Asteroseismic inference on the fundamental properties and magnetic activity of solar-like oscillating stars

by

Alexandra Elizabeth Louise Dixon

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

University of Birmingham

Birmingham, B15 2TT

March 2021

Abstract

Asteroseismology, the study of resonant oscillations of stars, is a vital tool for investigating interior stellar processes and accurately characterising stars. My thesis focuses on using asteroseismology of solar-like oscillators and advanced statistical methods to obtain fundamental properties of stars, including their magnetic activity distribution. I begin by setting up the theory and background of asteroseismology and stellar activity. I then follow with three applications all with the common theme of implementing Bayesian statistical inference.

The first study presents our newly developed method to estimate the latitudinal distribution of near-surface stellar magnetic activity. This makes use of the relative shifts of different asteroseismic frequencies which are sensitive to the level of activity in certain regions on the stellar surface. We were able to reproduce the known active latitude distribution on the Sun from helioseismic data and after applying our methods to data on the solar analogue HD 173701 as observed by the NASA *Kepler* Mission we find that its active bands extend across a wider range in latitude than those on the Sun. The second study builds on the concept of activity-induced frequency shifts to investigate the impact this has on predictions of fundamental stellar properties from stellar modelling pipelines. We find that although for most properties this bias should be small, with less than 0.5% in mass, age estimates can have up to a 5% error. We expect this to increase for stars with the highest and lowest inclination angles and with even stronger magnetic field strengths than those simulated in the study. The final chapter presents the use of machine learning

to build a prior which can be used when fitting a background model to a power spectrum in order to directly extract bulk stellar properties. Our method is able to distinguish between red clump and red giant branch stars to predict mass, effective temperature, and global asteroseismic parameters for targets observed by *Kepler* that are consistent with those of the APOKASC-2 catalogue.

For my wonderful husband,
my parents and my brother,
to whom I owe my successes.

Thanks and acknowledgements

Firstly, I would like to say a huge thank you to my supervisor Bill Chaplin. He has not only given me masses of guidance both professionally and personally, fantastic supervision and advice throughout my PhD, but he was the one who inspired me to pursue this magic thing called asteroseismology during my undergraduate degree.

Next, to Guy Davies my secondary supervisor who took me under his wing when he really didn't have to. Thank you for teaching me the right ways to do science (and most of the skills which got me employed!) and for being a brilliant mentor. Your advice has proved invaluable and I am always grateful for the time you have given me.

A special thanks is also due to Martin Nielsen, not only for being a wonderful friend but for your constant and limitless help which has been vital to me during my PhD.

I would like to extend a huge thank you to everyone in the Sun, Stars and Exoplanets group at the University of Birmingham, and particularly to Oli (sorry you had to cope with most of my worry rants), Ben, Mat, Eddie (my PhD twin), Vedad, Saniya, Matt, Walter, Martin, Warrick, Rachel, Emma, Alex, George, Lindsey and Lou. You have made my life at Birmingham hilarious and entertaining and so, so much fun! I have loved the drinks, games nights, stress unloads, the fights about vegetables, tea bag tossing and all of your silliness. I will miss you all very much but I know that I'll stay in touch and that you will all be well looked after by our mother hen Lou.

Of course, I would like to acknowledge the funding support of the Science Technology and Facilities Council (STFC) and the Stellar Astrophysics Centre (SAC) at the University of Aarhus which enabled my PhD. They also gave me the opportunity to attend multiple conferences over the 4 years which were not only very rewarding but a lot of fun! My gratitude also extends to Sarbani, Rachel and Angela for their collaboration during my studies.

To my Dad, Mum and Brother, thank you for always encouraging my passions, being unfailingly supportive and excited for my every achievement. I am incredibly lucky to have such a loving family and would not have succeeded without a lifetime of support from you all. I am a bit sorry though Dad that I won't be Dr Thomas.

Finally, thank you always to my best friend and partner in life, my husband Eliot. You have been my backbone for every highlight and trial of my PhD and it is clear to me and to everyone that I would not be in this position without you. I love you so much.

