_{nl}$ at $\tilde{f}(n,l)$ to predict $f(n,l)$. For example, $\tilde{f}(n,l)$ could be a K-th order polynomial in n with coefficients a_{lk} (e.g. Ulrich, 1986; Kjeldsen et al., 2005),

$$\tilde{f}(n,l) = \nu_0 \sum_{k=0}^{K} a_{lk} n^k. \tag{5.4}$$

The linear component of this is equivalent to the asymptotic expression (see Equation 1.2) and the remaining terms describe curvature in the mode frequencies. However, there are drawbacks to using a polynomial for $\tilde{f}(n,l)$. Whilst a polynomial with $K = \infty$ can represent any function, this is impractical here. If K is too low, then it will not be flexible enough, biasing the inference of $\delta\nu_{\rm He}$ and $\delta\nu_{\rm cz}$. Yet, if K is too high, then the model will over-fit and the glitch component may be missed. Regularising the polynomial is one solution to over-fitting, but this adds extra

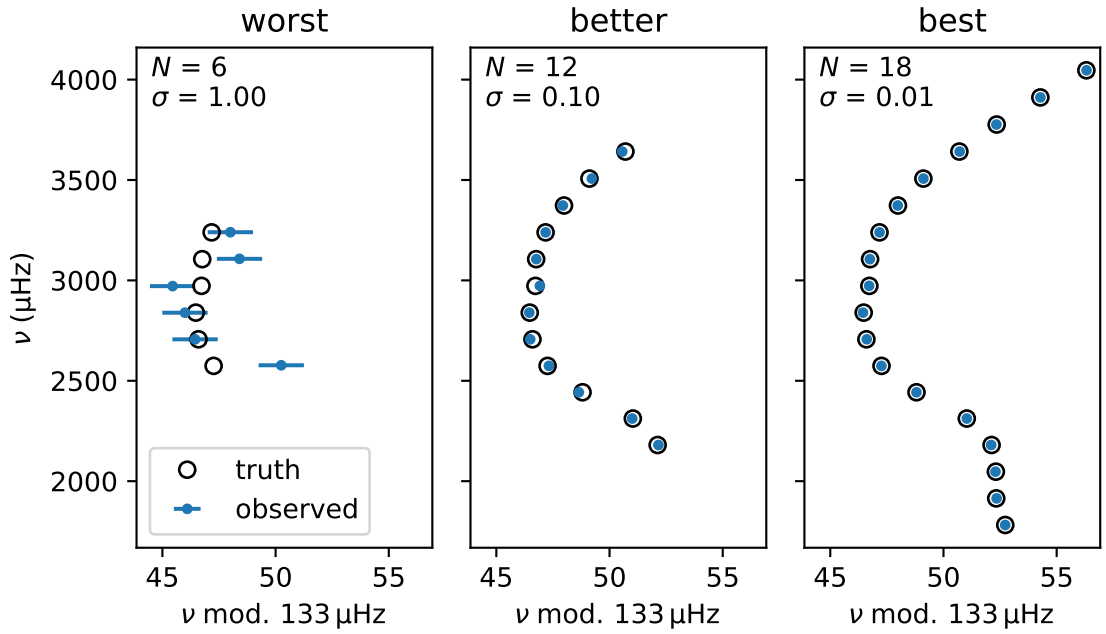


Figure 5.1: Echelle plot of simulated true and 'observed' radial mode frequency data from model star S for worst, better, and best case scenarios.

parameters to tune. Additionally, a finite polynomial represents only a small fraction of function space, systematically biasing our model to a particular functional form of $\tilde{f}$.

Directly fitting the glitch the above way has been done before (e.g. Mazumdar et al., 2014; Verma et al., 2014, 2017). In this work, we will compare the most recent version of this method (Verma et al., 2019, hereafter V19) with a new method. Our method will make use of a Gaussian Process (GP) to marginalise over the uncertainty in the functional form of f. In the next section, we introduce the data used in comparing the two methods. Then, we define both modelling in Section 5.3. Finally, we apply the methods to the data and present the results and discussion in Sections 5.4 and 5.5 respectively.

5.2 Data

In this section, we briefly outline the data used to compare our new method with the V19 method. For simplicity, we only use observations of radial mode frequencies in this work. However, both methods can be extended to use higher-degree modes.

Table 5.1: Observations of radial mode frequency ν_n at radial order n for model star S (*left*) and 16 Cyg A (*right*). N are the number of observed radial orders and the scale of the Gaussian noise added to each column is given by $\sigma_{\rm obs}$ where appropriate. The values and their uncertainties for 16 Cyg A come from Lund et al. (2017).

	ν_n (μHz)				
Star n	Worst	Better	Best	Truth	16 Cyg A
12	—	—	1781.726	1781.729	1390.808 ± 0.969
13	—	—	1914.348	1914.330	1495.053 ± 0.243
14	—	—	2047.296	2047.314	1598.690 ± 0.075
15	—	2180.16	2180.107	2180.125	1700.952 ± 0.101
16	—	2312.00	2312.030	2312.029	1802.351 ± 0.084
17	—	2442.64	2442.807	2442.808	1904.521 ± 0.058
18	2577.2	2574.32	2574.260	2574.263	2007.538 ± 0.047
19	2706.4	2706.50	2706.612	2706.596	2110.950 ± 0.044
20	2839.0	2839.46	2839.446	2839.476	2214.225 ± 0.055
21	2971.5	2972.92	2972.718	2972.727	2317.282 ± 0.059
22	3107.4	3105.77	3105.749	3105.759	2420.937 ± 0.082
23	3240.0	3239.18	3239.177	3239.175	2524.950 ± 0.164
24	—	3372.94	3372.980	3372.990	2628.930 ± 0.261
25	—	3507.22	3507.115	3507.112	2733.571 ± 0.445
26	—	3641.56	3641.715	3641.702	2840.148 ± 1.058
27	—	—	3776.339	3776.350	2944.937 ± 0.896
28	—	—	3911.281	3911.279	—
29	—	—	4046.319	4046.312	—
N	6	12	18	18	16
$\sigma_{\rm obs}$	1.00	0.10	0.01	—	—

Test Star We created three sets of test data for worst, better, and best case scenarios using stellar model R (defined in Section 4.3.1). Model R is similar to the Sun, with surface parameters of $T_{\rm eff} = 5682\,{\rm K}$, $\log g = 4.426$ and $[{\rm Fe/H}] = 0.03$, and bulk parameters of $M = 1.00\,{\rm M}_{\odot}$, $R = 1.01\,{\rm M}_{\odot}$ and $t_{\star} = 4.07\,{\rm Gyr}$. We calculated radial order modes ($l = 0$) for the test star using the GYRE oscillation code (Townsend and Teitler, 2013). Then, we selected different numbers of modes (N) symmetrically about a reference frequency, $\nu_{\rm ref} = 2900\,\mu{\rm Hz}$ (close to the expected frequency at maximum power of the star). To the frequencies, we added Gaussian noise scaled by $\sigma_{\rm obs} = 1.0$, 0.1, and 0.01 μHz, for the worst, better, and best cases respectively. The parameters and mode frequencies for each test case are shown in Table 5.1. We also plotted the modes on an echelle diagram in Figure 5.1. We can see the effect of the helium glitch in the echelle diagrams, where there is a periodic feature which is strongest at low frequencies. The glitch signature is visible in the better case and clearest in the best case scenario.

16 Cyg A We used the asteroseismic benchmark star 16 Cyg A as an example real star to test both methods. We adopted values for 16 radial mode frequencies identified by Lund et al. (2017) using observations from *Kepler* (Borucki et al., 2010, KIC 12069424). The mode frequencies and their associated uncertainties are given in Table 5.1. The glitch has been previously studied for 16 Cyg A with its binary companion 16 Cyg B in Verma et al. (2014), making it a useful subject for comparison. Similarly to model R, this target is a solar analogue. However, it is slightly hotter, more evolved and more metal-rich, with $T_{\rm eff} = 5825 \pm 50\,{\rm K}$, $\log g = 4.33 \pm 0.07$ dex and $[{\rm Fe/H}] = 0.10 \pm 0.03$ dex (Ramírez, Meléndez, and Asplund, 2009). Its fundamental stellar parameters are $M \approx 1.1\,{\rm M}_{\odot}$, $R \approx 1.2\,{\rm R}_{\odot}$ and $t_{\star} \approx 7\,{\rm Gyr}$ (Silva Aguirre et al., 2017).

5.3 Methods

In this section, we describe the two models for radial mode frequency, $\nu_n = f(n, 0)$, and the fitting methods being compared in this work. The first is the V19 method, originally used in Verma et al. (2014) to study the 16 Cyg binary star system. Secondly, we introduce our new 'GP method'. Both methods use different statistical philosophy and formulation of the smoothly

varying component, $\tilde{f}(n)$. However, they each fit the same glitch component $\delta\nu = \delta\nu_{\rm He} + \delta\nu_{\rm cz}$ as given in Equations 5.2 and 5.3.

5.3.1 The V19 Method

The V19 method is a specific case of Equation 5.1, where the smooth component is approximated by 4-th order polynomial,

$$f_A(n) = \tilde{f}_A(n) + \delta\nu_{\rm He} + \delta\nu_{\rm cz}; \quad \tilde{f}_A(n) = \sum_{k=0}^{4} b_k n^k, \tag{5.5}$$

where $b_k \equiv a_{0k}\nu_0$ from Equation 5.4. The model parameters are given by $\boldsymbol{\theta}_A = (b_0, \ldots, b_4, a_{\rm He}, \beta_{\rm He}, \tau_{\rm He}, \phi_{\rm He}, a_{\rm cz}, \tau_{\rm cz}, \phi_{\rm cz})$, where the glitch amplitude parameters are modified to include ν_0 such that $a_i \equiv \alpha_i\nu_0$. The ν_0 parameter is not explicitly included in the V19 model, but we find it useful to keep the scaling in mind, and we use ν_0 in the GP model.

The model parameters are optimised by minimising a χ^2 cost function with a regularisation term,

$$\chi^2 = \sum_n \left[\frac{\nu_n^{\rm obs} - f_A(n)}{\sigma_n^{\rm obs}}\right]^2 + \lambda^2 \sum_n \left[\frac{{\rm d}^3}{{\rm d}n^3}\tilde{f}_A(n)\right]^2, \tag{5.6}$$

where $\nu_n^{\rm obs}$ and $\sigma_n^{\rm obs}$ are the observed mode and its uncertainty at radial order n, and λ is the regularisation parameter. The regularisation was used to avoid the polynomial over-fitting and absorbing the glitch terms.

We fitted the model using the GlitchPy code[1]. The fitting method is described in Verma et al. (2019), in which we adopted the same value of $\lambda = 7$ and bounds for the selection of initial parameters. We chose the initial parameters randomly within their bounds and optimised them using a BFGS minimisation of χ^2 (Fletcher, 1987), repeated 200 times until a global minimum was found. We further repeated this for 1000 realisations of the observed ν_n with Gaussian noise scaled by $\sigma_n^{\rm obs}$ to obtain a range of possible solutions.

[1] https://github.com/alexlyttle/GlitchPy, adapted from https://github.com/kuldeepv89/GlitchPy.

5.3.2 The GP Method

In the second method, we used a Gaussian Process (GP; see Aigrain and Foreman-Mackey, 2022, for a recent review) instead of a high-order polynomial to model the smoothly varying component of the mode frequencies. The GP represents a probability distribution in function-space, meaning that we can quantify the uncertainty associated with the functional form of f. We write our model for a set of modes $\boldsymbol{n} = \{n_i\}_{i=1}^{N}$ as a random draw from a GP,

$$f_B(\boldsymbol{n}) \sim \mathcal{GP}\left[m(\boldsymbol{n}), k(\boldsymbol{n}, \boldsymbol{n}')\right], \tag{5.7}$$

where m and k are the mean and kernel functions, and $\boldsymbol{n}'$ is any other set of radial orders, $\boldsymbol{n}' = \{n'_j\}_{j=1}^{M}$. The mean function describes where we expect ν_n to be given some approximate physical reasoning. Therefore, we define our mean function for any n-th order mode as,

$$m(n) = \tilde{f}_B(n) + \delta\nu_{\rm He} + \delta\nu_{\rm cz}; \quad \tilde{f}_B(n) = (n + \varepsilon)\nu_0, \tag{5.8}$$

where $\tilde{f}_B(n)$ is the asymptotic approximation of the mode frequency (see Equation 1.2).

The kernel represents the expected covariance between the values of the function at different n. We chose the squared exponential kernel function to be compatible with a smoothly-varying function of n. Evaluating the kernel gives an $N \times M$ matrix with an element (i, j) given by,

$$k(n_i, n'_j) = \alpha_k \nu_0 \, \mathrm{e}^{-(n_i - n'_j)^2 / 2\lambda_k^2}, \tag{5.9}$$

where α_k is a dimensionless amplitude scale factor and λ_k is the length scale (in units of radial order). Both kernel parameters control the flexibility of the GP. The amplitude parameter scales the covariance and λ_k describes the breadth of correlation between different modes. As $\lambda_k \to 0$, the off-diagonal terms of the kernel approach zero producing uncorrelated noise with a variance of $\alpha_k \nu_0$. We found values of $\alpha_k = 0.5$ and $\lambda_k = 5$ predicted smoothly varying functions compatible with our prior expectation.

The GP likelihood for some set of observations $\boldsymbol{\nu}_{\rm obs} = \{\nu^{\rm obs}_{n_i}\}^N_{i=1}$ is a multivariate normal distribution centred on the mean function with covariance provided by the kernel function. We also added Gaussian noise terms to the diagonal of the covariance matrix,

$$\boldsymbol{\nu}_{\rm obs} \sim \mathcal{N}\left(\boldsymbol{\mu}, \boldsymbol{\Sigma}\right); \quad \boldsymbol{\Sigma} = \boldsymbol{K} + \mathrm{diag}(\sigma^2 + \boldsymbol{\sigma}^2_{\rm obs}), \tag{5.10}$$

where $\boldsymbol{\mu} = m(\boldsymbol{n})$ and $\boldsymbol{K} = k(\boldsymbol{n}, \boldsymbol{n})$. The scales of Gaussian noise in the model and observations are σ and $\boldsymbol{\sigma}_{\rm obs}$ respectively. We included σ to account for uncorrelated noise in the model, for example from the difference between $\delta\nu$ evaluated at $\tilde{f}_B(n)$ and at the 'true' mode frequencies.

To make noiseless predictions of mode frequencies $\boldsymbol{\nu}_\star$ at new radial orders $\boldsymbol{n}_\star$, we drew from the following multivariate normal distribution (Rasmussen and C. K. I. Williams, 2006),

$$\boldsymbol{\nu}_\star \mid \boldsymbol{\nu}_{\rm obs} \sim \mathcal{N}\left[\boldsymbol{\mu}_\star + \boldsymbol{K}_\star^{\mathsf{T}}\,\boldsymbol{\Sigma}^{-1}\,(\boldsymbol{\nu}_{\rm obs} - \boldsymbol{\mu}),\; \boldsymbol{K}_{\star\star} - \boldsymbol{K}_\star^{\mathsf{T}}\,\boldsymbol{\Sigma}^{-1}\,\boldsymbol{K}_\star\right] \tag{5.11}$$

where,

$$\boldsymbol{\mu}_\star = m(\boldsymbol{n}_\star), \quad \boldsymbol{K}_\star = k(\boldsymbol{n}, \boldsymbol{n}_\star), \quad \boldsymbol{K}_{\star\star} = k(\boldsymbol{n}_\star, \boldsymbol{n}_\star).$$

We used Bayes' theorem (see Equation 1.4) to estimate the probability of our model parameters, $\boldsymbol{\theta}_B = (\nu_0, \varepsilon, \alpha_{\rm He}, \beta_{\rm He}, \tau_{\rm He}, \phi_{\rm He}, a_{\rm cz}, \tau_{\rm cz}, \phi_{\rm cz}, \sigma)$, given observations of mode frequencies. Hence, we write the posterior probability density as,

$$p(\boldsymbol{\theta}_B \mid \boldsymbol{\nu}_{\rm obs}) = \frac{p(\boldsymbol{\nu}_{\rm obs} \mid \boldsymbol{\theta}_B)\, p(\boldsymbol{\theta}_B)}{p(\boldsymbol{\nu}_{\rm obs})} \equiv \frac{\mathcal{L}(\boldsymbol{\theta}_B)\, p(\boldsymbol{\theta}_B)}{\mathcal{Z}}, \tag{5.12}$$

where $\mathcal{L}(\boldsymbol{\theta}_B)$ is the likelihood of the data given the model, $p(\boldsymbol{\theta}_B)$ is the prior probability density of the model parameters, and $\mathcal{Z}$ is the 'evidence' (or normalisation). The true posterior density function cannot be derived analytically, so we estimated it with the 'nested sampling' algorithm (Skilling, 2004). By evaluating the prior volume at contours of constant likelihood, nested sampling estimates $\mathcal{Z}$ to weight samples according to their posterior probability. This method requires functions for the log-likelihood and a transform which maps random samples in the unit hypercube to prior parameter space.

Table 5.2: The GP model parameters and their prior distributions. Normal distributions are parametrised by the mean and variance (i.e. $\mathcal{N}(\mu, \sigma^2)$). Uniform distributions are parametrised by their lower and upper bounds.

Parameter	Prior	
	Test Star	16 Cyg A
$\nu_0/\mu\text{Hz}$	$\mathcal{N}(132.8, 0.01)$	$\mathcal{N}(103.28, 0.0025)$
ε	$\mathcal{N}(1.4, 0.0025)$	$\mathcal{N}(1.45, 0.0025)$
$\ln(\alpha_{\text{cz}}/\mu\text{Hz}^2)$	$\mathcal{N}(8.29, 0.64)$	$\mathcal{N}(8.04, 0.64)$
$\ln(\alpha_{\text{He}}/\text{Ms})$	$\mathcal{N}(-11.1, 0.04)$	$\mathcal{N}(-10.85, 0.04)$
$\ln(\beta_{\text{He}}/\text{Ms}^2)$	$\mathcal{N}(-15.07, 0.16)$	$\mathcal{N}(-14.56, 0.16)$
$\ln(\tau_{\text{cz}}/\text{Ms})$	$\mathcal{N}(-6.09, 0.04)$	$\mathcal{N}(-5.84, 0.04)$
$\ln(\tau_{\text{He}}/\text{Ms})$	$\mathcal{N}(-7.19, 0.04)$	$\mathcal{N}(-6.94, 0.04)$
$\ln(\sigma/\mu\text{Hz})$	$\mathcal{N}(-4.605, 4)$	$\mathcal{N}(-4.605, 4)$
$\phi_{\text{He}}, \phi_{\text{cz}}$	$\mathcal{U}(0, 2\pi)$	$\mathcal{U}(0, 2\pi)$

Following Equation 5.10, we defined the log-likelihood of the model as,

$$\ln \mathcal{L}_B = -\frac{1}{2}\left[(\boldsymbol{\nu}_{\text{obs}} - \boldsymbol{\mu})^{\mathsf{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\nu}_{\text{obs}} - \boldsymbol{\mu}) + \ln(|\boldsymbol{\Sigma}|) + N\ln(2\pi)\right], \tag{5.13}$$

where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ depend on $\boldsymbol{\theta}_B$. The prior transform for $\boldsymbol{\theta}_B$ is the inverse cumulative distribution function associated with the prior distribution $p(\boldsymbol{\theta}_B)$. We treated the prior for each of $\boldsymbol{\theta}_B$ independently. The total prior distribution is the product of distributions for each parameter, $p(\boldsymbol{\theta}_B) = \prod_j p(\theta_j)$. We give the prior distributions and their shape parameters in Table 5.2. In the following paragraphs, we justify our choice of prior distributions based on approximate scaling relations between the parameters.

Smooth Component Starting with priors for the parameters in $\tilde{f}_B(n)$, we chose normal distributions,

$$\nu_0 \sim \mathcal{N}(\overline{\nu}_0, s^2_{\nu_0}), \quad \varepsilon \sim \mathcal{N}(\overline{\varepsilon}, s^2_{\varepsilon}), \tag{5.14}$$

centred on $\overline{\nu}_0$ and $\overline{\varepsilon}$ and scaled by s_{ν_0} and s_{ε}. For example, the location and scale parameters could come from global estimates of $\langle\Delta\nu_n\rangle$ or a linear fit to the modes. We determined $\overline{\nu}_0$ and $\overline{\varepsilon}$ for the test stars from a linear fit to the true mode frequencies, and added representative

uncertainties of 10 and 5 per cent respectively. For 16 Cyg A, we used measurements of $\langle \Delta\nu_n \rangle$ and $\nu_{\max}$ from Lund et al. (2017) to estimate $\overline{\nu}_0$ and $\overline{\varepsilon}$.