Contents

List of Figures

List of Tables

Abbreviations

AIA	the Atmospheric Imaging Assembly.
AIMS	Asteroseismic Inference on a Massive Scale.
	.
BiSON	the Birmingham Solar Oscillations Network.
BNN	Bayesian Neural Network.
CME	coronal mass ejection.
CNN	Convolutional Neural Network.
CoRoT	Convection, Rotation and planetary Transits.
EUV	extreme ultraviolet.
FFI	full-frame images.
GOLF	Global Oscillations at Low Frequencies.
GONG	the Global Oscillation Network Group.
HMC	Hamilton Monte Carlo.
HMI	the Helioseismic and Magnetic Imager.
HR	Hertzsprung-Russell.
JWST	James Webb Space Telescope.
LKJ	Lewandowski-Kurowicka-Joe.
LOI	low-resolution imager.
MAST	Mikulski Archive for Space Telescopes.
MCMC	Markov Chain Monte Carlo.
MDI	Michelson Doppler Imager.
MVN	multivariate normal.
NSO	the National Solar Observatory.
NUTS	No-U-Turn Sampler.
PLATO	PLAnetary Transits and Oscillations of stars.

RC	red clump.
RGB	red giant branch.
rms	root mean square.
SDO	the Solar Dynamics Observatory.
SOHO	Solar and Heliospheric Observatory.
SONG	Stellar Observations Network Group.
SPM	sunphotometer.
SSM	Standard Solar Model.
TESS	Transiting Exoplanet Survey Satellite.
VIRGO	Variability of solar IRradiance and Gravity Oscillations.
YGM	Yale Grid-based Modelling.
ZDI	Zeeman-Doppler imaging.

1 Introduction and thesis outline

Our closest star, the Sun, has an enormous impact on our everyday lives being the fundamental source of most of the energy we use on Earth. To properly understand how the Sun interacts with its environment, including Earth, we must develop our theories of how it formed and is evolving. Stars are incredibly complex systems, involving processes we are only just beginning to grasp and which evolve over time. For example, solar magnetic activity, one of the main themes of this thesis, can have severe impacts on our environment but how it will affect the Earth depends on the magnetic nature of the Sun which is highly variable. With more detailed models we can make more reliable predictions for how the Sun will behave in the future.

However, we can only study the Sun in its current state so in order to fill in the gaps we must study other stars at different evolutionary stages and in different environments - to place the Sun in context. Only by exploring the characteristics and behaviours of other stars can we learn what the Sun alone cannot teach us.

1.1 Stellar astrophysics

Being so far away much about the inner workings of stars remains a mystery but one that astrophysicists are beginning to dissect. Even at a distance we can learn so much from the light that stars emit. By using telescopes which observe at different wavelengths we can examine the compositions and surface conditions of stars. Spectroscopic and astrometric measures provide kinematics for how stars move across the sky with respect to each other and in certain regions of the Milky Way. This helps to build a detailed picture of the formation and evolution of stars and their distribution

within the Galaxy which has a widespread impact on other fields in astrophysics. Observations of brightness changes over time can reveal many different situations, from energetic stellar activity events such as flares, through to binary systems, star spots traversing the stellar disc, orbiting exoplanets and other stellar companions. Most impressively, oscillations that are intrinsic to a star tell us directly about processes and conditions in their interiors. The study of these oscillations, known as *asteroseismology*, is the topic of this thesis.

Asteroseismology is a vital tool used to reveal details about the interiors of stars solely from observations of their stellar surfaces. The manner in which they oscillate is particular to every star but by analysing asteroseismic data on many stars we can refine our knowledge of the physical processes involved and build models of stars. Links can then be made to the Sun by studying stars which oscillate by similar mechanisms. Fortunately these are abundant in the Galaxy.

The contribution of asteroseismology to stellar physics has been widespread. For example, since they can be used to inspect internal stellar conditions, asteroseismic methods can distinguish between red giant branch stars which burn hydrogen in a shell around an inert helium core and red clump stars which are core helium burning. These stars are otherwise hard to distinguish based on their surface properties alone (e.g. Bedding et al., 2011; Mosser et al., 2012b, 2015; Elsworth et al., 2017). Asteroseismic analysis produces precise stellar ages (e.g. Miglio et al., 2013; Lebreton et al., 2014; Meibom et al., 2015; Davies et al., 2015; Silva Aguirre et al., 2015) and largely model-independent masses and radii for thousands of main sequence stars (e.g. Chaplin et al., 2011a; Mathur et al., 2012; Huber et al., 2017; Johnson et al., 2017) and red giants (e.g. Hekker et al., 2011a; Pinsonneault et al., 2014, 2018; Yu et al., 2018). Population studies of stars from across the Milky Way can provide insight into the fundamentals of how the Galaxy was formed and will evolve, a field known as 'Galactic archaeology' (e.g. Miglio et al., 2013; Mathur et al., 2016; Stello et al., 2015). What's more, there is a strong synergy between asteroseismology

and exoplanet studies. By obtaining accurate properties for host stars the characterisation of orbiting exoplanets can be improved or in the case of some detection methods (e.g. the transit or Doppler methods) relies on external measures of stellar radius or mass (e.g. Christensen-Dalsgaard et al., 2010a; Huber et al., 2013; Silva Aguirre et al., 2015; Van Eylen et al., 2014, 2015, 2018). One topic of this thesis is on how asteroseismology can contribute to developing our knowledge of magnetic activity. Chapter 3 discusses the signatures of activity present in asteroseismic data and provides an overview of past and ongoing studies using these concepts.

1.2 Bayesian methods for Big Data Astronomy

With modern astrophysics we are making more and more observations and as such we are entering an era of 'big data astronomy'. Missions such as *Kepler* (Borucki et al., 2010), the Transiting Exoplanet Survey Satellite (TESS) (Ricker et al., 2015), and Gaia (Gaia Collaboration, 2016, 2018), and catalogues like 2MASS (Skrutskie et al., 2006), APOGEE (Majewski et al., 2017) and LAMOST (Deng et al., 2012) are providing us with enormous quantities of data. Therefore we need better statistical techniques to fully utilise the information available to us. We can do this with Bayesian statistics and machine learning.

Simply put, machine learning is the process of building a model by training it on previously known data, i.e. it learns through experience. From the trained model, predictions based on new measurements can be made with minimal human interaction; perfect for the current data influx.

Bayesian methodology is based on the concept of describing data using a probability distribution. Crucially, this distribution includes prior information based on what is already known about the system and these beliefs can be updated as new evidence becomes available. For example, consider constructing a model describing star spots traversing a visible stellar disc in order to acquire the rotational period of the star and fitting this to a lightcurve. We can include a probability distribution

to restrict the rotation rate parameter from unphysical values, such as being faster than the break-up velocity of the star. This is known as the 'prior' distribution which will have higher probability at what we believe to be more likely rotation rates and zero for unreasonable values. Within a Bayesian framework we can construct a probability distribution describing the most likely rotation rate given the data whilst incorporating this prior belief. If any new information arises, for example additional measurements or observations using a different instrument, the prior can be updated to include this and the model refitted to refine the rotation rate prediction.

As in the example above, typically in astrophysics (and elsewhere in science) we want to fit a model to our data in order to have a good description of the system we are observing. To incorporate Bayesian statistics we define the posterior probability distribution of the model parameters with *Bayes theorem* (Bayes & Price, 1763) as being proportional to:

$$p(\Theta|\hat{d}) \propto p(\Theta)\, p(\hat{d}|\Theta), \tag{1.1}$$

where Θ are the model parameters and $\hat{d}$ are the observations. A lower case p indicates a probability distribution function. In the above $p(\Theta|\hat{d})$ is the probability of the parameters being true given our data and is called the 'posterior'. Often, this is the target distribution we want to obtain since it is a description of system underlying the data. $p(\Theta)$ is the 'prior', to incorporate any relevant prior belief we have on the model parameters. The remaining term $p(\hat{d}|\Theta)$ is the probability to obtain our observations given the choice of model parameters, known as the 'likelihood'. Using this equation, the shape of a joint posterior can be built over all parameters for the model based on the available data. This equation is used throughout my projects and will be called upon later.

Evaluating absolute values for this posterior requires normalising the relation in Equation 1.1 using a normalisation factor called the 'evidence', $p(\hat{d}) = \int p(\Theta)p(\hat{d}|\Theta)d\Theta$. This involves integrating, or 'marginalising', over all possible values of model pa-

rameters, however calculating the evidence can very rapidly become intractable. In order to obtain the posterior in a computationally efficient way techniques have been developed to estimate the target posterior distribution. These include sampling approaches, such as Markov Chain Monte Carlo (MCMC) or nested sampling, and approximation methods like variational inference. A full explanation of these methods is beyond the scope of this thesis but I give a brief description below and direct the reader to reviews such as Mackay (2003) or Gelman et al. (2013) for more details.

In MCMC representative samples are drawn from the prior by proposing a pseudo-random walk through parameter space. For each successive step in the walk the posterior probability is calculated before proposing a new position a small distance away in parameter space. Through this process Markov Chains are built and steps are continually taken until the chains reach a stationary distribution, one of the core requirements of MCMC. Often more than one chain for each parameter is run simultaneously. Over many iterations more samples are drawn from the regions with higher posterior probability density and as such the target posterior distribution is mapped out.