Acoustic Depths We derived a prior on $\tau_{\rm He}$ and $\tau_{\rm cz}$ by observing how they scale with acoustic radius (τ_0) in the grid of stellar models from Lyttle et al. (2021) and the results of Verma et al. (2022). Typically, the fractional depth of He II ionisation and the BCZ are $\tau_{\rm II}/\tau_0 \approx 0.2$ and $\tau_{\rm cz}/\tau_0 \approx 0.6$, for main sequence solar-like oscillators (see, e.g. Figure 4.8). We formed the prior distributions from these relations, estimating the acoustic radius from $\tau_0 = (2\nu_0)^{-1}$. To account for possible variance with stellar properties, we added a spread of 20 per cent. We defined their priors in natural log-space to ensure they remained positive,

$$\ln \tau_{\rm He} \sim \mathcal{N}\left[\ln(\overline{\tau}_{\rm He}),\, s^2_{\ln\tau}\right], \quad \ln \tau_{\rm cz} \sim \mathcal{N}\left[\ln(3\,\overline{\tau}_{\rm He}),\, s^2_{\ln\tau}\right], \tag{5.15}$$

$$\overline{\tau}_{\rm He} = (10\,\overline{\nu}_0)^{-1}, \quad s^2_{\ln\tau} = (1/5)^2 + (s_{\nu_0}/\overline{\nu}_0)^2. \tag{5.16}$$

Helium Glitch Amplitude Deriving a prior on the glitch amplitude parameters was less trivial. By observing γ profiles in the grid of stellar models, we assumed that the width of the helium ionisation region was about 8 per cent of the acoustic depth of the region, $\Delta_{\rm II}/\tau_{\rm II} \approx 0.08$. Thus, we centred the prior on $\overline{\beta}_{\rm He}$ obtained from this assumption and used the relation $\beta_{\rm He} = 8\pi^2\Delta^2_{\rm II}$ from comparison to Equation 4.44. We then propagated the variance from the prior on $\tau_{\rm He}$,

$$\ln \beta_{\rm He} \sim \mathcal{N}\left[\ln(\overline{\beta}_{\rm He}),\, 4s^2_{\ln\tau}\right], \quad \overline{\beta}_{\rm He} = \frac{32}{625}\,\pi^2\overline{\tau}^2_{\rm He}. \tag{5.17}$$

Then, we centred the prior for $\alpha_{\rm He}$ to satisfy a depth of 0.1 in $\delta\gamma/\gamma$ caused by helium ionisation using Equation 4.38,

$$\ln \alpha_{\rm He} \sim \mathcal{N}\left[1/2\,\ln(\overline{\beta}_{\rm He}/400\pi),\, s^2_{\ln\tau}\right]. \tag{5.18}$$

We verified that the priors on $\alpha_{\rm He}$ and $\beta_{\rm He}$ were compatible with our prior belief by checking that the amplitude at $\nu_{\rm ref}$ peaked from 0 to 1 μHz, decaying thereafter.

BCZ Glitch Amplitude We constructed the prior on α_{cz} such that the BCZ glitch amplitude was approximately 0.1 μHz at ν_{ref} (an order of magnitude smaller than the helium glitch). This occurred when $\alpha_{\mathrm{cz}}/\nu_0 \approx 30\,\mu$Hz. We also scaled the distribution by an additional 80 per cent,

$$\ln \alpha_{\mathrm{cz}} \sim \mathcal{N}\left[\ln(30\,\overline{\nu}_0),\ (4/5)^2 + (s_{\nu_0}/\overline{\nu}_0)^2\right].$$

Prior distributions for the remaining parameters were the same for both stars. For example, the prior on the phase parameters ϕ_{He} and ϕ_{cz} was uniformly distributed from 0 to 2π. We also used a weakly informative prior on the model uncertainty, $\ln\sigma \sim \mathcal{N}(-\ln 100, 4)$, centred on 0.01 μHz.

We sampled the posterior distribution using the Python nested sampling package `dynesty` (Speagle, 2020; Koposov et al., 2023). We applied the multi-ellipsoid bounding method (Feroz, Hobson, and Bridges, 2009) with 500 live points and the random walk sampling method (Skilling, 2006) with a minimum of 50 steps before proposing a new live point. In addition, we enabled periodic boundary conditions for ϕ_{He} and ϕ_{cz} with the prior transform projecting to periodic space using $\phi = \phi' \bmod 2\pi$, where ϕ' is unconstrained. Otherwise, the nested sampler ran with its default parameters. We used the Python package `jax` (Bradbury et al., 2018) to make use of accelerated linear algebra (XLA) and just-in-time (JIT) compilation, and `tinygp` (Foreman-Mackey et al., 2022) to build the GP. For analysis and comparison with the V19 method, we drew 1000 points randomly from the posterior samples according to their estimated weights.

5.4 Results

5.4.1 Test Star

In Figure 5.2, we plot the predicted total glitch component ($\delta\nu$) using 50 draws from the posterior distribution for both methods. We subtracted the median smooth component ($\tilde{f}(n)$) from the observed ν_n over-plot. For each scenario, both methods predicted similar glitch components. However, the V19 method showed more extreme multimodality in all cases. For example, the

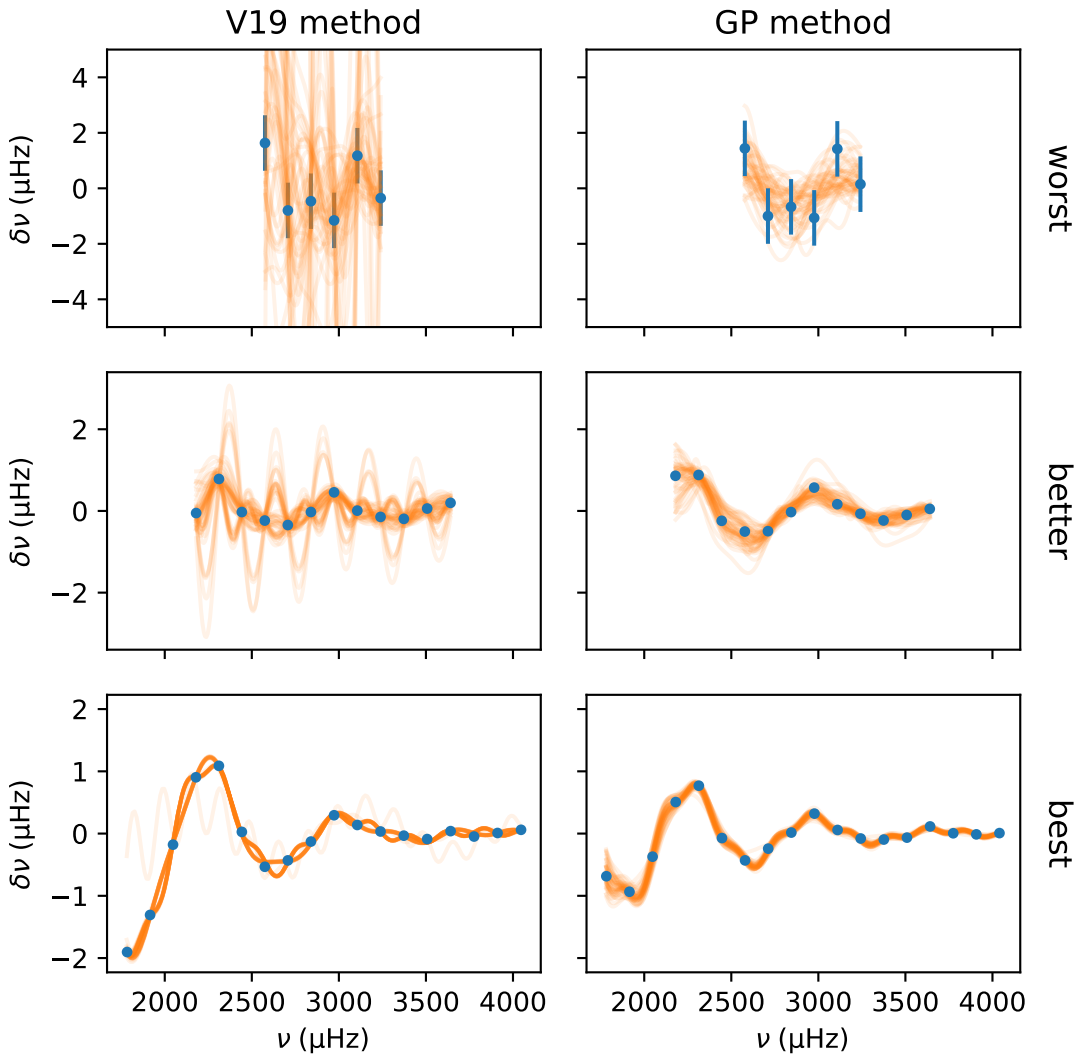


Figure 5.2: 50 random draws from the V19 and GP methods showing the total glitch signal, $\delta\nu = \delta_{\mathrm{He}}(\nu) + \delta_{\mathrm{cz}}(\nu)$, as a function of frequency, ν. The data points plot in blue are the median smooth component model at $\boldsymbol{n}$, subtracted from the observed modes ($\boldsymbol{\nu}_{\mathrm{obs}}$) with their observed uncertainty (σ_{obs}).

better case showed a high-amplitude ($\sim 2\,\mu$Hz) BCZ glitch solution which is not present with the GP method. Such a large BCZ signature would be unlikely. Additionally, the two methods differed by $\sim 1\,\mu$Hz at low frequency in the best case. Despite this, the V19 method appeared to be more confident in its glitch solutions compared to the GP method.

In Figure 5.3, we plotted posterior distributions for the glitch acoustic depths, τ_{He} and τ_{cz}, and compared them to the sound speed gradient of the test star from Figure 4.5. We expect the acoustic depths to approximately line up with the sharp structural changes. For the worst case,

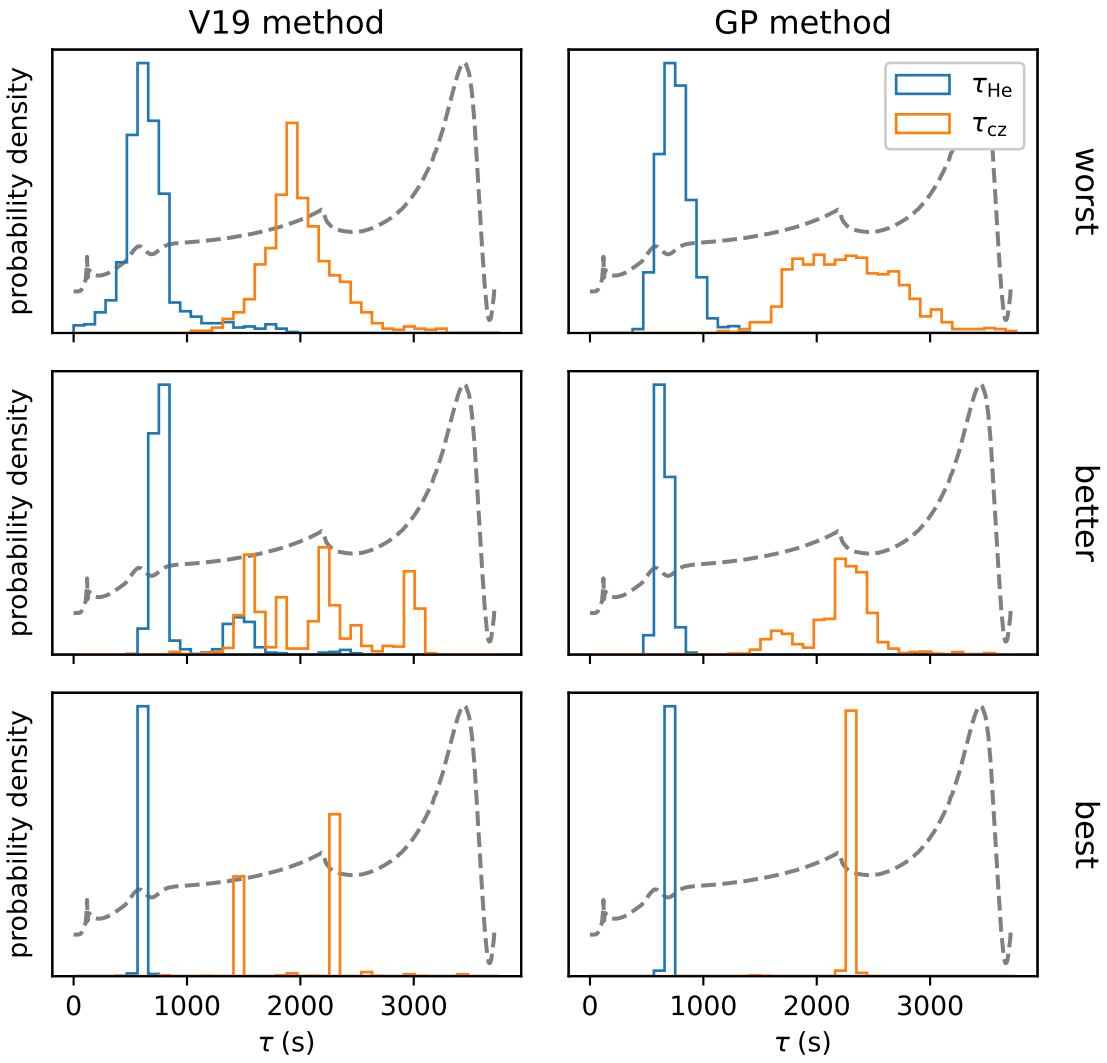


Figure 5.3: Samples from the V19 and GP methods for the acoustic depths of the helium glitch (τ_{He}) and base of the convective zone (τ_{cz}). The sound-speed gradient from Figure 4.5 is over-plot with a grey dashed line.

both methods gave broad distributions for the acoustic depths, compatible with their respective initial guesses and priors. The V19 method initial guesses appeared to underestimate τ_{cz}, whereas the GP method prior was broad enough to encompass a wide range of possible τ_{cz}. In the better and best cases, we found that the V19 solutions were multimodal. For example, the better case found solutions for τ_{He} far deeper into the star than we would expect, at around 1500 s and 2500 s.

The values for τ_{He} obtained were under-predicted compared to the location of the trough due to the second ionisation of helium. We can see this for the best star fit with the V19 method

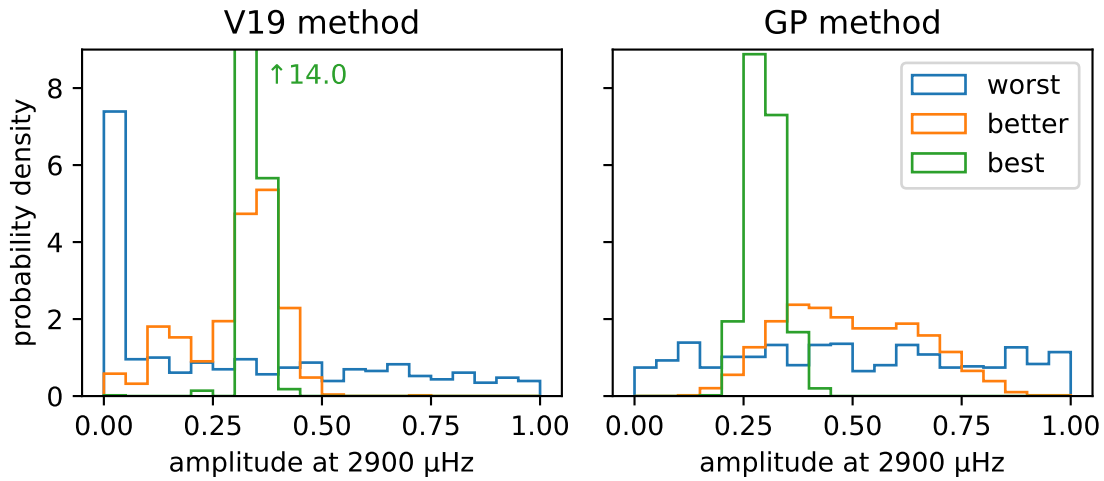


Figure 5.4: Samples of the helium glitch amplitude at a reference frequency of 3000 μHz fit with the V19 and GP methods for the worst, better, and best case test data. The tallest bar in the V19 method panel is cropped with its value displayed in green text.

which finds $\tau_{\mathrm{He}} = 619 \pm 15$ s. The depression in γ due to He II ionisation in the respective stellar model is located at 733 s. On the other hand, the GP method was closer with $\tau_{\mathrm{He}} = 696 \pm 19$ s.

We compared helium glitch amplitudes at ν_{ref} from both methods in Figure 5.4. The V19 method preferred low-amplitude solutions for the worst and better cases compared to the GP method. On the other hand, the GP method reflected our prior in the worst case, with the width of the distribution shrinking as the data improved. The GP method did show some bi-modality in the better case, with higher solutions at $A_{\mathrm{He}}^{\mathrm{ref}} \approx 0.7$ which were not found by the V19 method. In the best case, the V19 method obtained $A_{\mathrm{He}}^{\mathrm{ref}} = 0.347^{+0.006}_{-0.005}$, whereas the GP method found $A_{\mathrm{He}}^{\mathrm{ref}} = 0.296^{+0.042}_{-0.036}$.

5.4.2 16 Cyg A

We compared the results from both methods applied to the 16 Cyg A data in Figure 5.5. We found that the V19 method predicted extreme solutions for the glitch where the GP method did not. There was also a difference of about 1 μHz at the low frequency end such that the V19 predicted a larger glitch amplitude. Otherwise, the two methods gave similar solutions for the glitch function.

Both methods found similar values for $\tau_{\rm He}$ but relatively different distributions for $\tau_{\rm cz}$. Regarding the helium glitch, the V19 method found $\tau_{\rm He} = 917^{+50}_{-53}$ s, and our GP method obtained $\tau_{\rm He} = 931^{+59}_{-88}$ s, within 1-σ of each other. Similarly to the test star, the GP method found a slightly larger value of $\tau_{\rm He}$ than the V19 method. Verma et al. (2014) fit the glitch using $l = 0, 1, 2$ modes and found an acoustic depth of 930 ± 14 s, within 1-σ of both methods in this work.

We calculated the helium glitch amplitude at a reference frequency of 2188.5 μHz, equivalent to $\nu_{\rm max}$ obtained by Lund et al. (2017). Samples from both posteriors are shown in the bottom panel of Figure 5.5. We found $A^{\rm ref}_{\rm He} = 0.260^{+0.050}_{-0.065}$ μHz for the V19 method, and $A^{\rm ref}_{\rm He} = 0.333^{+0.081}_{-0.073}$ μHz for the GP method. Both values were about 1-σ apart. Additionally, both methods found $A^{\rm ref}_{\rm cz} \sim 0.1$ μHz, with the GP method favouring a smaller value. The V19 method found some solutions where $A^{\rm ref}_{\rm cz}$ was larger than $A^{\rm ref}_{\rm He}$, something we would not expect because the modes are less sensitive to structural changes deeper in the star.

5.5 Discussion

Throughout this work, we found a smaller $\delta\nu$ amplitude at low frequency with the GP method than with the V19 method. This was particularly visible in the best case and in 16 Cyg A. We expected this was a result of the different smooth background models. In Figure 5.6 we plotted the smooth component of each model extended to lower order, unobserved modes. We found the GP background component had a turning point at $\nu \approx 1900$ μHz, higher than the V19 method at $\nu \approx 1500$ μHz. The smooth component of the V19 method was confidently incorrect outside the observed frequencies. Conversely, the GP method predicted closer to the truth with increasing uncertainty further from the observations. It appeared that the GP provided a more accurate representation of the underlying function than the polynomial.

In the test star's best case and 16 Cyg A, the GP method found a higher $\tau_{\rm He}$ than the V19 method. This could be because the GP was better able to distinguish between the He II glitch and the smaller He I glitch. To explore this, we plotted the difference between each model's

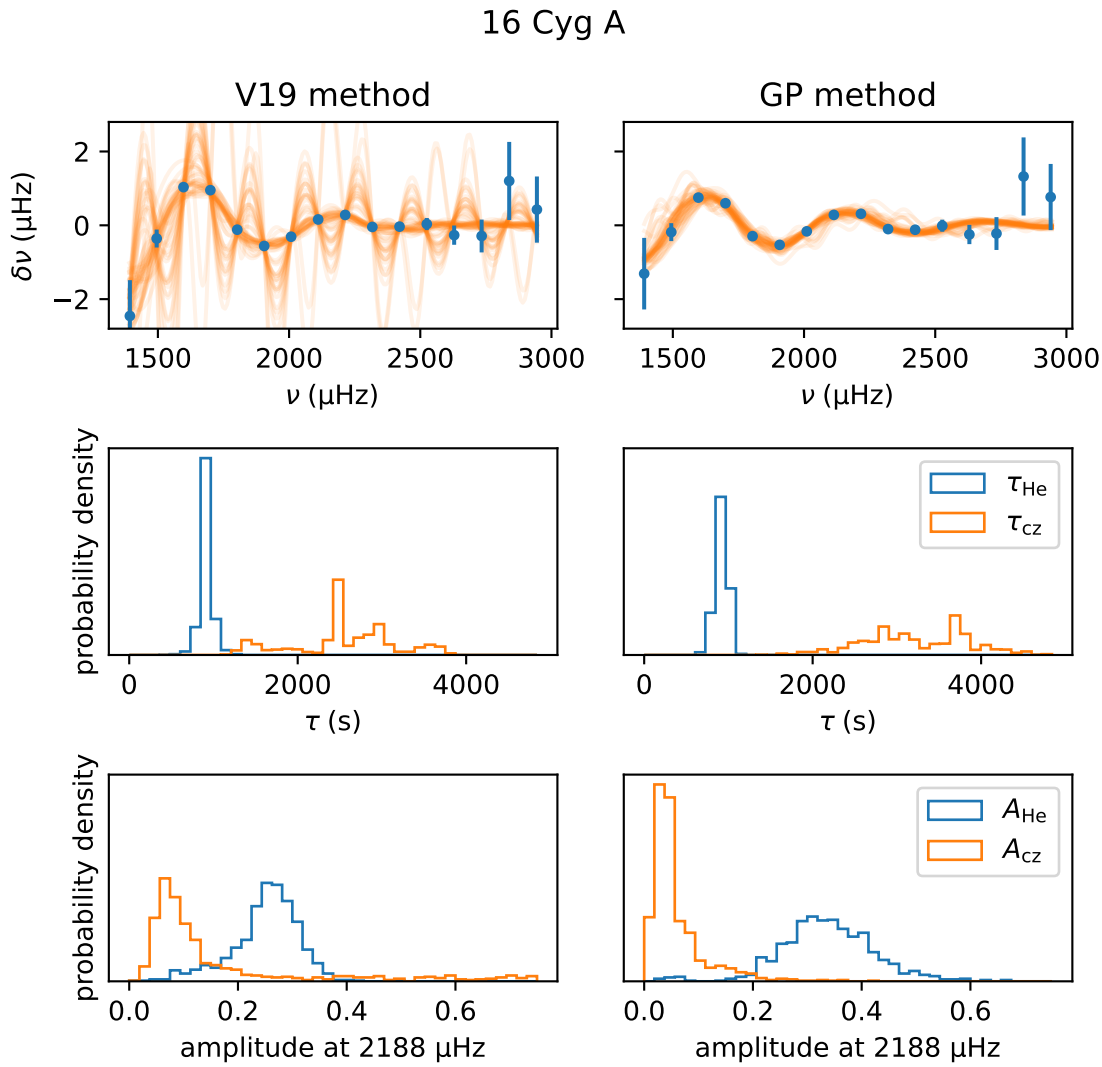


Figure 5.5: Results for the glitch fit to 16 Cyg A using the V19 method (*left*) and GP method (*right*). The top panel shows 50 draws from the posterior prediction of $\delta\nu = \delta\nu_{\mathrm{He}} + \delta\nu_{\mathrm{cz}}$ with the observed ν_n minus the median smooth background component of the model. The middle and bottom panels shows the posterior sample probability density for the acoustic depths and reference amplitudes of the glitches respectively.

predictions for the smooth component in Figure 5.7. We saw a clear periodic signal in the differences. The period of this signal in the test star corresponded to an acoustic depth of $\sim$ 500 s, matching the location of He II ionisation in model R. The periodic signal was less pronounced for 16 Cyg A, but corresponded to a plausible He II acoustic depth of $\sim$ 700 s. We expected that the polynomial was not flexible enough to pick up the He II ionisation signature, hence lowering its mean value for τ_{He}.

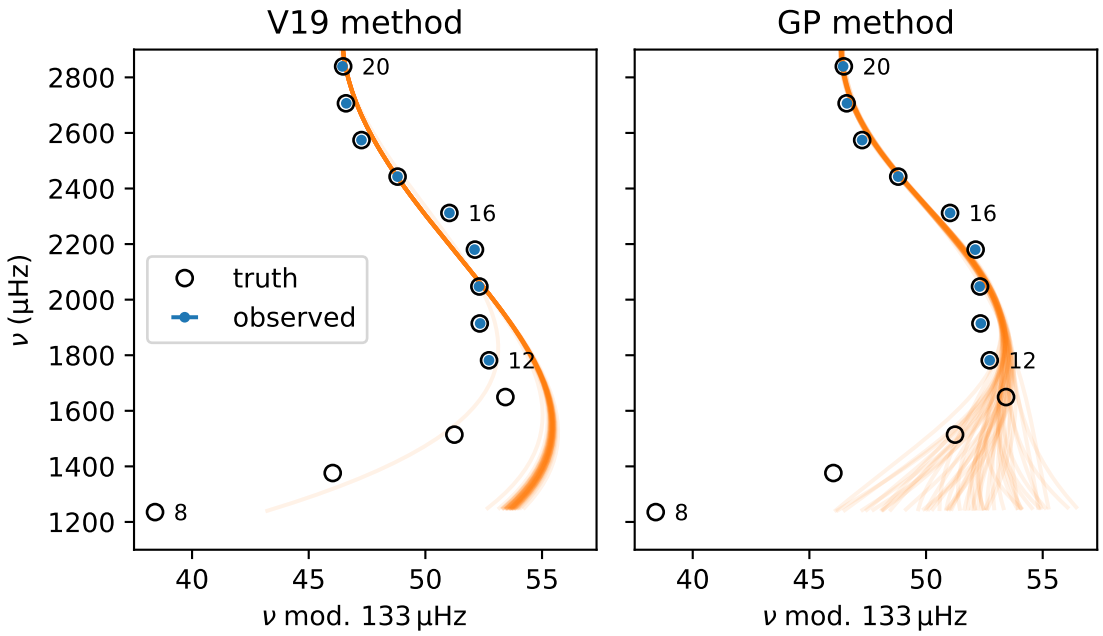


Figure 5.6: Echelle-like diagrams showing mode frequencies plot against their frequency modulo 133 μHz for the best case scenario. The orange lines show 50 draws from the model posterior predictions with the glitch component removed ($f(n) - \delta\nu$) for both methods. The predictions are made from a lower radial order $n = 8$ than the observed modes. The true values from model R are shown for each mode. Some radial orders (n) are shown next to the modes for context.

The V19 method found several extreme solutions for $\delta\nu$ whereas GP method did not. We expected this because the GP method used a prior over the model parameters. We tested relaxing the prior on $\boldsymbol{\theta}_B$ and found similar multimodal posteriors to the V19 method. This showed that the prior helped eliminate unrealistic solutions. However, care should be taken over the choice of prior on $\boldsymbol{\theta}_B$ to realistically reflect our expectation. If some of our prior assumptions are incorrect, they may bias the results. For example, the outer convective region gets shallower as stars get hotter (approaching $T_{\rm eff} \approx 7000\,{\rm K}$) making the assumption that $\tau_{\rm cz}/\tau_0 \approx 0.6$ an overestimate in these cases. We accommodate for this with a wide prior on $\tau_{\rm cz}$, but our prior could be more informed. There is a notable temperature dependence to $\tau_{\rm cz}/\tau_0$ observed in the grid of stellar models which could be exploited in the future when constructing the prior.

Additionally, the joint posteriors for $\alpha_{\rm He}$ and $\beta_{\rm He}$ were correlated in both methods. This was expected, because larger values of $\beta_{\rm He}$ can be compensated for with a larger amplitude factor

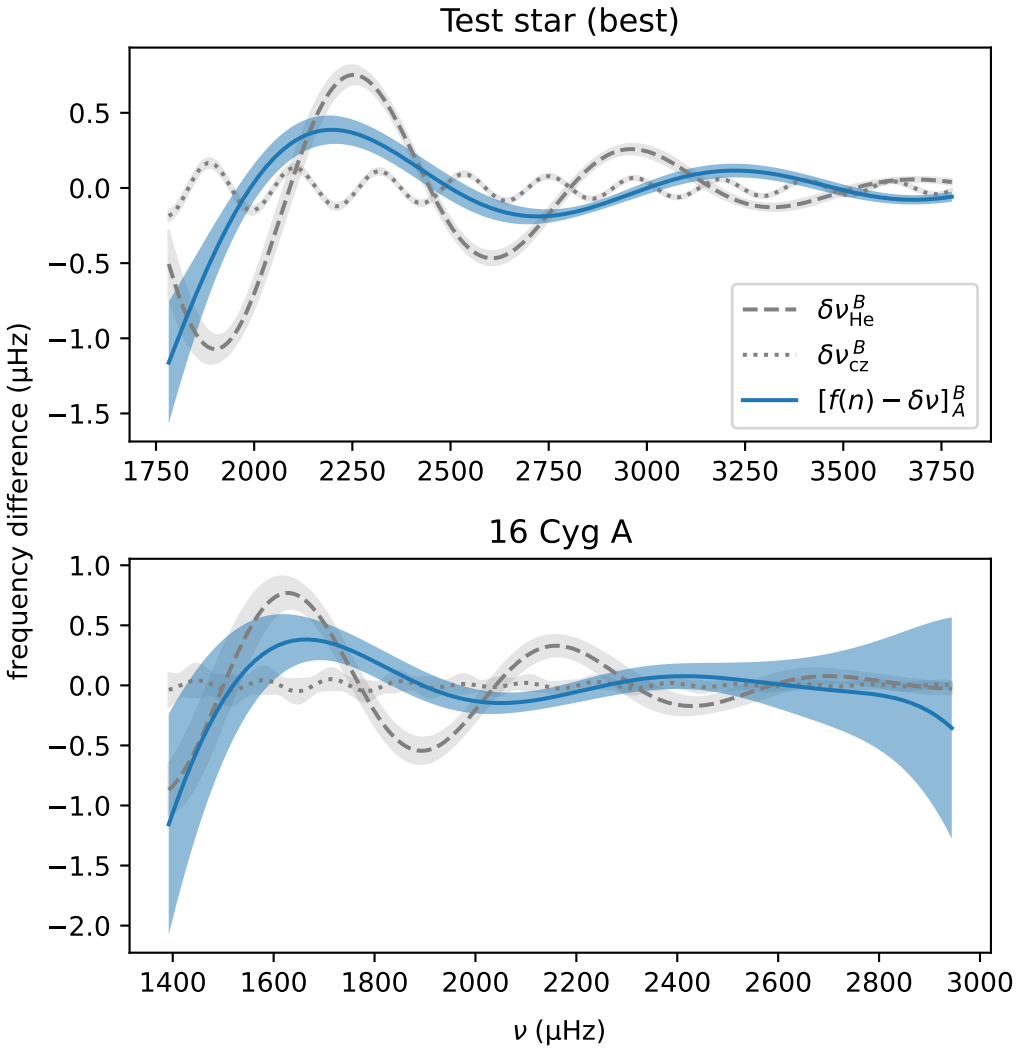


Figure 5.7: The difference between the model without the glitch components ($f(n) - \delta\nu$) predicted by the V19 method (A) and the GP method (B). The glitch components from the GP method are plot in grey for context. The 68 per cent confidence region is shaded for each line.

($\alpha_{\rm He}$). In the GP method, we did not include this expected correlation in our prior. Hence, we found that having broader priors on $\alpha_{\rm He}$ and $\beta_{\rm He}$ lead to the prior predicting unrealistic glitches. We could account for this by using a multivariate prior for the amplitude parameters. Despite this, our approach still improves on the V19 method by using a prior in the first place.

A potential limitation of the GP method is that we calculate the glitch at $\tilde{f}_B(n)$ from the linear asymptotic equation (Equation 1.2). In regions where the gradient of $\delta\nu$ is high, the difference between $\tilde{f}_B(n)$ and the true frequency can be up to $\sim 0.1\,\mu$Hz. Our choice of kernel function cannot absorb this difference because is varies on a short length-scale. Instead, we accounted for

this uncertainty by adding Gaussian noise to the model parametrised by σ. However, for the best test star, we found $\sigma \approx 0.05$, which was larger than the observational uncertainty, $\sigma_{\rm obs} = 0.01$. This could limit our method's inference ability for the best asteroseismic targets. One solution is to replace $\tilde{f}_B(n)$ with a quadratic (e.g. Nielsen et al., 2021) which better approximates ν_n. However, the GP kernel would need to be adjusted to account for this.

5.6 Conclusion

We introduced a new method for modelling acoustic glitches in solar-like oscillators using a Gaussian Process. Testing the method on a model star, we found that it more accurately characterised the underlying, smoothly-varying functional form of the radial modes than the V19 method. Furthermore, our method appeared able to absorb the glitch component from He I ionisation, for which the polynomial was not flexible enough. However, this questions whether the He I ionisation glitch should be explicitly included in the model.

Additionally, the GP method provided more believable uncertainties on the glitch parameters, whereas the V19 method was over-confident with the best data and under-confident with the worst. Robust uncertainties are important when using the results to make further inference about helium enrichment. In this case, the GP marginalised over correlated noise in the model, not possible with the V19 method.

Future development of the method could involve building a prior for the glitch parameters. For example, we could start with fitting the model to simulated stars and using the results to build an empirical prior. Then, we could run the model on a larger asteroseismic sample of main sequence stars (e.g. Davies et al., 2016; Lund et al., 2017) and compare our results to those from Verma et al. (2019). Additionally, we could add parameters from the GP model to the hierarchical model introduced in Lyttle et al. (2021). Ultimately, our goal is to scale this method in anticipation of the $\sim 10^4$ solar-like oscillators expected to be observed by *PLATO* (Rauer et al., 2014).

Chapter 6

CONCLUSION & FUTURE PROSPECTS

I conclude this thesis by providing a summary of the work and reflecting upon key results. Then, I consider possible improvements to our hierarchical model (from Chapter 3) and method for emulating stellar simulations. Finally, I discuss future prospects for applying our method to data from current and upcoming missions.

Summary

In this thesis, we built a hierarchical Bayesian model (HBM) to improve the inference of stellar parameters with asteroseismology. Introducing the concept of an HBM in Chapter 2, we showed how population-level distributions can be used as a prior over individual stellar parameters. We found that pooling parameters this way reduced their uncertainties by up to a factor of $\sqrt{N}$ where N is the number of stars in the population.

In Chapter 3, we built an HBM to estimate the masses, radii, and ages for a well-studied sample of ~ 60 dwarf and subgiant solar-like oscillators. Limited by observational noise, existing modelling techniques struggle to account for the uncertainty in initial helium abundance (Y) and mixing-length theory parameter ($\alpha_{\rm MLT}$) for these stars. We showed that applying a hierarchical prior over Y and $\alpha_{\rm MLT}$ allowed us to simultaneously marginalise over their uncertainty and model their distribution in the population. Pooling Y and $\alpha_{\rm MLT}$ in this way, we were still able to achieve statistical uncertainties of 1.2 per cent in radius, 2.5 per cent in mass, and 12 per cent in age.

Notably, our HBM halved the uncertainty in stellar mass compared to the same model without parameter pooling. This provided a scalable and reproducible framework for modelling large populations of stars at the same time.

The HBM required a function to map model parameters to observables. We built an emulator to approximate 1D numerical models of stellar evolution. Training a neural network on MESA stellar simulations, we were able to predict observable parameters ($T_{\rm eff}, \Delta\nu, L, [{\rm M/H}]$) with typical precisions of less than $\sim$ 0.1 per cent (see Appendix A). This provided a simple, continuous, and differentiable model well suited to modern, gradient-based MCMC algorithms. Another advantage to using a neural network emulator was its scalability. The simple matrix algebra involved was well suited to fast evaluations on a graphics processing unit (GPU) for large numbers of stars at the same time. Furthermore, the neural network could be scaled up to higher input and output dimensions with little performance impact. This made our method transferable to other kinds of stars. For example, we recently trained a neural network to emulate the regularly spaced mode frequencies as a part of a Bayesian stellar model of δ Scuti-type oscillators (Scutt et al., 2023).

In Chapter 4, we recalled that glitches in stellar structure cause a periodic signal ($\delta\nu$) in p mode frequencies. One such glitch arises from the second ionisation of helium, with the amplitude of $\delta\nu_{\rm He}$ correlating with helium abundance. Subsequently, we presented a new method for modelling the glitch signature using a Gaussian Process (GP) in Chapter 5. Past methods for measuring the glitch in the mode frequencies (ν_{nl}) used a polynomial in n to approximate the smooth component, over which the periodic glitch signature could be modelled (e.g. Mazumdar et al., 2014; Verma et al., 2019). We replaced the polynomial with a GP characterised by a kernel describing our prior belief of the function's smoothness and flexibility. We applied this method to model the glitch signature in radial mode frequencies for a simulated Sun-like star and 16 Cyg A. The GP allowed us to marginalise over our uncertainty in the functional form of $\nu_{n\,l=0}$ with n. In comparison, we found that the polynomial method was too restrictive and unable to marginalise over correlated noise in the model.

Improving the Hierarchical Model

The helium glitch parameters for a given star correlate with its near-surface helium abundance. Therefore, a natural next step would be to include helium glitch parameters as an additional observable in our HBM. Our GP glitch model can be applied to both observed and modelled mode frequencies, providing extra parameters to include in our stellar model emulator. Adding these should improve inference of helium abundance for stars with individual modes identified (e.g. Davies et al., 2016; Lund et al., 2017). Since our HBM simultaneously models the population distribution of helium, even a few stars with good helium constraint will in-turn improve helium estimates for the rest of the population. This introduces the possibility of testing more complex models of helium enrichment.

We also expect the HBM to scale to red giant solar-like oscillators for which observations are abundant. We trained the emulator on a grid of stellar models from the zero-age main sequence to the base of the red giant branch for masses from 0.8 to 1.2 $\mathrm{M}_{\odot}$. The upper mass limit was motivated by the diminishing outer convective envelope (responsible for driving solar-like oscillators) in these stars on the main sequence. However, extending the emulator to model red giant solar-like oscillators would require expanding the grid up to $\sim 2.0\,\mathrm{M}_{\odot}$. We would need to compute twice as many more evolutionary tracks and evolve existing models further. Stars with $M \gtrsim 1.1\,\mathrm{M}_{\odot}$ have a convective core on the main sequence which introduces an additional model uncertainty from mixing at its boundary. Parametrising this process would further multiply the number of input tracks, increasing dimensionality and grid computation time. Therefore, we should research ways of selectively computing stellar models. For example, we could augment the grid (e.g. Li et al., 2022) by upsampling where the neural network error is large.

There are a few additional systematic uncertainties we could also include in the HBM. For example, in Chapter 3 we did not consider the inaccuracies of near-surface physics which effect modelled mode frequencies. So-called 'surface correction' methods exist (e.g. Kjeldsen, Bedding, and Christensen-Dalsgaard, 2008; Ball and Gizon, 2014) but vary across the HR diagram when compared with 3D hydrodynamical simulations (Sonoi et al., 2015). Compton

et al. (2018) found a range of surface corrections can shift modelled frequencies at $\nu_{\rm max}$ by up to $\sim$ 0.5 per cent. This would amount to a systematic effect on $\Delta\nu$ which we would expect to correlate with other stellar parameters. Therefore, a future iteration of the HBM should account for the surface term systematic.

Current and Future Data

We tested the HBM on stars observed by *Kepler*, but there are a few current and upcoming missions which we can utilise to increase our sample size. With larger sample sizes, we can further increase the precision of pooled parameters and better characterise their spread in the population distribution. Recently, Hatt et al. (2023) identified a sample of $\sim$ 4000 solar-like oscillators in 120- and 20-second cadence *TESS* data. Of these, around 50 are dwarf and subgiant stars which we could include in a future iteration of the HBM. However, we anticipate greater improvements with future missions expected to launch in a few years time.

Towards the end of the 2020s, the *PLATO* mission will observe tens of thousands of dwarf and subgiant solar-like oscillators (Rauer et al., 2014). *PLATO* aims to discover hundreds of exoplanets orbiting solar-type stars across a wider proportion of the sky than observed by *Kepler*. Among its targets are around 20 000 bright ($V < 11$) oscillating F-K dwarf stars to be observed over a baseline of around 2 years (Goupil, 2017). Using our HBM method on a sample this size could see a reduction in uncertainty (σ) on helium abundance from 0.01 to 0.0005. While this is the maximum expected uncertainty reduction (as discussed in Chapter 2), it shows that we can start to consider more complex population distributions in helium and other stellar parameters.