One of the most important steps when using MCMC is to establish that convergence has been achieved such that the posterior has reached a stationary state and one which is independent of the starting position. To do this we can visually inspect the chains of parameters making sure they match an expected distribution (e.g. Gaussian) and disregarding the first part of the chain which may not be independent, known as the 'burn-in'. The Gelman-Rubin statistic, $\hat{R}$, is another diagnostic of convergence which measures the variance within chains in comparison to the variance between chains (Gelman & Rubin, 1992). Only if $\hat{R} \approx 1$, has convergence been reached. Typically the tolerance on this quantity is ± 0.02.

Although MCMC is effective at low dimensions ($\lesssim 100$), for some problems, such as those with more free parameters, multimodal posteriors or degeneracies, a nested

sampling approach can be more effective. This estimates the posterior by sampling model parameters from the prior and successively removing points based on which produces the worst likelihood. Discarded points are replaced by sampling from a reduced prior space until, after multiple iterations, the peaks of the posterior are reached. The removed points are ranked and used to form the estimated posterior.

Although not used in this thesis, for completeness I include a brief description of an alternative approach to using sampling methods which is to use approximation techniques. Variational inference methods approximate the posterior with another function from a family of distributions which are easier to calculate, e.g. a Gaussian. The best fitting solution is reached by minimising the difference between this estimate and the target posterior. Once fitted, the approximation is used as the posterior estimate from which summary statistics can be taken. Variational inference is usually able to find a solution more quickly than sampling methods, and is more effective for problems with a large number of dimensions. However, since it is an approximation it can produce biased results if the true posterior is not well represented by the chosen family of distributions.

Once the joint posterior has been estimated, the corresponding model parameters are retrieved by marginalising over other variables thereby producing individual posterior distributions. For each parameter an appropriate summary statistic, e.g. the mean or median, can be reported along with an uncertainty estimate, e.g. the standard deviation.

An advantage of Bayesian methods is that we obtain a full description of the system since we acquire a distribution, not just a maximum. This is advantageous particularly if the posterior has a complex distribution shape. In addition, through Bayesian implementation it is possible to marginalize over nuisance parameters and still take them into account in uncertainty propagation. Correlations between parameters can also be identified easily.

1.3 Thesis outline

In this thesis I present three studies which were all conducted using Bayesian techniques and the third involved the implementation of machine learning to train a model prior for stellar parameter estimation. I begin with an introduction to asteroseismology, providing the theoretical foundation necessary for the rest of the thesis. Following this, Chapter 3 provides the background context for studying magnetic activity of the Sun and other stars and describes the contributions of asteroseismology to this rapidly developing field. The thesis then moves to the analysis chapters, each of which details the methodology and results of the study along with a discussion on the related future outlook. My individual contributions are detailed at the beginning of each analysis chapter.

Chapter 4 presents our new method to constrain the distribution of near-surface stellar activity using asteroseismic measurements (Thomas et al. 2019, published in Monthly Notices of the Royal Astronomical Society, Volume 485, Issue 3, pages 3857-3868). Within a Bayesian framework, our model is fitted to observations of low-l frequency shifts over time and used to reproduce the expected band of active latitudes for the Sun. We also apply our technique to the solar-like star HD 173701.

In addition to being useful for studying magnetic fields and activity cycles, research in the field of stellar magnetic activity is also important to determine how oscillation parameters are affected when fitting to the acoustic modes of active stars. If we disregard any presence of magnetic field our inference may be biased. In Chapter 5 we build on principles from Chapter 4 to investigate the impacts of magnetic activity on stellar property estimates from asteroseismic data. We show that the bias on estimates depends on many factors, including the latitudinal activity distribution, strength of magnetic field and stellar inclination angle (Thomas et al. 2021, published in Monthly Notices of the Royal Astronomical Society, Volume 502, Issue 4, pages 5808-5820).

Finally, Chapter 6 presents our method which combines machine learning techniques (by making use of trends within existing populations) with a Bayesian approach to produce a statistically rigorous method of extracting information from power spectra with limited human interaction. Our method makes use of a prior built by training it on existing data to guide the fitting. We show that we can make predictions of bulk stellar properties for evolved stars by fitting a simple model to a global spectrum with solely the power spectrum as input. The product of this work is a tool that can be used as a step before implementing programs which identify modes in power spectra since these usually require global seismic parameters and stellar properties as input.

2 Asteroseismology

Developing our theories of the formation, behaviour and evolution of the Sun relies on studying other stars to place the Sun in context. Since these concepts depend on the physical processes ongoing within stars, techniques which can probe stellar interiors are crucial to advance our understanding of these complex systems. One such tool is asteroseismology which uses the idea that stars resonate in unique but identifiable ways, revealing the nature of their interiors.