While *PLATO* will offer unprecedented numbers of main sequence solar-like oscillators, we already have large samples of more evolved asteroseismic stars to include in a future HBM. Combined, *Kepler*, *K2*, and *TESS* have yielded $\sim$ 150, 000 red giant solar-like oscillators to date (Yu et al., 2018; Hon et al., 2021). Providing that we can extend our stellar model emulator to more these stars, expanding our dataset will allow us to test more complex population-distributions. For example, since *TESS* is an all-sky survey, we could include kinematics and

galactic positions in the helium enrichment law. Additionally, observations of open clusters and binary star systems motivate population distributions over age, distance and chemical abundances.

Extending our HBM to red giants may be particularly worthwhile with the upcoming *Nancy Grace Roman Space Telescope* (*Roman*, formerly *WFIRST*; Spergel et al., 2015). Although asteroseismology is not its main focus, *Roman* will observe the galactic bulge for a few years to search for free-floating planets with micro-lensing. Its light curves will be noisier than *Kepler* and oscillation amplitudes are lower in infrared light, but Gould et al. (2015) still predict a yield of $\sim 10^6$ red giant oscillators. With these huge asteroseismic datasets expected in the next 10 years, it has never been a better time to research scalable methods for modelling populations of stars.

DATA AVAILABILITY & SOFTWARE

The data and code underlying this thesis are available in the online supplementary material of Lyttle et al. (2021), and in the Zenodo database at https://doi.org/10.5281/zenodo.4746352. The code used to produce the remainder of this work is available at https://github.com/alexlyttle/thesis, and in the Zenodo database at https://doi.org/10.5281/zenodo.8027956. This thesis includes data collected by the *Kepler* mission. Funding for the *Kepler* mission is provided by the NASA Science Mission directorate This work has also used data from the European Space Agency (ESA) mission *Gaia* (https://www.cosmos.esa.int/gaia), processed by the *Gaia* Data Processing and Analysis Consortium (DPAC, https://www.cosmos.esa.int/web/gaia/dpac/consortium). Funding for the DPAC has been provided by national institutions, in particular the institutions participating in the *Gaia* Multilateral Agreement. To query other published datasets, this research used the VizieR catalogue access tool, CDS, Strasbourg, France. The original description of the VizieR service was published in Ochsenbein, Bauer, and Marcout (2000). Finally, this work used the *Gaia-Kepler* crossmatch database at https://gaia-kepler.fun created by Megan Bedell.

I acknowledge use of the Python programming language (Python Software Foundation, https://www.python.org) for the majority of code written for this work. Specific Python packages used are referenced in-text, except for: `matplotlib` (v3.6.2; Hunter, 2007; Caswell et al., 2022) and `seaborn` (Waskom, 2021) for creating plots; `scipy` (Virtanen et al., 2020) for general scientific computational methods; `astropy` (Astropy Collaboration et al., 2022) for reading and writing astronomical data; `astroquery` (Ginsburg et al., 2019) for querying astronomy databases; and `lightkurve` (Lightkurve Collaboration et al., 2018) for processing *Kepler* light curves. Finally, this thesis was compiled directly to PDF from LaTeX source code based on the

`uob-thesis-template` template (https://github.com/alexlyttle/uob-thesis-template).

REFERENCES

Abadi, M. et al. (May 2016). "TensorFlow: A System for Large-Scale Machine Learning". In: *arXiv e-prints*. arXiv: 1605.08695 [cs.DC].

Aerts, C. (Jan. 2021). "Probing the Interior Physics of Stars through Asteroseismology". In: *Rev. Mod. Phys.* 93, p. 015001. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.93.015001.

Aguirre Børsen-Koch, V. et al. (Jan. 2022). "The BAyesian STellar Algorithm (BASTA): A Fitting Tool for Stellar Studies, Asteroseismology, Exoplanets, and Galactic Archaeology". In: *MNRAS* 509, pp. 4344–4364. ISSN: 0035-8711. DOI: 10.1093/mnras/stab2911.

Aigrain, S. and D. Foreman-Mackey (Sept. 2022). "Gaussian Process Regression for Astronomical Time-Series". In: *arXiv e-prints*. DOI: 10.48550/arXiv.2209.08940. arXiv: 2209.08940 [astro-ph].

Albareti, F. D. et al. (Dec. 2017). "The 13th Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-IV Survey Mapping Nearby Galaxies at Apache Point Observatory". In: *ApJS* 233, p. 25. DOI: 10.3847/1538-4365/aa8992.

Anderson, L. et al. (Oct. 2018). "Improving Gaia Parallax Precision with a Data-driven Model of Stars". In: *AJ* 156, p. 145. ISSN: 0004-6256. DOI: 10.3847/1538-3881/aad7bf.

Angus, R. et al. (June 2015). "Calibrating Gyrochronology Using Kepler Asteroseismic Targets". In: *MNRAS* 450, pp. 1787–1798. ISSN: 0035-8711. DOI: 10.1093/mnras/stv423.

Appourchaux, T. et al. (July 2012). "Oscillation Mode Frequencies of 61 Main-Sequence and Subgiant Stars Observed by Kepler". In: *A&A* 543, A54. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201218948.

Appourchaux, T. et al. (Oct. 2015). "A Seismic and Gravitationally Bound Double Star Observed by Kepler. Implication for the Presence of a Convective Core". In: *A&A* 582, A25. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201526610.

Asplund, M. et al. (Sept. 2009). "The Chemical Composition of the Sun". In: *ARA&A* 47, pp. 481–522. ISSN: 0066-4146. DOI: 10.1146/annurev.astro.46.060407.145222.

Astropy Collaboration et al. (Aug. 2022). "The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package". In: *ApJ* 935, p. 167. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac7c74.

Aver, E., K. A. Olive, and E. D. Skillman (July 2015). "The Effects of He I λ10830 on Helium Abundance Determinations". In: *J. Cosmology Astropart. Phys.* 07, p. 011. ISSN: 1475-7516. DOI: 10.1088/1475-7516/2015/07/011.

Baglin, A. et al. (Jan. 2006). "CoRoT: A High Precision Photometer for Stellar Ecolution and Exoplanet Finding". In: *36th COSPAR Sci. Assem.* Vol. 36. Beijing, China, p. 3749.

Bahcall, J. N., M. H. Pinsonneault, and G. J. Wasserburg (Oct. 1995). "Solar Models with Helium and Heavy-Element Diffusion". In: *Rev. Mod. Phys.* 67, pp. 781–808. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.67.781.

Bahcall, J. N. and R. K. Ulrich (Apr. 1988). "Solar Models, Neutrino Experiments, and Helioseismology". In: *Rev. Mod. Phys.* 60, pp. 297–372. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.60.297.

Ball, W. H. and L. Gizon (Aug. 2014). "A New Correction of Stellar Oscillation Frequencies for Near-Surface Effects". In: *A&A* 568, A123. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201424325.

Balser, D. S. (Dec. 2006). "The Chemical Evolution of Helium". In: *AJ* 132, pp. 2326–2332. ISSN: 0004-6256. DOI: 10.1086/508515.

Basu, S. (July 1997). "Seismology of the Base of the Solar Convection Zone". In: *MNRAS* 288, pp. 572–584. ISSN: 0035-8711. DOI: 10.1093/mnras/288.3.572.

Basu, S. and H. M. Antia (May 2004). "Constraining Solar Abundances Using Helioseismology". In: *ApJ* 606, pp. L85–L88. DOI: 10.1086/421110.

Basu, S., H. M. Antia, and D. Narasimha (Mar. 1994). "Helioseismic Measurement of the Extent of Overshoot below the Solar Convection Zone". In: *MNRAS* 267, p. 209. ISSN: 0035-8711. DOI: 10.1093/mnras/267.1.209.

Basu, S. et al. (May 2004). "Asteroseismic Determination of Helium Abundance in Stellar Envelopes". In: *MNRAS* 350, pp. 277–286. DOI: 10.1111/j.1365-2966.2004.07644.x.

Bazot, M., S. Bourguignon, and J. Christensen-Dalsgaard (Jan. 2008). "Estimation of Stellar Parameters Using Monte Carlo Markov Chains". In: *Mem. Della Soc. Astron. Ital.* 79, p. 660. ISSN: 0037-8720. DOI: 10.48550/arXiv.0803.2529.

Bedding, T. R. et al. (Mar. 2011). "Gravity Modes as a Way to Distinguish between Hydrogen- and Helium-Burning Red Giant Stars". In: *Nature* 471, pp. 608–611. ISSN: 0028-0836. DOI: 10.1038/nature09935.

Bellinger, E. P. et al. (Oct. 2016). "Fundamental Parameters of Main-Sequence Stars in an Instant with Machine Learning". In: *ApJ* 830, p. 31. DOI: 10.3847/0004-637X/830/1/31.

Berger, T. A. et al. (June 2020). "The Gaia-Kepler Stellar Properties Catalog. I. Homogeneous Fundamental Properties for 186,301 Kepler Stars". In: *AJ* 159, p. 280. DOI: 10.3847/1538-3881/159/6/280.

Bingham, E. et al. (2019). "Pyro: Deep Universal Probabilistic Programming". In: *J. Mach. Learn. Res.* 20, 28:1–28:6.

Blanton, M. R. et al. (July 2017). "Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe". In: *AJ* 154, p. 28. DOI: 10.3847/1538-3881/aa7567.

Bonaca, A. et al. (Aug. 2012). "Calibrating Convective Properties of Solar-like Stars in the Kepler Field of View". In: *ApJ* 755, p. L12. ISSN: 0004-637X. DOI: 10.1088/2041-8205/755/1/L12.

Borucki, W. J. et al. (Feb. 2010). "Kepler Planet-Detection Mission: Introduction and First Results". In: *Science* 327, p. 977. DOI: 10.1126/science.1185402.

Bouchy, F. and F. Carrier (July 2001). "P-Mode Observations on Cen A". In: *A&A* 374.1, pp. L5–L8. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361:20010792.

Bradbury, J. et al. (2018). *JAX: Composable Transformations of Python+NumPy Programs*.

Brogaard, K. et al. (Jan. 2011). "Age and Helium Content of the Open Cluster NGC 6791 from Multiple Eclipsing Binary Members. I. Measurements, Methods, and First Results". In: *A&A* 525, A2. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201015503.

Brogaard, K. et al. (July 2012). "Age and Helium Content of the Open Cluster NGC 6791 from Multiple Eclipsing Binary Members. II. Age Dependencies and New Insights". In: *A&A* 543, A106. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201219196.

Broomhall, A.-M. et al. (June 2011). "Solar-Cycle Variations of Large Frequency Separations of Acoustic Modes: Implications for Asteroseismology". In: *MNRAS* 413, pp. 2978–2986. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2011.18375.x.

Brown, T. M. et al. (Feb. 1991). "Detection of Possible P-Mode Oscillations on Procyon". In: *ApJ* 368, p. 599. ISSN: 0004-637X. DOI: 10.1086/169725.

Burke, C. J. et al. (Feb. 2014). "Planetary Candidates Observed by Kepler IV: Planet Sample from Q1-Q8 (22 Months)". In: *ApJS* 210, p. 19. ISSN: 0067-0049. DOI: 10.1088/0067-0049/210/2/19.

Campante, T. L. et al. (Mar. 2016). "Spin-Orbit Alignment of Exoplanet Systems: Ensemble Analysis Using Asteroseismology". In: *ApJ* 819, p. 85. ISSN: 0004-637X. DOI: 10.3847/0004-637X/819/1/85.

Casagrande, L. et al. (Mar. 2010). "An Absolutely Calibrated Teff Scale from the Infrared Flux Method. Dwarfs and Subgiants". In: *A&A* 512, A54. ISSN: 0004-6361. DOI: 10.1051/0004-6361/200913204.

Casagrande, L. et al. (Dec. 2007). "The Helium Abundance and $\Delta Y/\Delta Z$ in Lower Main-Sequence Stars". In: *MNRAS* 382, pp. 1516–1540. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2007.12512.x.

Caswell, T. A. et al. (Nov. 2022). *matplotlib/matplotlib: REL: v3.6.2*. Zenodo. DOI: 10.5281/zenodo.7275322.

Chabrier, G. (July 2003). "Galactic Stellar and Substellar Initial Mass Function". In: *Publ. Astron. Soc. Pac.* 115, pp. 763–795. ISSN: 0004-6280. DOI: 10.1086/376392.

Chan, V. C. and J. Bovy (Apr. 2020). "The Gaia DR2 Parallax Zero-Point: Hierarchical Modelling of Red Clump Stars". In: *MNRAS* 493, pp. 4367–4381. ISSN: 0035-8711. DOI: 10.1093/mnras/staa571.

Chandrasekhar, S. (1939). *An Introduction to the Study of Stellar Structure*. The University of Chicago Press.

— (Feb. 1964). "A General Variational Principle Governing the Radial and the Non-Radial Oscillations of Gaseous Masses." In: *ApJ* 139, p. 664. ISSN: 0004-637X. DOI: 10.1086/147792.

Chaplin, W. J. et al. (Apr. 2007). "Solar P-Mode Frequencies over Three Solar Cycles". In: *ApJ* 659, pp. 1749–1760. ISSN: 0004-637X. DOI: 10.1086/512543.

Chaplin, W. J. et al. (Apr. 2011). "Ensemble Asteroseismology of Solar-Type Stars with the NASA Kepler Mission". In: *Science* 332, p. 213. ISSN: 0036-8075. DOI: 10.1126/science.1201827.

Chaplin, W. J. et al. (Jan. 2014). "Asteroseismic Fundamental Properties of Solar-type Stars Observed by the NASA Kepler Mission". In: *ApJS* 210, p. 1. ISSN: 0067-0049. DOI: 10.1088/0067-0049/210/1/1.

Chaplin, W. J. and A. Miglio (Aug. 2013). "Asteroseismology of Solar-Type and Red-Giant Stars". In: *ARA&A* 51, pp. 353–392. DOI: 10.1146/annurev-astro-082812-140938.

Chaplin, W. J. et al. (Jan. 2020). "Age Dating of an Early Milky Way Merger via Asteroseismology of the Naked-Eye Star ν Indi". In: *Nat. Astron.* 4, pp. 382–389. ISSN: 2397-3366. DOI: 10.1038/s41550-019-0975-9.

Chiosi, C. and F. M. Matteucci (Jan. 1982). "The Helium to Heavy Element Enrichment Ratio, Delta Y/Delta Z". In: *A&A* 105, pp. 140–148. ISSN: 0004-6361.

Choi, J. et al. (June 2016). "Mesa Isochrones and Stellar Tracks (MIST). I. Solar-scaled Models". In: *ApJ* 823, p. 102. DOI: 10.3847/0004-637X/823/2/102.

Christensen-Dalsgaard, J. (Jan. 1984). "What Will Asteroseismology Teach Us". In: *Space Res. Stellar Act. Var.* P. 11.

— (Jan. 2014). *Lecture Notes on Stellar Oscillations*. Fifth. University of Aarhus, Denmark: Unpublished manuscript.

Christensen-Dalsgaard, J., C. R. Proffitt, and M. J. Thompson (Feb. 1993). "Effects of Diffusion on Solar Models and Their Oscillation Frequencies". In: *ApJ* 403, pp. L75–L78. DOI: 10.1086/186725.

Claverie, A. et al. (Dec. 1979). "Solar Structure from Global Studies of the 5-Minute Oscillation". In: *Nature* 282.5739, pp. 591–594. ISSN: 1476-4687. DOI: 10.1038/282591a0.

Clevert, D.-A., T. Unterthiner, and S. Hochreiter (Nov. 2015). "Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)". In: *arXiv e-prints*. arXiv: 1511.07289 [cs.LG].

Compton, D. L. et al. (Oct. 2018). "Surface Correction of Main-Sequence Solar-like Oscillators with the Kepler LEGACY Sample". In: *MNRAS* 479, pp. 4416–4431. ISSN: 0035-8711. DOI: 10.1093/mnras/sty1632.

Connelly, J. N. et al. (Nov. 2012). "The Absolute Chronology and Thermal Processing of Solids in the Solar Protoplanetary Disk". In: *Science* 338, p. 651. DOI: 10.1126/science.1226919.

Cooke, R. J. and M. Fumagalli (Oct. 2018). "Measurement of the Primordial Helium Abundance from the Intergalactic Medium". In: *Nat. Astron.* 2, pp. 957–961. ISSN: 2397-3366. DOI: 10.1038/s41550-018-0584-z.

Cyburt, R. H. et al. (Jan. 2016). "Big Bang Nucleosynthesis: Present Status". In: *Rev. Mod. Phys.* 88, p. 015004. ISSN: 0034-6861. DOI: 10.1103/RevModPhys.88.015004.

Davies, G. R. et al. (Jan. 2015). "Asteroseismic Inference on Rotation, Gyrochronology and Planetary System Dynamics of 16 Cygni". In: *MNRAS* 446, pp. 2959–2966. ISSN: 0035-8711. DOI: 10.1093/mnras/stu2331.

Davies, G. R. et al. (Feb. 2016). "Oscillation Frequencies for 35 *Kepler* Solar-Type Planet-Hosting Stars Using Bayesian Techniques and Machine Learning". In: *MNRAS* 456.2, pp. 2183–2195. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stv2593.

De Ridder, J. et al. (May 2009). "Non-Radial Oscillation Modes with Long Lifetimes in Giant Stars". In: *Nature* 459.7245, pp. 398–400. ISSN: 1476-4687. DOI: 10.1038/nature08022.

Deheuvels, S. et al. (Sept. 2012). "Seismic Evidence for a Rapidly Rotating Core in a Lower-giant-branch Star Observed with Kepler". In: *ApJ* 756, p. 19. ISSN: 0004-637X. DOI: 10.1088/0004-637X/756/1/19.

Deubner, F.-L., R. K. Ulrich, and E. J. Rhodes Jr. (Feb. 1979). "Solar P-Mode Oscillations as a Tracer of Radial Differential Rotation". In: *A&A* 72, pp. 177–185. ISSN: 0004-6361.

Deubner, F.-L. and D. O. Gough (1984). "Helioseismology: Oscillations as a Diagnostic of the Solar Interior". In: *ARA&A* 22, pp. 593–619. DOI: 10.1146/annurev.aa.22.090184.003113.

Dillon, J. V. et al. (Nov. 2017). "TensorFlow Distributions". In: *arXiv e-prints*. arXiv: 1711.10604 [cs.LG].

Dotter, A. (Jan. 2016). "MESA Isochrones and Stellar Tracks (MIST) 0: Methods for the Construction of Stellar Isochrones". In: *ApJS* 222, p. 8. DOI: 10.3847/0067-0049/222/1/8.

Dotter, A. et al. (May 2017). "The Influence of Atomic Diffusion on Stellar Ages and Chemical Tagging". In: *ApJ* 840, p. 99. DOI: 10.3847/1538-4357/aa6d10.

Eggleton, P. P., J. Faulkner, and B. P. Flannery (Mar. 1973). "An Approximate Equation of State for Stellar Material". In: *A&A* 23, p. 325. ISSN: 0004-6361.

Ferguson, J. W. et al. (Apr. 2005). "Low-Temperature Opacities". In: *ApJ* 623, pp. 585–596. ISSN: 0004-637X. DOI: 10.1086/428642.

Feroz, F., M. P. Hobson, and M. Bridges (Oct. 2009). "MULTINEST: An Efficient and Robust Bayesian Inference Tool for Cosmology and Particle Physics". In: *MNRAS* 398, pp. 1601–1614. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.14548.x.

Feuillet, D. K. et al. (Jan. 2016). "Determining Ages of APOGEE Giants with Known Distances". In: *ApJ* 817, p. 40. ISSN: 0004-637X. DOI: 10.3847/0004-637X/817/1/40.

Fletcher, R. (Jan. 1987). *Practical Methods of Optimization*. Second. Chichester, NY: Wiley.

Foreman-Mackey, D. et al. (Mar. 2021). *daft-dev/daft: daft v0.1.2*. Zenodo. DOI: 10.5281/zenodo.4615289.

Foreman-Mackey, D. et al. (Oct. 2022). *dfm/tinygp: v0.2.3*. Zenodo. DOI: 10.5281/zenodo.7269074.

Foreman-Mackey, D. (June 2016). "corner.py: Scatterplot Matrices in Python". In: *J. Open Source Softw.* 1.2, p. 24. DOI: 10.21105/joss.00024.

Frebel, A. (May 2010). "Stellar Archaeology: Exploring the Universe with Metal-Poor Stars". In: *Astron. Nachrichten* 331, pp. 474–488. DOI: 10.1002/asna.201011362.

Gaia Collaboration et al. (Nov. 2016). "The Gaia Mission". In: *A&A* 595, A1. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201629272.

Gaia Collaboration et al. (Aug. 2018). "Gaia Data Release 2. Summary of the Contents and Survey Properties". In: *A&A* 616, A1. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201833051.

Gaia Collaboration et al. (July 2022). "Gaia Data Release 3: Summary of the Content and Survey Properties". In: *arXiv e-prints*. DOI: 10.48550/arXiv.2208.00211.

García, R. A. et al. (Dec. 2014). "Rotation and Magnetism of Kepler Pulsating Solar-like Stars. Towards Asteroseismically Calibrated Age-Rotation Relations". In: *A&A* 572, A34. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201423888.

García, R. A. and J. Ballot (Sept. 2019). "Asteroseismology of Solar-Type Stars". In: *Living Rev. Sol. Phys.* 16, p. 4. DOI: 10.1007/s41116-019-0020-1.

García, R. A. et al. (Aug. 2010). "CoRoT Reveals a Magnetic Activity Cycle in a Sun-Like Star". In: *Science* 329, p. 1032. ISSN: 0036-8075. DOI: 10.1126/science.1191064.

García Pérez, A. E. et al. (June 2016). "ASPCAP: The APOGEE Stellar Parameter and Chemical Abundances Pipeline". In: *AJ* 151, p. 144. DOI: 10.3847/0004-6256/151/6/144.

Gelly, B., G. Grec, and E. Fossat (Aug. 1986). "Evidence for Global Pressure Oscillations in Procyon and Alpha Centauri." In: *A&A* 164, pp. 383–394. ISSN: 0004-6361.

Gelman, A. and D. B. Rubin (1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Stat. Sci.* 7, pp. 457–472. DOI: 10.1214/ss/1177011136.

Ginsburg, A. et al. (Mar. 2019). "Astroquery: An Astronomical Web-querying Package in Python". In: *AJ* 157, p. 98. ISSN: 0004-6256. DOI: 10.3847/1538-3881/aafc33.

Glorot, X., A. Bordes, and Y. Bengio (Jan. 2011). "Deep Sparse Rectifier Neural Networks". In: *Proc. 14th Int. Conf. Artif. Intell. Statisitics AISTATS 2011*. Vol. 15, pp. 315–323.

Goldreich, P. and D. A. Keeley (Feb. 1977). "Solar Seismology. II. The Stochastic Excitation of the Solar p-Modes by Turbulent Convection." In: *ApJ* 212, pp. 243–251. ISSN: 0004-637X. DOI: 10.1086/155043.

Goodfellow, I., Y. Bengio, and A. Courville (2016). *Deep Learning*. Cambridge, MA: MIT Press.

Gough, D. O. (May 1977). "Mixing-Length Theory for Pulsating Stars". In: *ApJ* 214, pp. 196–213. DOI: 10.1086/155244.

— (Jan. 1986). "Solar and Solar-like Oscillations: Theory." In: *Highlights Astron.* 7, pp. 283–293.

— (Jan. 1993). "Linear Adiabatic Stellar Pulsation." In: *Astrophysical Fluid Dynamics. Les Houches Session LXVII*. Ed. by J.-P. Zahn and J. Zinn-Justin. Vol. 47. Lecture Notes of the Les Houches Summer School. Netherlands: Elsevier, pp. 399–560.

Gough, D. O. and M. J. Thompson (Jan. 1991). "The Inversion Problem". In: *Solar Interior and Atmosphere*. Ed. by A. N. Cox, W. C. Livingston, and M. S. Matthews. Space Science Series. Tucson, AZ: University of Arizona Press, pp. 519–561.

Gould, A. et al. (Apr. 2015). "WFIRST Ultra-Precise Astrometry II: Asteroseismology". In: *J. Korean Astron. Soc.* 48, pp. 93–104. ISSN: 1225-4614. DOI: 10.5303/JKAS.2015.48.2.93.

Goupil, M. (Oct. 2017). "Expected Asteroseismic Performances with the Space Project PLATO". In: *Eur. Phys. J. Web Conf.* Vol. 160. European Physical Journal Web of Conferences, p. 01003. DOI: 10.1051/epjconf/201716001003.

Green, G. M. et al. (July 2018). "Galactic Reddening in 3D from Stellar Photometry - an Improved Map". In: *MNRAS* 478, pp. 651–666. DOI: 10.1093/mnras/sty1008.

Grevesse, N. and A. J. Sauval (May 1998). "Standard Solar Composition". In: *Space Sci. Rev.* 85, pp. 161–174. ISSN: 0038-6308. DOI: 10.1023/A:1005161325181.

Gruberbauer, M., D. B. Guenther, and T. Kallinger (Apr. 2012). "Toward a New Kind of Asteroseismic Grid Fitting". In: *ApJ* 749, p. 109. ISSN: 0004-637X. DOI: 10.1088/0004-637X/749/2/109.

Gruberbauer, M. et al. (Oct. 2013). "Bayesian Asteroseismology of 23 Solar-like Kepler Targets". In: *MNRAS* 435, pp. 242–254. ISSN: 0035-8711. DOI: 10.1093/mnras/stt1289.

Hahnloser, R. H. R. et al. (June 2000). "Digital Selection and Analogue Amplification Coexist in a Cortex-Inspired Silicon Circuit". In: *Nature* 405, pp. 947–951. DOI: 10.1038/35016072.

Hall, O. J. et al. (July 2019). "Testing Asteroseismology with Gaia DR2: Hierarchical Models of the Red Clump". In: *MNRAS* 486, pp. 3569–3585. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1092.

Hall, O. J. et al. (Jan. 2021). "Weakened Magnetic Braking Supported by Asteroseismic Rotation Rates of Kepler Dwarfs". In: *Nat. Astron.* 5, pp. 707–714. ISSN: 2397-3366. DOI: 10.1038/s41550-021-01335-x.

Hatt, E. et al. (Jan. 2023). "Catalogue of Solar-like Oscillators Observed by TESS in 120-s and 20-s Cadence". In: *A&A* 669, A67. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/202244579.

Hawkins, K. et al. (Oct. 2017). "Red Clump Stars and Gaia: Calibration of the Standard Candle Using a Hierarchical Probabilistic Model". In: *MNRAS* 471, pp. 722–729. ISSN: 0035-8711. DOI: 10.1093/mnras/stx1655.

Haykin, S. (2007). *Neural Networks: A Comprehensive Foundation*. Third. Upper Saddle River, NJ: Prentice-Hall, Inc. ISBN: 0-13-147139-2.

Helmi, A. et al. (Nov. 2018). "The Merger That Led to the Formation of the Milky Way's Inner Stellar Halo and Thick Disk". In: *Nature* 563.7729, pp. 85–88. ISSN: 1476-4687. DOI: 10.1038/s41586-018-0625-x.

Hendriks, L. and C. Aerts (Oct. 2019). "Deep Learning Applied to the Asteroseismic Modeling of Stars with Coherent Oscillation Modes". In: *Publ. Astron. Soc. Pac.* 131, p. 108001. ISSN: 0004-6280. DOI: 10.1088/1538-3873/aaeeec.

Hoffman, M. D. and A. Gelman (Jan. 2014). "The No-U-turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo". In: *J. Mach. Learn. Res.* 15.1, pp. 1593–1623. ISSN: 1532-4435.

Hogg, D. W., A. D. Myers, and J. Bovy (Dec. 2010). "Inferring the Eccentricity Distribution". In: *ApJ* 725, pp. 2166–2175. ISSN: 0004-637X. DOI: 10.1088/0004-637X/725/2/2166.

Hon, M., D. Stello, and J. Yu (Aug. 2017). "Deep Learning Classification in Asteroseismology". In: *MNRAS* 469.4, pp. 4578–4583. ISSN: 0035-8711, 1365-2966. DOI: 10.1093/mnras/stx1174.

— (May 2018). "Deep Learning Classification in Asteroseismology Using an Improved Neural Network: Results on 15 000 Kepler Red Giants and Applications to K2 and TESS Data". In: *MNRAS* 476, pp. 3233–3244. ISSN: 0035-8711. DOI: 10.1093/mnras/sty483.

Hon, M. et al. (Oct. 2021). "A "Quick Look" at All-sky Galactic Archeology with TESS: 158,000 Oscillating Red Giants from the MIT Quick-look Pipeline". In: *ApJ* 919.2, p. 131. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac14b1.

Houdayer, P. S. et al. (Nov. 2021). "Properties of the Ionisation Glitch. I. Modelling the Ionisation Region". In: *A&A* 655, A85. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202141711.

Houdek, G. and D. O. Gough (Mar. 2007). "An Asteroseismic Signature of Helium Ionization". In: *MNRAS* 375, pp. 861–880. DOI: 10.1111/j.1365-2966.2006.11325.x.

Howell, S. B. et al. (Apr. 2014). "The K2 Mission: Characterization and Early Results". In: *Publ. Astron. Soc. Pac.* 126, p. 398. ISSN: 0004-6280. DOI: 10.1086/676406.

Hoyle, F. (Jan. 1946). "The Synthesis of the Elements from Hydrogen". In: *MNRAS* 106, p. 343. ISSN: 0035-8711. DOI: 10.1093/mnras/106.5.343.

Huber, D. et al. (Dec. 2011). "Testing Scaling Relations for Solar-like Oscillations from the Main Sequence to Red Giants Using Kepler Data". In: *ApJ* 743, p. 143. ISSN: 0004-637X. DOI: 10.1088/0004-637X/743/2/143.

Huber, D. et al. (Aug. 2017). "Asteroseismology and Gaia: Testing Scaling Relations Using 2200 Kepler Stars with TGAS Parallaxes". In: *ApJ* 844, p. 102. DOI: 10.3847/1538-4357/aa75ca.

Huber, D. et al. (June 2019). "A Hot Saturn Orbiting an Oscillating Late Subgiant Discovered by TESS". In: *AJ* 157, p. 245. DOI: 10.3847/1538-3881/ab1488.

Hunter, J. D. (May 2007). "Matplotlib: A 2D Graphics Environment". In: *Comput. Sci. Eng.* 9.3, pp. 90–95. ISSN: 1558-366X. DOI: 10.1109/MCSE.2007.55.

Iben Jr., I. (Feb. 1967). "Stellar Evolution.VI. Evolution from the Main Sequence to the Red-Giant Branch for Stars of Mass 1 M_{sun}, 1.25 M_{sun}, and 1.5 M_{sun}". In: *ApJ* 147, p. 624. ISSN: 0004-637X. DOI: 10.1086/149040.

Jermyn, A. S. et al. (Mar. 2023). "Modules for Experiments in Stellar Astrophysics (MESA): Time-dependent Convection, Energy Conservation, Automatic Differentiation, and Infrastructure". In: *ApJS* 265, p. 15. ISSN: 0067-0049. DOI: 10.3847/1538-4365/acae8d.

Jönsson, H. et al. (Sept. 2020). "APOGEE Data and Spectral Analysis from SDSS Data Release 16: Seven Years of Observations Including First Results from APOGEE-South". In: *AJ* 160, p. 120. ISSN: 0004-6256. DOI: 10.3847/1538-3881/aba592.

Kiefer, R. et al. (Feb. 2017). "Stellar Magnetic Activity and Variability of Oscillation Parameters: An Investigation of 24 Solar-like Stars Observed by Kepler". In: *A&A* 598, A77. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201628469.

Kingma, D. P. and J. Ba (Dec. 2014). "Adam: A Method for Stochastic Optimization". In: *arXiv e-prints*. arXiv: 1412.6980 [cs.LG].

Kippenhahn, R., A. Weigert, and E. Hofmeister (1967). "Methods for Calculating Stellar Evolution". In: *Methods in Computational Physics: Astrophysics*. Vol. 7. New York and London: Academic Press, pp. 129–190.

Kippenhahn, R., A. Weigert, and A. Weiss (Jan. 2013). *Stellar Structure and Evolution*. Second. Astronomy and Astrophysics Library. Heidelberg, Germany: Springer Berlin. ISBN: 978-3-642-30304-3. DOI: 10.1007/978-3-642-30304-3.

Kjeldsen, H. and T. R. Bedding (Jan. 1995). "Amplitudes of Stellar Oscillations: The Implications for Asteroseismology." In: *A&A* 293, pp. 87–106. ISSN: 0004-6361.

Kjeldsen, H., T. R. Bedding, and J. Christensen-Dalsgaard (Aug. 2008). "Correcting Stellar Oscillation Frequencies for Near-Surface Effects". In: *ApJ* 683, p. L175. DOI: 10.1086/591667.

Kjeldsen, H. et al. (Dec. 2005). "Solar-like Oscillations in α Centauri B". In: *ApJ* 635, pp. 1281–1290. DOI: 10.1086/497530.

Koposov, S. et al. (Feb. 2023). *joshspeagle/dynesty: v2.1.0*. Zenodo. DOI: 10.5281/zenodo.7600689.

Korn, A. J. et al. (Dec. 2007). "Atomic Diffusion and Mixing in Old Stars. I. Very Large Telescope FLAMES-UVES Observations of Stars in NGC 6397". In: *ApJ* 671, pp. 402–419. DOI: 10.1086/523098.

Kuszlewicz, J. S. et al. (Sept. 2019). "Bayesian Hierarchical Inference of Asteroseismic Inclination Angles". In: *MNRAS* 488, pp. 572–589. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1689.

Leavitt, H. S. (Jan. 1908). "1777 Variables in the Magellanic Clouds". In: *Ann. Harv. Coll. Obs.* 60, pp. 87–108.3.

Lebreton, Y., M. Goupil, and J. Montalbán (2014). "How Accurate Are Stellar Ages Based on Stellar Models?: I. The Impact of Stellar Models Uncertainties". In: *EAS Publ. Ser.* 65. Ed. by Y. Lebreton, D. Valls-Gabaud, and C. Charbonnel, pp. 99–176. ISSN: 1633-4760, 1638-1963. DOI: 10.1051/eas/1465004.

Leighton, R. B., R. W. Noyes, and G. W. Simon (Mar. 1962). "Velocity Fields in the Solar Atmosphere. I. Preliminary Report." In: *ApJ* 135, p. 474. ISSN: 0004-637X. DOI: 10.1086/147285.

Leistedt, B. and D. W. Hogg (Dec. 2017). "Hierarchical Probabilistic Inference of the Color-Magnitude Diagram and Shrinkage of Stellar Distance Uncertainties". In: *AJ* 154, p. 222. ISSN: 0004-6256. DOI: 10.3847/1538-3881/aa91d5.

Li, T. et al. (Mar. 2018). "Modelling Kepler Red Giants in Eclipsing Binaries: Calibrating the Mixing-Length Parameter with Asteroseismology". In: *MNRAS* 475, pp. 981–998. DOI: 10.1093/mnras/stx3079.

Li, T. et al. (Apr. 2022). "Modelling Stars with Gaussian Process Regression: Augmenting Stellar Model Grid". In: *MNRAS* 511.4, pp. 5597–5610. ISSN: 0035-8711. DOI: 10.1093/mnras/stac467.

Lightkurve Collaboration et al. (Dec. 2018). "Lightkurve: Kepler and TESS Time Series Analysis in Python". In: *Astrophys. Source Code Libr.*, ascl:1812.013.

Lindegren, L. et al. (Aug. 2018). "Gaia Data Release 2. The Astrometric Solution". In: *A&A* 616, A2. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201832727.

Lund, M. N. and D. R. Reese (Jan. 2018). "Tutorial: Asteroseismic Stellar Modelling with AIMS". In: *Asteroseismol. Exopl. List. Stars Search. New Worlds*. Vol. 49. Astrophysics and Space Science Proceedings. eprint: arXiv:1711.01896, p. 149. DOI: 10.1007/978-3-319-59315-9_8.

Lund, M. N. et al. (Feb. 2017). "Standing on the Shoulders of Dwarfs: The Kepler Asteroseismic LEGACY Sample. I. Oscillation Mode Parameters". In: *ApJ* 835, p. 172. ISSN: 0004-637X. DOI: 10.3847/1538-4357/835/2/172.

Lyttle, A. J. et al. (Aug. 2021). "Hierarchically Modelling Kepler Dwarfs and Subgiants to Improve Inference of Stellar Properties with Asteroseismology". In: *MNRAS* 505, pp. 2427–2446. ISSN: 0035-8711. DOI: 10.1093/mnras/stab1368.

Magic, Z., A. Weiss, and M. Asplund (Jan. 2015). "The Stagger-grid: A Grid of 3D Stellar Atmosphere Models. III. The Relation to Mixing Length Convection Theory". In: *A&A* 573, A89. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201423760.

Majewski, S. R. et al. (Sept. 2017). "The Apache Point Observatory Galactic Evolution Experiment (APOGEE)". In: *AJ* 154, p. 94. DOI: 10.3847/1538-3881/aa784d.

Martić, M. et al. (Nov. 1999). "Evidence for Global Pressure Oscillations on Procyon". In: *A&A* 351, pp. 993–1002.

Masters, D. and C. Luschi (Apr. 2018). "Revisiting Small Batch Training for Deep Neural Networks". In: *arXiv e-prints*. arXiv: 1804.07612 [cs.LG].

Mazumdar, A. et al. (Dec. 2012a). "Acoustic Glitches in Solar-Type Stars from Kepler". In: *Astron. Nachrichten* 333, pp. 1040–1043. ISSN: 0004-6337. DOI: 10.1002/asna.201211825.

Mazumdar, A. et al. (Apr. 2012b). "Seismic Detection of Acoustic Sharp Features in the CoRoT Target HD 49933". In: *A&A* 540, A31. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201118495.

Mazumdar, A. et al. (Feb. 2014). "Measurement of Acoustic Glitches in Solar-type Stars from Oscillation Frequencies Observed by Kepler". In: *ApJ* 782, p. 18. DOI: 10.1088/0004-637X/782/1/18.

Metcalfe, T. S. et al. (Mar. 2012). "Asteroseismology of the Solar Analogs 16 Cyg A and B from Kepler Observations". In: *ApJ* 748, p. L10. ISSN: 0004-637X. DOI: 10.1088/2041-8205/748/1/L10.

Metcalfe, T. S. et al. (Oct. 2014). "Properties Of 42 Solar-Type Kepler Targets From The Asteroseismic Modeling Portal". In: *ApJS* 214.2, p. 27. ISSN: 0067-0049. DOI: 10.1088/0067-0049/214/2/27.

Meurer, A. et al. (Jan. 2017). "SymPy: Symbolic Computing in Python". In: *PeerJ Comput. Sci.* 3, e103. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.103.

Miglio, A. et al. (Sept. 2010). "Evidence for a Sharp Structure Variation inside a Red-Giant Star". In: *A&A* 520, p. L6. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201015442.

Montalbán, J. et al. (Jan. 2021). "Chronologically Dating the Early Assembly of the Milky Way". In: *Nat. Astron.* 5, pp. 640–647. ISSN: 2397-3366. DOI: 10.1038/s41550-021-01347-7.

Monteiro, M. J. P. F. G., J. Christensen-Dalsgaard, and M. J. Thompson (Mar. 1994). "Seismic Study of Overshoot at the Base of the Solar Convective Envelope". In: *A&A* 283, pp. 247–262. ISSN: 0004-6361.

Monteiro, M. J. P. F. G. and M. J. Thompson (Jan. 1998). "On the Seismic Signature of the HeII Ionization Zone in Stellar Envelopes". In: *New Eyes See Sun Stars*. Vol. 185, p. 317.

— (Aug. 2005). "Seismic Analysis of the Second Ionization Region of Helium in the Sun - I. Sensitivity Study and Methodology". In: *MNRAS* 361, pp. 1187–1196. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2005.09246.x.

Morton, T. D. and J. N. Winn (Nov. 2014). "Obliquities of Kepler Stars: Comparison of Single- and Multiple-transit Systems". In: *ApJ* 796, p. 47. ISSN: 0004-637X. DOI: 10.1088/0004-637X/796/1/47.

Mosser, B. et al. (July 2010). "Red-Giant Seismic Properties Analyzed with CoRoT". In: *A&A* 517, A22. ISSN: 0004-6361, 1432-0746. DOI: 10.1051/0004-6361/201014036.

Mosser, B. et al. (Aug. 2011). "Mixed Modes in Red-Giant Stars Observed with CoRoT". In: *A&A* 532, A86. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201116825.

Mosser, B. et al. (Dec. 2015). "Period Spacings in Red Giants. I. Disentangling Rotation and Revealing Core Structure Discontinuities". In: *A&A* 584, A50. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201527075.

Nielsen, M. B. et al. (Feb. 2021). "PBjam: A Python Package for Automating Asteroseismology of Solar-like Oscillators". In: *AJ* 161, p. 62. ISSN: 0004-6256. DOI: 10.3847/1538-3881/abcd39.