2.1 Pulsating stars

Stars are known to vary in the amount of light they emit. This can be due to extrinsic effects, such as being in a multiple-star system, hosting exoplanets, or material obscuring line-of-sight observations. Investigating these variations reveals details about a star's environment. Brightness variations can also be intrinsic to stars as a result of internal processes causing their surfaces to physically expand and contract. These are known as pulsating variables and are spread across the Hertzsprung-Russell (HR) diagram as shown in Figure 2.1. Various classes of oscillating stars exist and their mechanism for pulsation depends on their size and internal characteristics. However, all pulsating variables require a continual mechanism by which fluctuations are excited in their interiors.

The focus of this thesis is on one particular class of pulsating star: 'solar-like oscillators', which pulsate in a similar manner to the Sun. Solar-like oscillators are those that are cool enough to have an outer convective region ($< 1.5 M_{\odot}$) where convective motion drives the transport of energy outwards. This includes main-

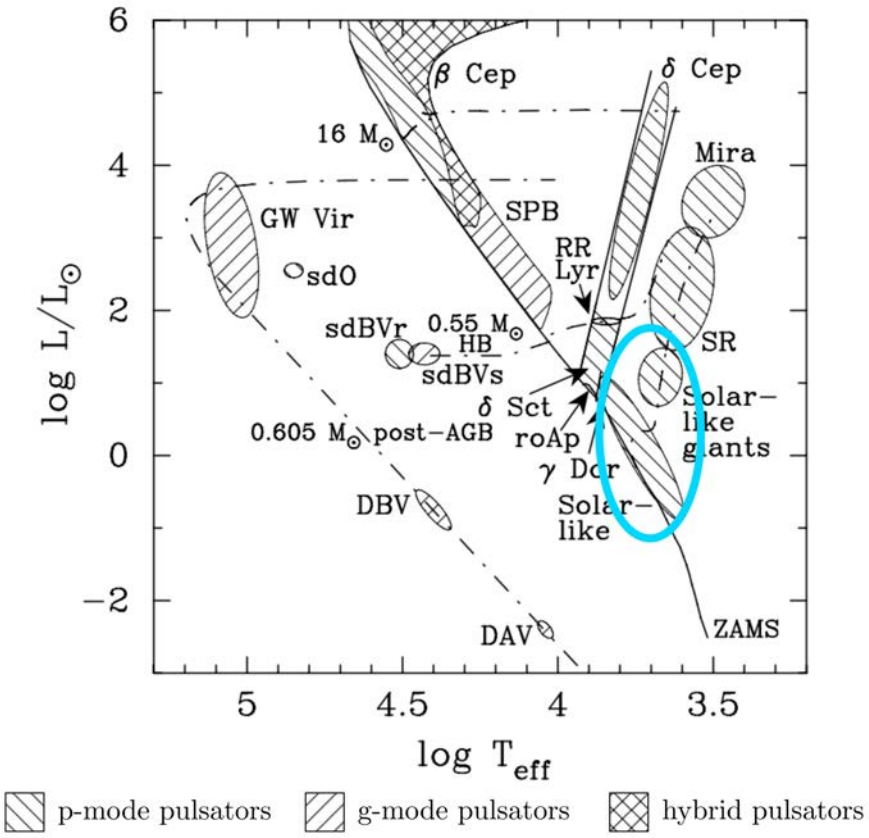


Figure 2.1: A HR diagram showing the regions occupied by different types of pulsating star. The position of solar-like oscillators, including main sequence and red giant stars, is shown by the blue ellipse. Hatched areas indicate different pulsation mechanisms (where 'hybrid' refers to modes with a mixed character). The solid line indicates the Zero-Age Main Sequence and dashed-dotted line shows parts of evolutionary tracks. Also shown is the instability strip. Figure adapted from Handler (2009).

sequence stars which are fuelled by hydrogen fusion in the core, sub-giants that have evolved to burn hydrogen in a shell surrounding a helium core, through to evolved red giant stars which are undergoing core helium fusion.

Solar-like pulsations are both stochastically driven and damped by turbulent motion within their near-surface convection layers (e.g. Goldreich & Keeley, 1977; Houdek et al., 1999). Their stochastic nature produces small amplitude oscillations that are short-lived (as opposed to those stellar pulsators that occupy the instability strip (e.g. Goldreich & Keeley, 1977)). Pulsations cause variations in brightness over time which we detect with ground-based or space telescopes (see Section 2.5).