Nissen, P. E. et al. (Dec. 2017). "High-Precision Abundances of Elements in Kepler LEGACY Stars. Verification of Trends with Stellar Age". In: *A&A* 608, A112. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201731845.

Nsamba, B. et al. (July 2018). "Asteroseismic Modelling of Solar-Type Stars: Internal Systematics from Input Physics and Surface Correction Methods". In: *MNRAS* 477, pp. 5052–5063. DOI: 10.1093/mnras/sty948.

Ochsenbein, F., P. Bauer, and J. Marcout (Apr. 2000). "The VizieR Database of Astronomical Catalogues". In: *A&AS* 143, pp. 23–32. ISSN: 0365-0138. DOI: 10.1051/aas:2000169.

Önehag, A., B. Gustafsson, and A. J. Korn (Feb. 2014). "Abundances and Possible Diffusion of Elements in M 67 Stars". In: *A&A* 562, A102. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201322663.

Paquette, C. et al. (May 1986). "Diffusion Coefficients for Stellar Plasmas". In: *ApJS* 61, pp. 177–195. DOI: 10.1086/191111.

Paxton, B. et al. (Jan. 2011). "Modules for Experiments in Stellar Astrophysics (MESA)". In: *ApJS* 192, p. 3. ISSN: 0067-0049. DOI: 10.1088/0067-0049/192/1/3.

Paxton, B. et al. (Sept. 2013). "Modules for Experiments in Stellar Astrophysics (MESA): Planets, Oscillations, Rotation, and Massive Stars". In: *ApJS* 208, p. 4. ISSN: 0067-0049. DOI: 10.1088/0067-0049/208/1/4.

Paxton, B. et al. (Sept. 2015). "Modules for Experiments in Stellar Astrophysics (MESA): Binaries, Pulsations, and Explosions". In: *ApJS* 220, p. 15. ISSN: 0067-0049. DOI: 10.1088/0067-0049/220/1/15.

Paxton, B. et al. (Feb. 2018). "Modules for Experiments in Stellar Astrophysics (MESA): Convective Boundaries, Element Diffusion, and Massive Star Explosions". In: *ApJS* 234, p. 34. ISSN: 0067-0049. DOI: 10.3847/1538-4365/aaa5a8.

Paxton, B. et al. (July 2019). "Modules for Experiments in Stellar Astrophysics (MESA): Pulsating Variable Stars, Rotation, Convective Boundaries, and Energy Conservation". In: *ApJS* 243, p. 10. ISSN: 0067-0049. DOI: 10.3847/1538-4365/ab2241.

Payne, C. H. (Jan. 1925). "Stellar Atmospheres; a Contribution to the Observational Study of High Temperature in the Reversing Layers of Stars." PhD thesis. Harvard University.

Peimbert, A., M. Peimbert, and V. Luridiana (Oct. 2016). "The Primordial Helium Abundance and the Number of Neutrino Families". In: *Rev. Mex. Astron. Astrofisica* 52, pp. 419–424.

Phan, D., N. Pradhan, and M. Jankowiak (Dec. 2019). "Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro". In: *arXiv e-prints*, arXiv:1912.11554. arXiv: 1912.11554 [stat.ML].

Pinsonneault, M. H. et al. (Apr. 2012). "A Revised Effective Temperature Scale for the Kepler Input Catalog". In: *ApJS* 199, p. 30. ISSN: 0067-0049. DOI: 10.1088/0067-0049/199/2/30.

Pinsonneault, M. H. et al. (Dec. 2014). "The APOKASC Catalog: An Asteroseismic and Spectroscopic Joint Survey of Targets in the Kepler Fields". In: *ApJS* 215, p. 19. DOI: 10.1088/0067-0049/215/2/19.

Pinsonneault, M. H. et al. (Dec. 2018). "The Second APOKASC Catalog: The Empirical Approach". In: *ApJS* 239, p. 32. DOI: 10.3847/1538-4365/aaebfd.

Pitrou, C. et al. (Sept. 2018). "Precision Big Bang Nucleosynthesis with Improved Helium-4 Predictions". In: *Phys. Rep.* 754, pp. 1–66. ISSN: 0370-1573. DOI: 10.1016/j.physrep.2018.04.005.

Qian, N. (Jan. 1999). "On the Momentum Term in Gradient Descent Learning Algorithms". In: *Neural Netw.* 12.1, pp. 145–151. ISSN: 0893-6080. DOI: 10.1016/S0893-6080(98)00116-6.

Ramírez, I., J. Meléndez, and M. Asplund (Dec. 2009). "Accurate Abundance Patterns of Solar Twins and Analogs. Does the Anomalous Solar Chemical Composition Come from Planet Formation?" In: *A&A* 508.1, pp. L17–L20. ISSN: 0004-6361. DOI: 10.1051/0004-6361/200913038.

Rasmussen, C. E. and C. K. I. Williams (2006). *Gaussian Processes for Machine Learning*. Adaptive Computation and Machine Learning. Cambridge, MA: MIT Press. ISBN: 13 978-0-262-18253-9.

Rauer, H. et al. (Nov. 2014). "The PLATO 2.0 Mission". In: *Exp. Astron.* 38, pp. 249–330. ISSN: 0922-6435. DOI: 10.1007/s10686-014-9383-4.

Rauer, H. et al. (Sept. 2016). "The PLATO Mission". In: *Astron. Nachrichten* 337, p. 961. ISSN: 0004-6337. DOI: 10.1002/asna.201612408.

Rendle, B. M. et al. (Mar. 2019). "Aims – a New Tool for Stellar Parameter Determinations Using Asteroseismic Constraints". In: *MNRAS* 484.1, pp. 771–786. ISSN: 0035-8711. DOI: 10.1093/mnras/stz031.

Ribas, I. et al. (Mar. 2000). "Chemical Composition of Eclipsing Binaries: A New Approach to the Helium-to-Metal Enrichment Ratio". In: *MNRAS* 313, pp. 99–111. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2000.03195.x.

Ricker, G. R. et al. (Jan. 2015). "Transiting Exoplanet Survey Satellite (TESS)". In: *J. Astron. Telesc. Instrum. Syst.* 1, p. 014003. DOI: 10.1117/1.JATIS.1.1.014003.

Riess, A. G. et al. (July 2018). "Milky Way Cepheid Standards for Measuring Cosmic Distances and Application to Gaia DR2: Implications for the Hubble Constant". In: *ApJ* 861, p. 126. DOI: 10.3847/1538-4357/aac82e.

Rogers, F. J. and A. Nayfonov (Sept. 2002). "Updated and Expanded OPAL Equation-of-State Tables: Implications for Helioseismology". In: *ApJ* 576, pp. 1064–1074. ISSN: 0004-637X. DOI: 10.1086/341894.

Ruder, S. (Sept. 2016). "An Overview of Gradient Descent Optimization Algorithms". In: *arXiv e-prints*. arXiv: 1609.04747 [cs.LG].

Rumelhart, D. E., G. E. Hinton, and R. J. Williams (Oct. 1986). "Learning Representations by Back-Propagating Errors". In: *Nature* 323.6088, pp. 533–536. ISSN: 1476-4687. DOI: 10.1038/323533a0.

Russell, H. N. (May 1914). "Relations Between the Spectra and Other Characteristics of the Stars". In: *Pop. Astron.* 22, pp. 275–294. ISSN: 0197-7482.

Salvatier, J., T. V. Wiecki, and C. Fonnesbeck (Apr. 2016). "Probabilistic Programming in Python Using PyMC3". In: *PeerJ Comput. Sci.* 2, e55. ISSN: 2376-5992. DOI: 10.7717/peerj-cs.55.

Sandage, A. (Sept. 1957). "Observational Approach to Evolution. III. Semiempirical Evolution Tracks for M67 and M3." In: *ApJ* 126, p. 326. ISSN: 0004-637X. DOI: 10.1086/146405.

Schofield, M. et al. (Mar. 2019). "The Asteroseismic Target List for Solar-like Oscillators Observed in 2 Minute Cadence with the Transiting Exoplanet Survey Satellite". In: *ApJS* 241, p. 12. ISSN: 0067-0049. DOI: 10.3847/1538-4365/ab04f5.

Schwarzschild, M. (Sept. 1946). "On the Helium Content of the Sun." In: *ApJ* 104, p. 203. ISSN: 0004-637X. DOI: 10.1086/144846.

Scott, P. et al. (Jan. 2015). "The Elemental Composition of the Sun. I. The Intermediate Mass Elements Na to Ca". In: *A&A* 573, A25. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201424109.

Scutt, O. J. et al. (Feb. 2023). "Asteroseismology of δ Scuti Stars: Emulating Model Grids Using a Neural Network". In: *arXiv e-prints*. DOI: 10.48550/arXiv.2302.11025.

Serenelli, A. M. and S. Basu (Aug. 2010). "Determining the Initial Helium Abundance of the Sun". In: *ApJ* 719, pp. 865–872. ISSN: 0004-637X. DOI: 10.1088/0004-637X/719/1/865.

Serenelli, A. M. et al. (Nov. 2009). "New Solar Composition: The Problem with Solar Models Revisited". In: *ApJ* 705, pp. L123–L127. DOI: 10.1088/0004-637X/705/2/L123.

Serenelli, A. M. et al. (Dec. 2017). "The First APOKASC Catalog of Kepler Dwarf and Subgiant Stars". In: *ApJS* 233, p. 23. ISSN: 0067-0049. DOI: 10.3847/1538-4365/aa97df.

Silva Aguirre, V. et al. (Sept. 2015). "Ages and Fundamental Properties of Kepler Exoplanet Host Stars from Asteroseismology". In: *MNRAS* 452.2, p. 2127. DOI: 10.1093/mnras/stv1388.

Silva Aguirre, V. et al. (Feb. 2017). "Standing on the Shoulders of Dwarfs: The Kepler Asteroseismic LEGACY Sample. II.Radii, Masses, and Ages". In: *ApJ* 835, p. 173. ISSN: 0004-637X. DOI: 10.3847/1538-4357/835/2/173.

Silva Aguirre, V. et al. (Jan. 2020). "Detection and Characterization of Oscillating Red Giants: First Results from the TESS Satellite". In: *ApJ* 889.2, p. L34. ISSN: 2041-8205. DOI: 10.3847/2041-8213/ab6443.

Skilling, J. (Nov. 2004). "Nested Sampling". In: *24th Int. Workshop Bayesian Inference Maximum Entropy Methods Sci. Eng.* Vol. 735. American Institute of Physics Conference Series, pp. 395–405. DOI: 10.1063/1.1835238.

— (Dec. 2006). "Nested Sampling for General Bayesian Computation". In: *Bayesian Anal.* 1.4, pp. 833–859. ISSN: 1936-0975, 1931-6690. DOI: 10.1214/06-BA127.

Skrutskie, M. F. et al. (Feb. 2006). "The Two Micron All Sky Survey (2MASS)". In: *AJ* 131, pp. 1163–1183. DOI: 10.1086/498708.

Sonoi, T. et al. (Nov. 2015). "Surface-Effect Corrections for Solar-like Oscillations Using 3D Hydrodynamical Simulations. I. Adiabatic Oscillations". In: *A&A* 583, A112. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201526838.

Speagle, J. S. (Apr. 2020). "DYNESTY: A Dynamic Nested Sampling Package for Estimating Bayesian Posteriors and Evidences". In: *MNRAS* 493, pp. 3132–3158. ISSN: 0035-8711. DOI: 10.1093/mnras/staa278.

Spergel, D. et al. (Mar. 2015). *Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report*. DOI: 10.48550/arXiv.1503.03757.

Stancliffe, R. J. et al. (Feb. 2016). "Confronting Uncertainties in Stellar Physics. II. Exploring Differences in Main-Sequence Stellar Evolution Tracks". In: *A&A* 586, A119. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201527099.

Stanton, L. G. and M. S. Murillo (Apr. 2016). "Ionic Transport in High-Energy-Density Matter". In: *Phys. Rev. E* 93, p. 043203. ISSN: 1063-651X. DOI: 10.1103/PhysRevE.93.043203.

Tassoul, M. (Aug. 1980). "Asymptotic Approximations for Stellar Nonradial Pulsations". In: *ApJS* 43, pp. 469–490. DOI: 10.1086/190678.

Tayar, J. et al. (May 2017). "The Correlation between Mixing Length and Metallicity on the Giant Branch: Implications for Ages in the Gaia Era". In: *ApJ* 840, p. 17. DOI: 10.3847/1538-4357/aa6a1e.

Tayar, J. et al. (Dec. 2020). "A Guide to Realistic Uncertainties on Fundamental Properties of Solar-Type Exoplanet Host Stars". In: *ArXiv e-prints*, arXiv:2012.07957. arXiv: 2012.07957 [astro-ph.EP].

— (Mar. 2022). "A Guide to Realistic Uncertainties on the Fundamental Properties of Solar-type Exoplanet Host Stars". In: *ApJ* 927, p. 31. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ac4bbc.

Thoul, A. A., J. N. Bahcall, and A. Loeb (Feb. 1994). "Element Diffusion in the Solar Interior". In: *ApJ* 421, pp. 828–842. DOI: 10.1086/173695.

Townsend, R. H. D. and S. A. Teitler (Nov. 2013). "GYRE: An Open-Source Stellar Oscillation Code Based on a New Magnus Multiple Shooting Scheme". In: *MNRAS* 435, pp. 3406–3418. ISSN: 0035-8711. DOI: 10.1093/mnras/stt1533.

Trampedach, R. et al. (Dec. 2014). "Improvements to Stellar Structure Models, Based on a Grid of 3D Convection Simulations - II. Calibrating the Mixing-Length Formulation". In: *MNRAS* 445, pp. 4366–4384. ISSN: 0035-8711. DOI: 10.1093/mnras/stu2084.

Trumpler, R. J. (Jan. 1930). "Preliminary Results on the Distances, Dimensions and Space Distribution of Open Star Clusters". In: *Lick Obs. Bull.* 420, pp. 154–188. ISSN: 0075-9317. DOI: 10.5479/ADS/bib/1930LicOB.14.154T.

Ulrich, R. K. (July 1986). "Determination of Stellar Ages from Asteroseismology". In: *ApJ* 306, pp. L37–L40. DOI: 10.1086/184700.

Ulrich, R. K. (Dec. 1970). "The Five-Minute Oscillations on the Solar Surface". In: *ApJ* 162, p. 993. DOI: 10.1086/150731.

Valle, G. et al. (Mar. 2015). "Uncertainties in Asteroseismic Grid-Based Estimates of Stellar Ages. SCEPtER: Stellar CharactEristics Pisa Estimation gRid". In: *A&A* 575, A12. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201424686.

van Saders, J. L. et al. (Jan. 2016). "Weakened Magnetic Braking as the Origin of Anomalously Rapid Rotation in Old Field Stars". In: *Nature* 529, pp. 181–184. ISSN: 0028-0836. DOI: 10.1038/nature16168.

Verma, K. et al. (Aug. 2014). "Asteroseismic Estimate of Helium Abundance of a Solar Analog Binary System". In: *ApJ* 790, p. 138. DOI: 10.1088/0004-637X/790/2/138.

Verma, K. et al. (Oct. 2016). "Asteroseismic Determination of Fundamental Parameters of Sun-like Stars Using Multilayered Neural Networks". In: *MNRAS* 461, pp. 4206–4214. ISSN: 0035-8711. DOI: 10.1093/mnras/stw1621.

Verma, K. et al. (Mar. 2017). "Seismic Measurement of the Locations of the Base of Convection Zone and Helium Ionization Zone for Stars in the Kepler Seismic LEGACY Sample". In: *ApJ* 837, p. 47. DOI: 10.3847/1538-4357/aa5da7.

Verma, K. et al. (Mar. 2019). "Helium Abundance in a Sample of Cool Stars: Measurements from Asteroseismology". In: *MNRAS* 483, pp. 4678–4694. ISSN: 0035-8711. DOI: 10.1093/mnras/sty3374.

Verma, K. et al. (Sept. 2022). "Advanced Asteroseismic Modelling: Breaking the Degeneracy between Stellar Mass and Initial Helium Abundance". In: *MNRAS* 515, pp. 1492–1509. ISSN: 0035-8711. DOI: 10.1093/mnras/stac1860.

Verner, G. A. (2005). "Helioseismic Inferences on Structural Change over the Solar Cycle". PhD thesis. Birmingham, United Kingdom: University of Birmingham.

Viani, L. S. et al. (May 2018). "Investigating the Metallicity-Mixing-length Relation". In: *ApJ* 858, p. 28. DOI: 10.3847/1538-4357/aab7eb.

Villante, F. L. et al. (May 2014). "The Chemical Composition of the Sun from Helioseismic and Solar Neutrino Data". In: *ApJ* 787, p. 13. DOI: 10.1088/0004-637X/787/1/13.

Virtanen, P. et al. (2020). "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nat. Methods* 17, pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

Waskom, M. L. (2021). "Seaborn: Statistical Data Visualization". In: *J. Open Source Softw.* 6.60, p. 3021. DOI: 10.21105/joss.03021.

Weiss, A. and H. Schlattl (Aug. 2008). "GARSTEC—the Garching Stellar Evolution Code. The Direct Descendant of the Legendary Kippenhahn Code". In: *Ap&SS* 316, pp. 99–106. DOI: 10.1007/s10509-007-9606-5.

West, C. and A. Heger (Sept. 2013). "Metallicity-Dependent Galactic Isotopic Decomposition for Nucleosynthesis". In: *ApJ* 774, p. 75. DOI: 10.1088/0004-637X/774/1/75.

White, T. R. et al. (Dec. 2011). "Calculating Asteroseismic Diagrams for Solar-like Oscillations". In: *ApJ* 743, p. 161. ISSN: 0004-637X. DOI: 10.1088/0004-637X/743/2/161.

Woodard, M. and H. Hudson (Oct. 1983). "Frequencies, Amplitudes and Linewidths of Solar Oscillations from Total Irradiance Observations". In: *Nature* 305, pp. 589–593. ISSN: 0028-0836. DOI: 10.1038/305589a0.

Yu, J. et al. (June 2018). "Asteroseismology of 16,000 Kepler Red Giants: Global Oscillation Parameters, Masses, and Radii". In: *ApJS* 236, p. 42. ISSN: 0067-0049. DOI: 10.3847/1538-4365/aaaf74.

Zinn, J. C. et al. (June 2019a). "Confirmation of the Gaia DR2 Parallax Zero-point Offset Using Asteroseismology and Spectroscopy in the Kepler Field". In: *ApJ* 878, p. 136. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab1f66.

Zinn, J. C. et al. (Nov. 2019b). "Testing the Radius Scaling Relation with Gaia DR2 in the Kepler Field". In: *ApJ* 885, p. 166. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab44a9.

Appendix A

HIERARCHICALLY MODELLING *KEPLER* DWARFS AND SUBGIANTS TO IMPROVE INFERENCE OF STELLAR PROPERTIES WITH ASTEROSEISMOLOGY

In this chapter, I present the appendix of Lyttle et al. (2021) with minor modification. It follows from and relates to Chapter 3. Firstly, I explain the methodology behind the neural network emulator in Section A.1. In Section A.2, I briefly show the beta distribution which was used as a prior for some parameters in the main body of work. Finally, I test the hierarchical model on a synthetic population of stars in Section A.3.

A.1 Artificial Neural Network

Once we constructed our grid of models, we needed a way to continuously sample the grid for use in our statistical model. We opted to train an Artificial Neural Network (ANN). The ANN is advantageous over interpolation because it scales well with dimensionality, training and evaluation is fast, and gradient evaluation is easy due to its roots in linear algebra (Haykin, 2007). We trained an ANN on the data generated by the grid of stellar models to map fundamentals to observables. Firstly, we split the grid into a *train* and *validation* dataset for tuning the

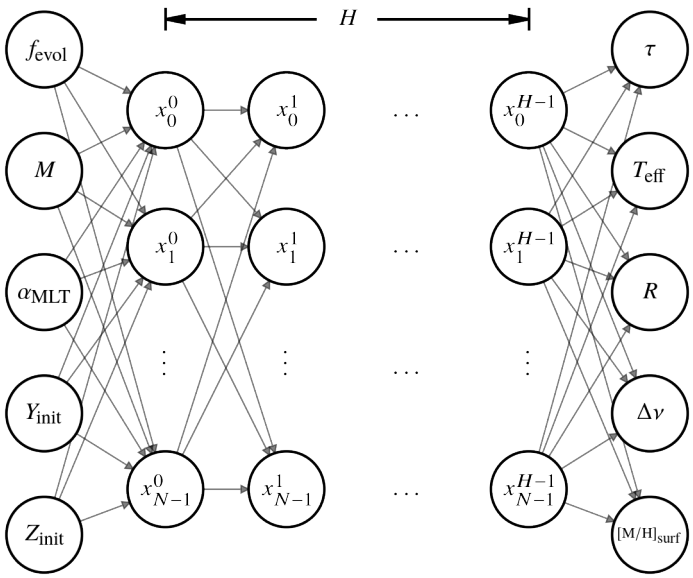


Figure A.1: An artificial neural network comprising H hidden layers with N neurons per layer. Arrows connecting the nodes represent tunable weights.