2.2 Theory of solar-like oscillations

2.2.1 Excitation of modes

Within stellar interiors there is balance between the inward force of gravity and an outward pressure-gradient force. Oscillations arise when there is a perturbation about this equilibrium state and a restoring force acts on the stellar medium.

Once excited, these perturbations form waves which propagate within spherical cavities through a star. Each wave has a set of boundary conditions to satisfy and therefore occupies its own region within the star and propagates to a particular depth as indicated by dashed circular regions in Figure 2.2. When under the influence of a constant driving force, each type of oscillation self-interferes and forms standing waves within the stellar interior. The possible resonant frequencies depends on the physical properties of the cavity, for example its size or the speed of sound within it. Hence, the frequencies of these resonant oscillations will tell us about the properties of a star and reveal its inner workings.

There are two principal types of oscillation mode in solar-like pulsators and they occupy different cavities in main sequence stars due to their excitation mechanisms. In the outermost layers of near-surface convective regions turbulence generates acoustic waves producing 'pressure'/'acoustic'/'p modes'. Conversely, modes

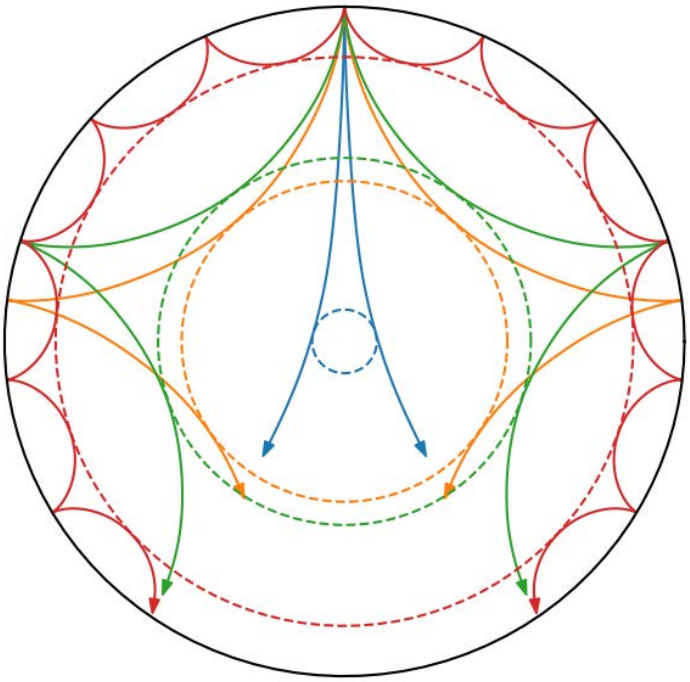


Figure 2.2: A cross section of a star showing the propagation of modes of different angular degree, l, within the stellar interior which are excited in the near-surface layers. The modes shown here are $l = 2, 20, 25, 75$ in blue, orange, green and red respectively. Modes of smaller l travel deeper into the star and have larger cavities. The dashed circles indicate the lower turning point of each mode. Image courtesy of Warrick Ball.

which rely on buoyancy as the restoring force are known as 'gravity'/'g modes'. These waves are confined to the radiative interior, becoming evanescent in the convective regions. In the near-surface layers of a solar-type star, the amplitude of g modes is very small making them notoriously difficult to observe (e.g. Turck-Chièze et al., 2004; Appourchaux et al., 2010). There have been some reports of solar observations of dipolar g modes using the Global Oscillations at Low Frequencies (GOLF) instrument on board the Solar and Heliospheric Observatory (SOHO) (García et al., 2007, 2008) and more recently of perturbations to p modes due to g modes in the Sun (Fossat et al., 2017), however these are still controversial (e.g. Schunker et al., 2018; Appourchaux & Corbard, 2019).

For evolved stars, where the p and g mode cavities have expanded and become

close enough for the waves to couple, it is possible to observe 'mixed modes' - those which show both g mode-like behaviour in core regions and act like p modes in outer layers. Such modes can be important for exploring the interiors of evolved stars since their g mode components probe stellar cores whilst their p mode properties allow them to have observable amplitudes at the surface (e.g. Beck et al., 2011; Bedding et al., 2011; Mosser et al., 2011). However, for the studies presented in this thesis only pure acoustic modes were used.

2.2.2 Characterizing modes

Since waves propagate in 3D stellar interiors we characterize them in terms of spherical coordinates: the radius, r, colatitude, θ, and longitude, ϕ. To describe the angular dependence of solar-like oscillations we use spherical harmonics, $Y_l^m(\theta, \phi)$ (e.g. Christensen-Dalsgaard, 2014) which originate from solving the Laplace equation in spherical coordinates. Modes are typically labelled using three integer numbers: the radial order, n, the angular degree, l, and the azimuthal order, m. Figure 2.3 shows the projection of oscillations on a sphere for different combinations of l and m.

The angular degree denotes the number of nodes on the spherical surface, where $l = 0$ are radial modes (where the entire star expands and contracts as a sphere), $l = 1$ are known as dipole modes, $l = 2$ as quadrupole modes, and so on. Angular degree continues in integer numbers where larger l correspond to oscillations which occur on smaller spatial scales on the surface and propagate to shallower depths. This can be seen in Figure 2.2 where lower angular degree modes have deeper cavities and fewer nodes on the sphere's surface. The azimuthal order indicates the number of nodal lines that intersect the equator. The number of available azimuthal orders for a particular angular degree is equal to $2l + 1$, where $-l \leq m \leq l$ and m can equal zero. Finally, the radial order describes the number of nodes in the radial direction as shells of material oscillate within the star.

By making simplifying assumptions, such as spherical symmetry and that the

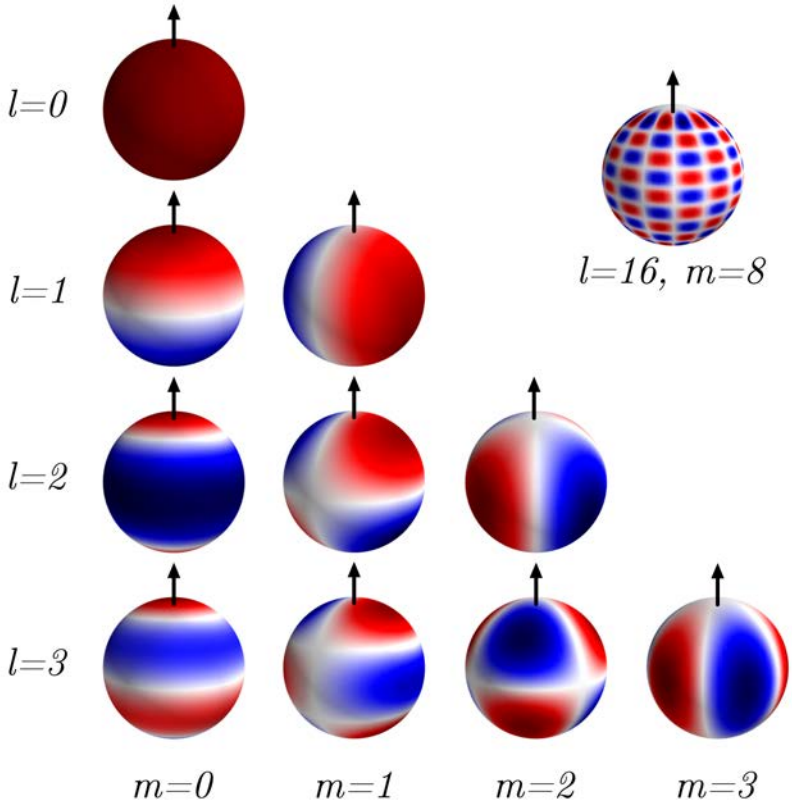


Figure 2.3: Schematic of spherical harmonics projected onto a sphere for different combinations of l and m. The strong red or blue colour indicates the maximum or minimum extent of oscillation. Shown here are the $l \leq 3$ modes which can be observed with *Kepler*, along with an example of higher degree and smaller spatial scale oscillations. Images courtesy of Warrick Ball.

perturbation to the gravitational potential is negligible (the Cowling approximation (Cowling, 1941)), we can set up wave equations for oscillations within stars. The solutions are the eigenfrequencies and associated eigenmodes that describe propagation (Aerts et al., 2010) and they must obey boundary conditions.

At the upper turning point of the cavity, waves are reflected back inwards when their frequencies are smaller than a threshold frequency known as the 'acoustic cut-off frequency' as given by:

$$\nu_{\rm ac} = \frac{c_s}{4\pi H_\rho}, \tag{2.1}$$

where H_ρ is the density scale height. Near the surface of the Sun there is a sharp drop in density causing H_ρ to reduce and incurring a sudden rise in $\nu_{\rm ac}$. At this point, the period of the wave is longer than the characteristic timescales in that region and perturbations can no longer continue as outward travelling waves. They become evanescent beyond the upper boundary and are reflected inwards. For low-l modes the height of the upper turning point is largely independent of angular degree.