ANN, as described in Appendix A.1.1. We then tested a multitude of ANN configurations and training data inputs, repeatedly evaluating them with the validation dataset in Appendix A.1.2. In Appendix A.1.3, we reserved a set of randomly generated, off-grid stellar models as our final *test* dataset to evaluate the approximation ability of the best-performing ANN independently from our train and validation data. Here, we briefly describe the theory and motivation behind the ANN.

An ANN is a network of artificial *neurons* which each transform an input vector, $\boldsymbol{x}$ based on trainable weights, $\boldsymbol{w}$ and a bias, b. The weights are represented by the connections between neurons and the bias is a unique scalar associated with each neuron. A multi-layered ANN is where neurons are arranged into a series of layers such that any neuron in layer $j-1$ is connected to at least one of the neurons in layer j.

In this work, we considered a fully-connected ANN, where each neuron in layer $j-1$ is connected to every neuron in layer j. The output of the k-th neuron in layer j is,

$$x_{j,k} = f_j(\boldsymbol{w}_{j,k} \cdot \boldsymbol{x}_{j-1} + b_{j,k}), \tag{A.1}$$

where f_j is the *activation* function for the j-th layer, $\boldsymbol{w}_{j,k}$ are the weights connecting all the neurons in layer $j-1$ to the current neuron, and $b_{j,k}$ is the bias. This generalises such that the output of the j-th layer is,

$$\boldsymbol{x}_j = f_j(\boldsymbol{W}_j \cdot \boldsymbol{x}_{j-1} + \boldsymbol{b}_j), \tag{A.2}$$

where $\boldsymbol{W}_j$ is the matrix of weights leading to all neurons in the j-th layer. For a regression ANN, we typically have a linear activation function applied to the final output layer. Layers of neurons between the input and output layers are called *hidden* layers. Therefore, the output of a network of H hidden layers with initial input $\mathbb{X}$ is,

$$\widetilde{\mathbb{Y}} = \boldsymbol{W}_H \cdot f_{H-1}(\ldots f_1(\boldsymbol{W}_1 \cdot f_0(\boldsymbol{W}_0 \cdot \mathbb{X} + \boldsymbol{b}_0) + \boldsymbol{b}_1)) + \boldsymbol{b}_H. \tag{A.3}$$

We also restricted our configuration to an ANN with the same number of neurons, N in each hidden layer. Hereafter, we refer to our choice of neurons per layer, N and hidden layers, H as the *architecture* (see Fig. A.1).

To fit the ANN, we used a set of training data, $\mathbb{D}_{\text{train}} = \{\mathbb{X}_i, \mathbb{Y}_i\}_{i=1}^{N_{\text{train}}}$ comprising N_{train} input-output pairs. We split the training data into pseudo-random batches, $\mathbb{D}_{\text{batch}}$ because this has been shown to improve ANN stability and computational efficiency (Masters and Luschi, 2018). The set of predictions made for each batch is evaluated using a *loss* function which primarily comprises an error function, $E(\mathbb{D}_{\text{batch}})$ to quantify the difference between the training data outputs ($\mathbb{Y}$) and predictions ($\widetilde{\mathbb{Y}}$). We also considered an additional term to the loss called *regularisation* which helps reduce over-fitting (Goodfellow, Bengio, and Courville, 2016). During fitting, the weights are updated after each batch using an algorithm called the *optimizer*, back-propagating the error with the goal of minimising the loss such that $\widetilde{\mathbb{Y}} \approx \mathbb{Y}$ (see e.g. Rumelhart, Hinton, and R. J. Williams, 1986).

We trained the ANN using TensorFlow (Abadi et al., 2016). We varied the architecture, number of batches, choice of loss function, optimizer, and regularisation during the optimisation phase. For each set of ANN parameters, we initialised the ANN with a random set of weights and biases and minimised the loss over a given number of *epochs*. An epoch is defined as one

iteration through the entire training dataset, $\mathbb{D}_{\text{train}}$. We tracked the loss for each ANN using an independent validation dataset to determine the most effective choice of ANN parameters (see Appendix A.1.2).

A.1.1 Train, Validation, and Test Data

We built the train and validation datasets from the outputs of the grid of stellar models in Section 3.3.1. This included the input parameters: M, α_{MLT}, Y_{init}, and the initial heavy-elements fraction, Z_{init}. We also included the T_{eff}, $\log g$, $\Delta\nu$, stellar age (τ), radius (R), surface metallicity ($[\text{M/H}]_{\text{surf}}$), and other chemical composition information generated by the models. We determined the fractional MS lifetime, $f_{\text{MS}} = \tau/\tau_{\text{MS}}$, of each evolutionary track by taking τ_{MS} as the age when the central hydrogen fraction, $X_c < 0.01$. We then cut data where $f_{\text{MS}} < 0.01$ to remove points on the grid prior to the MS. Once we had refined the data from the grid of models, we randomly sampled 7.736×10^6 points to use as the training dataset, with the remaining $\sim 2 \times 10^6$ points given to the validation dataset. We varied our choice of ANN input and output parameters among those available in the dataset during tuning (see Appendix A.1.2).

Additionally, we produced a test dataset of $\sim 2 \times 10^6$ stellar models evolved using MESA. Values for the initial M, [M/H], Y, and α_{MLT} were chosen randomly within the range of grid parameters described in Table 3.2 such that they spanned the breadth of the grid in an unbiased manner. We prepared this dataset in the same way as the training set, but also constrained it to $\tau < 15$ Gyr because we consider ages above ~ 15 Gyr unphysical and such points are sparse in the training data. The test dataset was set aside and evaluated on the final ANN.

A.1.2 Tuning

We trained an ANN to reproduce stellar observables according to our choice of physics with greater accuracy than typical observational precisions. We experimented with a variety of ANN parameter choices, such as the architecture, activation function, optimisation algorithm, and loss function. We tuned the ANN parameters by varying them in both a grid-based and heuristic approach, each time evaluating the accuracy using the validation dataset.

During initial tuning, we found that having stellar age as an input was unstable, because it varied heavily with the other input parameters. We mitigated this by introducing an input to describe the fraction of time a star had spent in a given evolutionary phase, $f_{\rm evol}$.

$$f_{\rm evol} = \begin{cases} f_{\rm MS}, & f_{\rm MS} \leq 1 \\ 1 + \frac{\tau - \tau_{\rm MS}}{\tau_{\rm end} - \tau_{\rm MS}}, & f_{\rm MS} > 1 \end{cases} \tag{A.4}$$

where $\tau_{\rm end}$ is the age of the star at the end of the track,

$$f_{\rm MS} = \frac{\tau}{\tau_{\rm MS}}, \tag{A.5}$$

and $\tau_{\rm MS}$ is the MS lifetime. A star with $0.01 < f_{\rm evol} \leq 1.0$ is in its MS phase, burning hydrogen in its core, and $1.0 < f_{\rm evol} \leq 2.0$ has left the MS. Consequently, $f_{\rm evol}$ gives the ANN information about the internal state of the star which affects the output observables. Otherwise, $f_{\rm evol}$ has little physical meaning, although it could be interpreted as a measure of the evolutionary phase of the star.

We also observed that the ANN trained poorly in areas with a high rate of change in observables, likely because of poor grid coverage in those areas. To bias training to such areas, we calculated the gradient in $T_{\rm eff}$ and $\log g$ between each point for each stellar evolutionary track and used them as optional weights to the loss during tuning. These weights multiplied the difference between the ANN prediction and the training data in our chosen loss function.

After preliminary tuning, we chose the ANN input and output parameters to be $\mathbb{X} = \{f_{\rm evol}, M, \alpha_{\rm MLT}, Y_{\rm init}, Z_{\rm init}\}$ and $\mathbb{Y} = \{\log(\tau), T_{\rm eff}, R, \Delta\nu, [{\rm M/H}]_{\rm surf}\}$ respectively. A generalised form of our neural network is depicted in Fig. A.1. The inputs corresponded to initial conditions in the stellar modelling code and the outputs corresponded to surface conditions throughout the lifetime of the star, except for age which is mapped from $f_{\rm evol}$.

We standardised the training dataset by subtracting the median, $\mu_{1/2}$ and dividing by the standard deviation, σ for each input and output parameter. We found that the ANN performed better when the training data was scaled in this way as opposed to no scaling at all. We present

Table A.1: The median, $\mu_{1/2}$ and standard deviation, σ for each parameter in the training data, used to standardise the dataset.

Input parameters.

	$f_{\rm evol}$	M ($M_\odot$)	$\alpha_{\rm MLT}$	$Y_{\rm init}$	$Z_{\rm init}$
$\mu_{1/2}$	0.865	1.000	1.900	0.280	0.017
σ	0.651	0.118	0.338	0.028	0.011

Output parameters.

	$\log(\tau/{\rm Gyr})$	$T_{\rm eff}$ (K)	R ($R_\odot$)	$\Delta\nu$ (μHz)	$[{\rm M/H}]_{\rm surf}$ (dex)
$\mu_{1/2}$	0.790	5566.772	1.224	100.720	0.081
σ	0.467	601.172	0.503	42.582	0.361

the parameters used to standardise the training dataset in Table A.1.

We found that the optimal choice of architecture (N and H) varied depending on our choice of other ANN parameters. Therefore, each time we explored a new parameter, we trained an ANN with a grid of (N, H) ranging from $(32, 2)$ to $(512, 10)$.

We evaluated the performance of three activation functions: the hyperbolic-tangent, the rectified linear unit (ReLU; Hahnloser et al., 2000; Glorot, Bordes, and Bengio, 2011) and the exponential linear unit (ELU; Clevert, Unterthiner, and Hochreiter, 2015). Although the ReLU activation function out-performed the other two in speed and accuracy, the resulting ANN output was not smooth. The discontinuity in the ReLU function, $f(x) = \max(0, x)$ in turn caused the output to be discontinuous. This made the ANN difficult to sample for our choice of statistical model (see Section 3.3.2). Out of the remaining activation functions, ELU performed the best, providing a smooth output which was well-suited to our probabilistic sampling methods.

We compared the performance of two optimisers: Adam (Kingma and Ba, 2014) and stochastic gradient descent (SGD; see e.g. Ruder, 2016) with and without momentum (Qian, 1999). Both optimizers required a choice of *learning rate* which determined the rate at which the weights were adjusted during training. We found that Adam performed well, but the validation loss was noisy as a function of epochs as it struggled to converge. The SGD optimizer was less noisy than Adam, but it was difficult to tune the learning rate. However, SGD with momentum allowed for more adaptive weight updates and out-performed the other configurations.

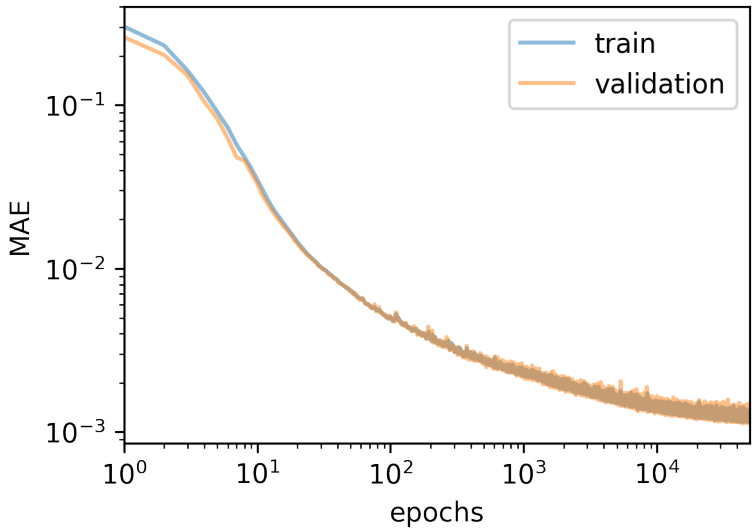


Figure A.2: The MAE as a function of epochs for the train and validation datasets.

There are several ways to reduce over-fitting, from minimising the complexity of the architecture, to increasing the size and coverage of the training dataset. One alternative is to introduce weight regularisation. So-called L2 regularisation adds a term, $\sim \lambda_k \sum_i w_{i,k}^2$ to the loss function for a given hidden layer, k which acts to keep the weights small. We varied the magnitude of λ_k and found that if it was too large it would dominate the loss function, but if it was too small then performance on the validation dataset was poorer.

We compared the choice of two error functions: mean squared error (MSE) and mean absolute error (MAE). The former is widely used among ANNs because it is more sensitive to large errors. However, we tracked both metrics regardless of which was added to the loss function and found that MAE converged faster. Although MAE is less effective at large errors, we found that these were typically at the edges of the grid and the accuracy was good enough everywhere else.

After extensive tuning, we opted for an ANN with $N = 128$ neurons in each of $H = 6$ hidden layers. Each of the hidden layers used an ELU activation function and L2 weight regularisation with $\lambda = 1 \times 10^{-6}$. We trained the ANN for 50 000 epochs with a 500 training data batches each containing 15 472 input-output pairs. To fit the ANN, we used an SGD optimiser with an initial learning rate of 1×10^{-4} and momentum of 0.999 with an MAE loss function. Training took $\sim$ 48 h on an NVidia Tesla V100 graphics processing unit (GPU). In Fig. A.2 we show

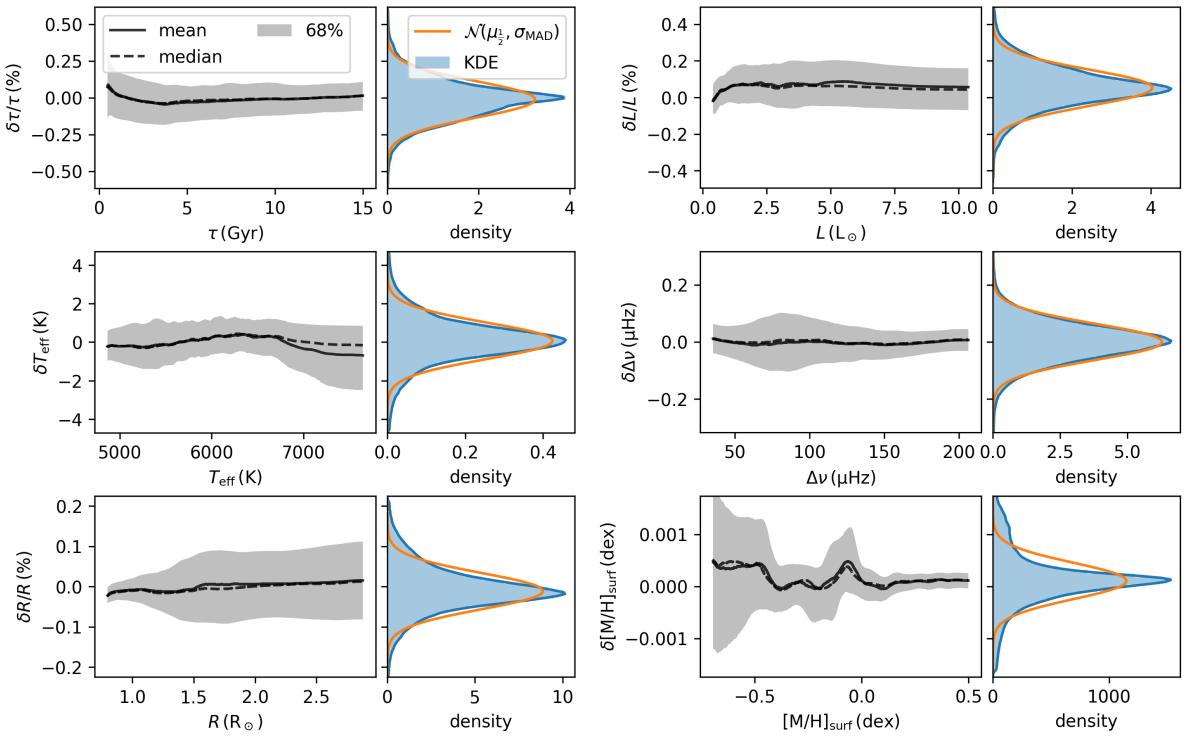


Figure A.3: *Left*: the rolling error between a given parameter in the test dataset ($\mathbb{Y}$) and the ANN prediction for that parameter ($\widetilde{\mathbb{Y}}$) where $\delta\mathbb{Y} = \widetilde{\mathbb{Y}} - \mathbb{Y}$. *Right*: a kernel density estimate (KDE) of the error and a normal distribution centred on the median, $\mu_{1/2}$ with an estimator for the standard deviation from the median absolute deviation, σ_{MAD}.

the training and validation MAE as a function of epochs for the final ANN configuration. The training and validation loss were comparable throughout training.

A.1.3 Testing

The test dataset contained $\sim 2{\times}10^6$ stellar models evolved in the same way as the training dataset, but with initial conditions chosen randomly across the range of the grid. We made predictions for the test dataset, deriving luminosity from the output radius and effective temperature, using the final trained ANN as described in Appendix A.1.2. We then evaluated the accuracy of the ANN by taking the difference between the test truth and ANN prediction, $x_{\mathrm{true}} - x_{\mathrm{pred}}$.

We found good agreement between the test dataset and ANN predictions, within typical observational uncertainties. We noted that the largest errors lay at the boundaries of the training data and in areas sparsely populated by the grid. This is apparent in Fig. A.3 where we plot the test error against each parameter. For example, the spread in error increases for

Table A.2: The median error, $\mu_{1/2}$ and median absolute deviation of the error, $\sigma_{\rm MAD} = 1.4826 \cdot \text{median}(|E(\mathbb{Y}) - \mu_{1/2}|)$ for a given ANN output parameter, $\mathbb{Y}$ from the test dataset. The error, $E(\mathbb{Y})$, is given in the table header, where $\delta\mathbb{Y} = \widetilde{\mathbb{Y}} - \mathbb{Y}$.

Error	$\delta\tau/\tau$ (%)	$\delta T_{\rm eff}$ (K)	$\delta R/R$ (%)	$\delta L/L$ (%)	$\delta\Delta\nu$ (μHz)	$\delta[\text{M/H}]_{\rm surf}$ (dex)
$\mu_{1/2}$	-0.012	0.070	-0.011	0.053	0.00022	0.00010
$\sigma_{\rm MAD}$	0.123	0.941	0.045	0.099	0.06341	0.00035

$[\text{M/H}]_{\rm surf} < -0.5$ where training data is sparse at the edge of the grid. However, the accuracy is very good within the observed range covered by our sample of 81 dwarfs and subgiants. Hence, we chose the median absolute deviation (MAD) as an estimator of the spread in error because it is less sensitive to large errors at the grid boundary than the standard deviation.

To represent the accuracy of the ANN, we present the median, $\mu_{1/2}$ and MAD estimator, $\sigma_{\rm MAD} = 1.4826 \cdot \text{median}(|E(x) - \mu_{1/2}|)$ of the error ($E(x)$) in Table A.2. The median is close to zero for all parameters, showing little systematic bias in the ANN. The MAD is also lower than observational uncertainties quoted in Section 3.2. The spread in error for $\Delta\nu$ of 0.06 μHz is comparable to that of observations with the best signal-to-noise. However, the error in $\Delta\nu$ predictions is also comparable to other systematic uncertainties in $\Delta\nu$ discussed in Section 3.3.1. Therefore, a robust model which takes account of systematic uncertainties pertaining to $\Delta\nu$, including those from the ANN, will be explored in future work.

A.2 Prior Distributions

We chose a transformed beta distribution (see Equation 3.4) as the prior for the non-pooled stellar parameters as an alternative to a uniform distribution. Fig. A.4 shows the beta distribution compared with a uniform distribution for some parameter x from 0 to 1. We found that the continuously differentiable nature of the beta distribution was preferred by the NUTS over the uniform distribution.

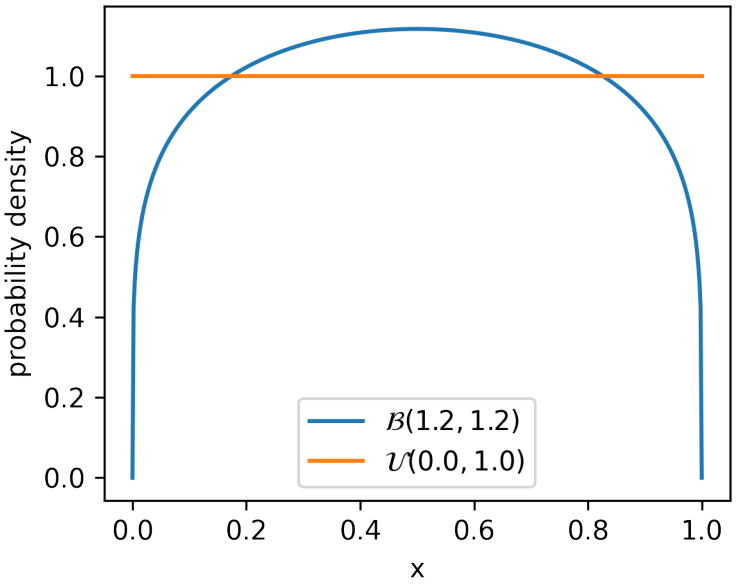


Figure A.4: A beta distribution ($\mathcal{B}$) with $\alpha = \beta = 1.2$ for some parameter x and a uniform distribution ($\mathcal{U}$) from 0 to 1.

A.3 The Synthetic Population

In this section, we present the results for the NP, PP, and MP models run on a synthetic sample of 100 stars with the following initial conditions. We randomly generated initial M and $[\mathrm{M/H}]_{\mathrm{init}}$ uniformly. We drew initial values for Y_{init} from a normal distribution centred on the helium enrichment law from Equation 3.10 with $\Delta Y/\Delta Z = 1.8$ and $Y_P = 0.247$, and scaled by $\sigma_Y = 0.008$. We also generated initial values for α_{MLT} from a normal distribution centred on $\mu_\alpha = 2.0$ and scaled by $\sigma_\alpha = 0.08$.

We evolved the synthetic stars to randomly chosen ages using MESA. We then took the output τ, T_{eff}, L, $\Delta\nu$, and $[\mathrm{M/H}]_{\mathrm{surf}}$ from the models and used these as true values for each of the stars. We added random noise to the observed quantities centred on the true values with a standard deviation of 2.2 per cent in T_{eff}, 3.5 per cent in L, 0.9 μHz in $\Delta\nu$, and 0.07 dex in $[\mathrm{M/H}]_{\mathrm{surf}}$ chosen to be representative of the APOKASC sample.

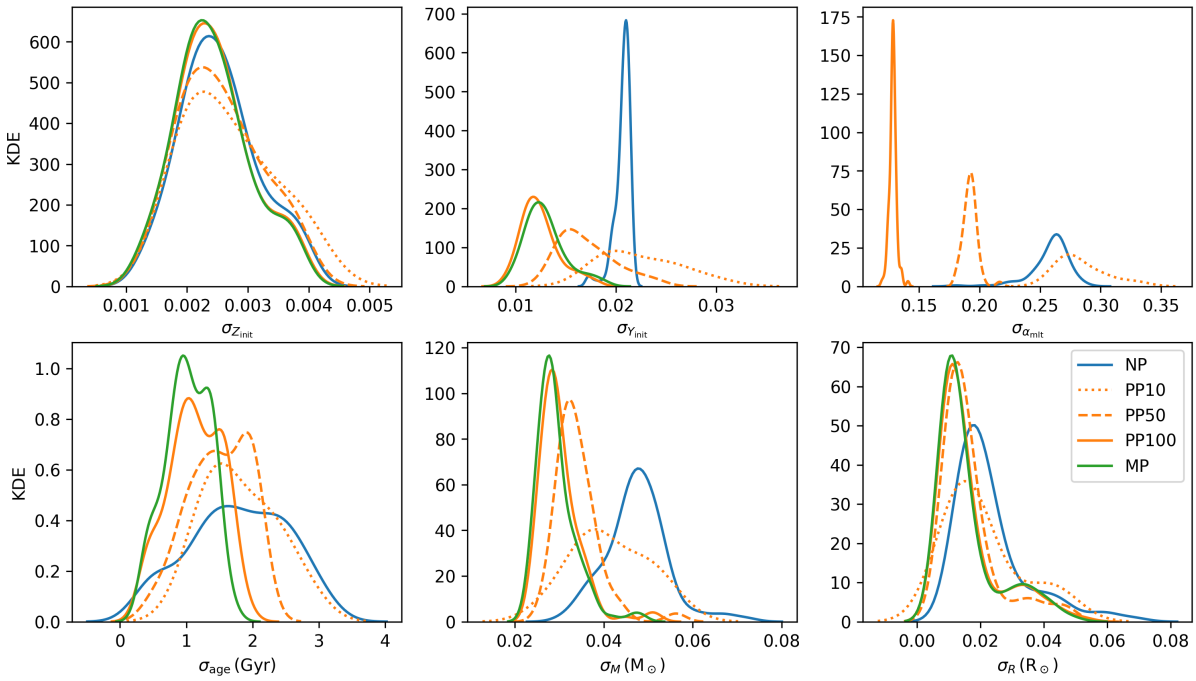


Figure A.5: Kernel density estimates (KDEs) of the distributions of statistical uncertainties from each model for the sample of synthetic stars. The PP model was run with 10, 50 and 100 stars and is denoted PP10, PP50, and PP100 respectively. The NP and MP models were both run with the full set of 100 stars.

A.3.1 Stellar Parameters

We found that the NP model recovered the true values for the individual stellar parameters, but the uncertainties were unreliable. The observational quantities alone were not good enough to constrain $Y_{\rm init}$ and $\alpha_{\rm MLT}$. As a result, their distributions were truncated at the bounds of their priors. These boundary effects skewed the marginalised posterior means for $Y_{\rm init}$ and $\alpha_{\rm MLT}$ towards the centre of the prior (0.28 and 2.0 respectively).

The PP model recovered true values for the synthetic stars with more reliable uncertainty than the NP model. The addition of pooling $Y_{\rm init}$ and $\alpha_{\rm MLT}$ between the stars improved their uncertainty which reduced the effects of the prior as seen in the NP model. We found little difference between the results of the PP and MP models.

We reran the PP model with 10 and 50 stars chosen randomly from the sample of synthetic stars. In Fig. A.5, we show the uncertainties in the several parameters from the results of each of the models. For the two pooled parameters, $Y_{\rm init}$ and $\alpha_{\rm MLT}$, the uncertainty reduction due to

pooling is most obvious. We see the PP model repeatedly improves on the uncertainties from the NP model when $N_{\rm stars}$ is increased.

In Fig. A.5 we also see a similar reduction in uncertainty for τ, M, and R, with all models improve upon the NP model. However, we do not see the same effect in $Z_{\rm init}$ for which the uncertainty appears dominated by observations of $[{\rm M/H}]_{\rm surf}$.

A.3.2 Population Parameters

In Fig. A.6, we see that the PP model also recovers the hyperparameter truths well, with some noise due to random realisation error. Fitting the model this way has the added benefit over the NP model of improving the inference of the individual stellar parameters, as shown in the previous two sections. We also found that when we ran the PP model with 10 and 50 stars, the uncertainties on the hyperparameters also shrank with increasing $N_{\rm stars}$.

Fig. A.7 shows the hyperparameter results for the MP model. Here, $\alpha_{\rm MLT}$ was assumed to be the same for all stars. This model also recovers the true hyperparameters for helium well, and the assumed value for $\alpha_{\rm MLT}$ is within uncertainty of the true μ_α.

A.4 The Sun as a Star

Our model consistently recovers the Sun when modelled in each of the NP, PP, and MP models. In Table A.3 we present the results for the Sun as a star from the NP model to show what we obtain without the influence of any other stars in the sample. We show the marginal and joint posterior distributions for the solar parameters from the NP model in the corner plot in Fig. A.8.

We found some differences between $\alpha_{\rm MLT}$ from our solar model and solar calibrations in the literature produced using MESA with similar input physics. For example, the solar calibration in Stancliffe et al. (2016) using Asplund et al. (2009) abundances yields compatible initial abundances, $Z_{\rm init} = 0.0149$ and $Y_{\rm init} = 0.266$ but $\alpha_{\rm MLT} = 1.783$ which differs from our results by about 10-σ. This is likely because of a few differences in observed values used for the calibration. Stancliffe et al. (2016) used observed helium abundance and convection zone

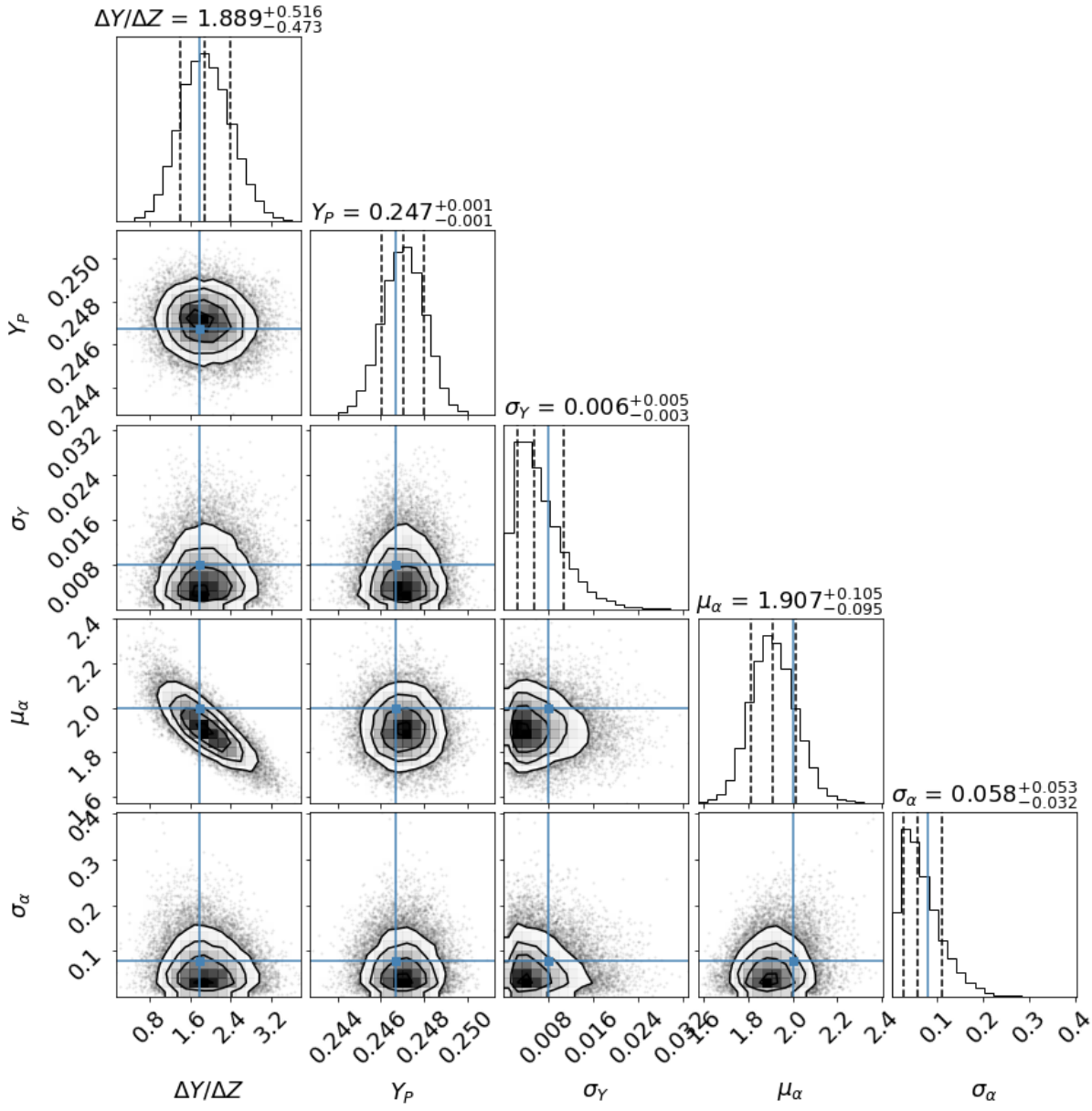


Figure A.6: Corner plot showing the marginalised and joint posterior distributions between the PP model hyperparameters for 100 synthetic stars. The true values are shown by the blue lines.

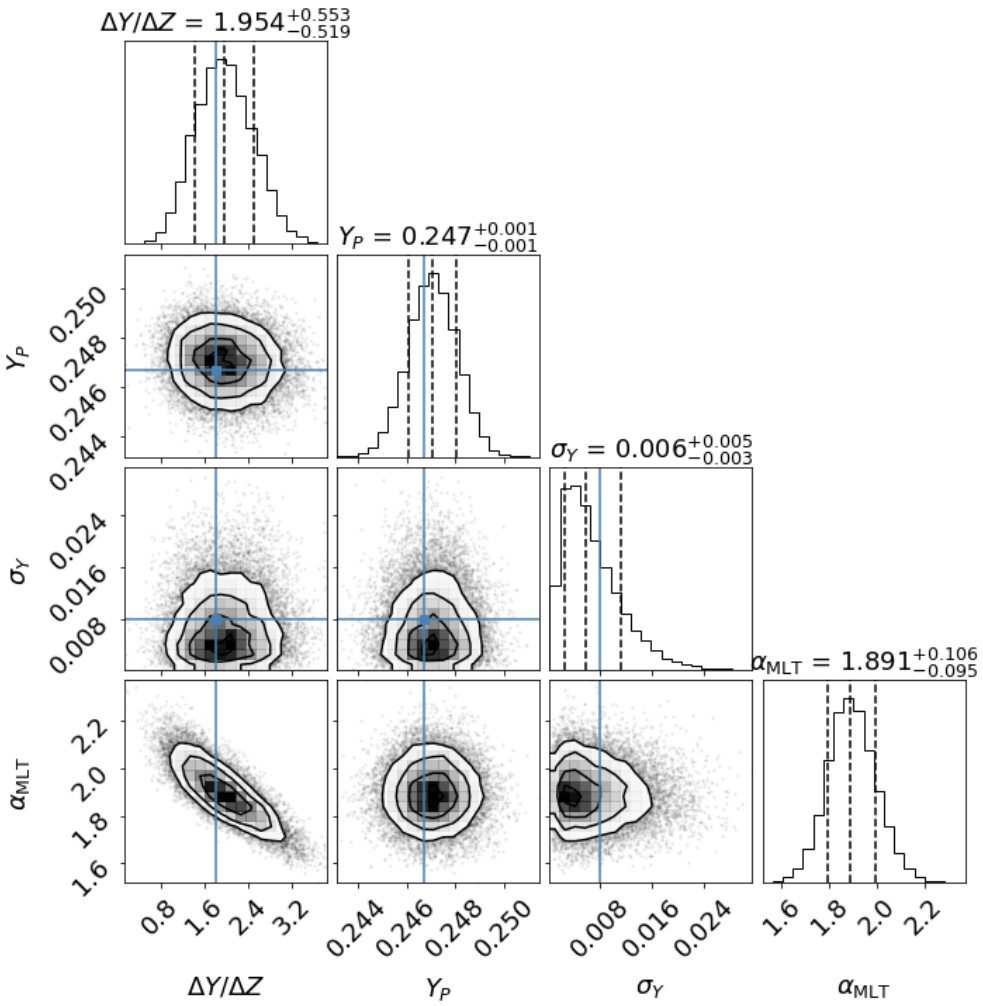


Figure A.7: The same as Fig. A.6 but for the MP model.

Table A.3: Solar results from the NP model. The second column shows the median marginalised posterior samples for each parameter with their respective upper and lower 68 per cent credible intervals.

Input parameters.

$f_{\rm evol}$	M ($\mathrm{M}_\odot$)	$\alpha_{\rm MLT}$	$Y_{\rm init}$	$Z_{\rm init}$
$0.517^{+0.009}_{-0.008}$	$1.000^{+0.001}_{-0.001}$	$2.12^{+0.03}_{-0.03}$	$0.262^{+0.002}_{-0.002}$	$0.0150^{+0.0003}_{-0.0003}$

Output parameters.

τ (Gyr)	$T_{\rm eff}$ (K)	R ($\mathrm{R}_\odot$)	$\Delta\nu$ (μHz)	$[\mathrm{M/H}]_{\rm surf}$ (dex)
$4.6^{+0.1}_{-0.1}$	5777^{+12}_{-12}	$1.001^{+0.001}_{-0.001}$	$135.37^{+0.14}_{-0.14}$	$0.00^{+0.01}_{-0.01}$

depth measurements from helioseismology. Furthermore, solar calibrations typically include convective envelope overshooting. Presuming overshooting increases mixing in the star, we might expect a lower $\alpha_{\rm MLT}$ to compensate this. Therefore, we stress that the addition of the Sun as a star in our model is with the assumption of our choice of input physics.

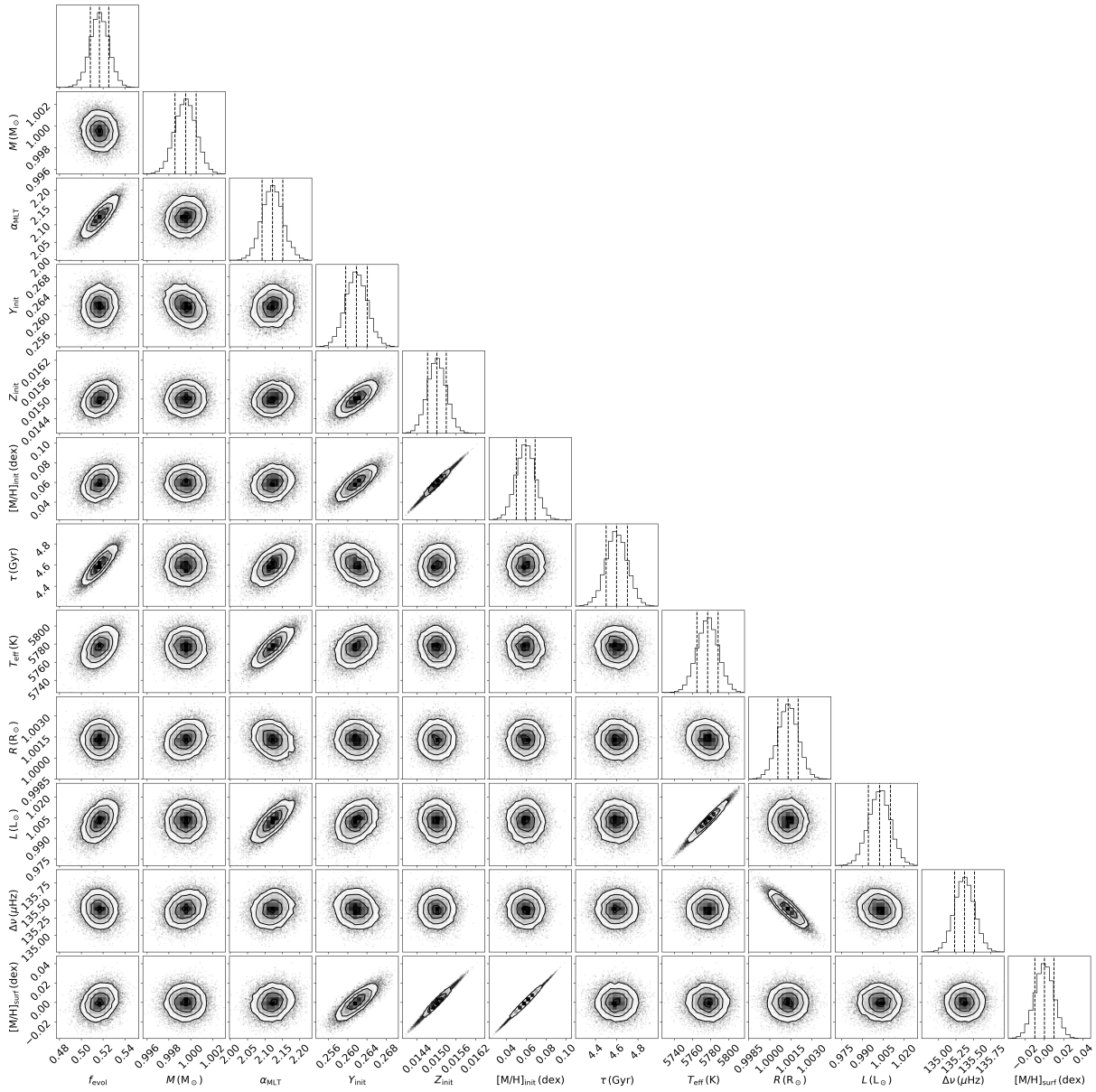


Figure A.8: A corner plot showing the sampled marginal and joint posterior distributions for the Sun as a part of the NP model.