As waves travel towards the centre of the star, the temperature rises and therefore the sound speed increases with depth. Consequently, non-radial ($l \neq 0$) waves are refracted and forced to follow curved paths. The depth to which waves propagate, i.e. the lower turning point, depends on the angular degree of the mode. For the same radial order, lower-l modes travel more deeply into the stellar interior (as shown in Figure 2.2).

The inertia of a mode describes how difficult it is to excite and how susceptible it is to perturbations (e.g. Aerts et al., 2010). It is a function of the angular degree and frequency in that lower-l modes have higher inertia as do lower frequency modes of the same angular degree. Any perturbations to oscillations will affect lower inertia, and hence higher l and higher frequency, modes more significantly. We see this frequency dependence in shifts caused by the presence of magnetic activity (as discussed in Chapter 5) (Libbrecht & Woodard, 1990; Chaplin et al., 1998).

2.3 Characteristics of a frequency spectrum

2.3.1 Constructing a power spectrum

We observe solar-like oscillations through variations in stellar brightness or line-of-sight velocity over time. After data preprocessing to remove systematics (see Wohler & Jenkins 2016) the Fourier transform of the lightcurve is computed in order to decompose the oscillatory signal into its different frequency contributions. Although light emitted by stars is continuous, astronomical observations are made by taking many discrete flux measurements, so instead the *Discrete* Fourier transform (DFT) of the lightcurve is calculated. The result is a frequency-power spectrum (e.g. Figure 2.4) showing a rich collection of pulsations from which asteroseismic measurements are made. By interpreting these data we can infer stellar characteristics.

The noise in a lightcurve or timeseries is normally distributed in the time domain. When taking the DFT both the real and imaginary parts will be normally distributed with zero mean and a variance which depends on the level of noise in the time domain. To construct a power spectrum, each part is made up of the complex components of the DFT squared and then exponentiated. This generates an observed power spectrum that is distributed about the underlying limit spectrum with negative exponential statistics, hence the likelihood function follows a χ^2_k distribution where $k = 2$ is the number of degrees-of-freedom (Appourchaux et al., 1998). The χ^2_k distribution is a special case of the gamma distribution which has the general posterior density function:

$$p(x) = \frac{\beta^\alpha}{(\alpha - 1)!}\, x^{\alpha-1}\, e^{-\beta x}, \tag{2.2}$$

where $\alpha = \frac{k}{2}$ is a positive integer which acts as a shape parameter, β is a rate parameter, and x is the variable with the negative exponentially-distributed noise when $k = 2$. For un-binned power, P, this reduces with $\alpha = \frac{k}{2} = \frac{2}{2}$ such that the

noise properties are described by:

$$p(P) = \frac{1}{2}\, e^{-\frac{P}{2}}. \tag{2.3}$$

Whereas Gaussian noise is additive, noise following χ^2_2 statistics is multiplicative. Within a frequency bin the variance is equal to the underlying power squared, therefore higher amplitude signals in the power spectrum will have more scatter.

Often in asteroseismology data are rebinned in order to smooth the spectrum, for example by taking the mean power over several bins. However, the noise properties change with rebinning such that the number of degrees-of-freedom is $2s$ where s is the number of bins over which the average power is calculated, a.k.a. the rebinning factor. In general, to statistically describe the power distribution we use $\alpha = \frac{k}{2} = s$ and $\beta = \frac{s}{2 \times \mathcal{M}(\Theta,\nu)}$ where $\mathcal{M}$ is the modelled limit spectrum. Hence, the probability density that the observed power is equal to P_i in a particular frequency bin, ν_i, is given by

$$p(P_i|\Theta) = \frac{s^s}{(s-1)!}\, \frac{P_i^{s-1}}{\mathcal{M}(\Theta,\nu_i)^s}\, e^{-\frac{sP_i}{2\mathcal{M}(\Theta,\nu_i)}}, \tag{2.4}$$

(Anderson et al., 1990; Appourchaux, 2003). For the simple case with no rebinning (i.e. $s = 1$) this relation returns to the standard χ^2_2 distribution.

We can calculate the joint probability density function, or the *likelihood*, of the entire spectrum by taking the product of $p(P_i)$ over a range of frequency bins (up to a total of N bins) and assuming that bins are uncorrelated:

$$p(\hat{d}|\Theta) = \prod_{i=1}^{N} p(P_i|\Theta). \tag{2.5}$$

This equation can be used when forward-fitting model power spectra to asteroseismic observations. An alternative method, via the central limit theorem, is to assume that a rebinned spectrum (when using $s \gtrsim 30$) can be modelled as having

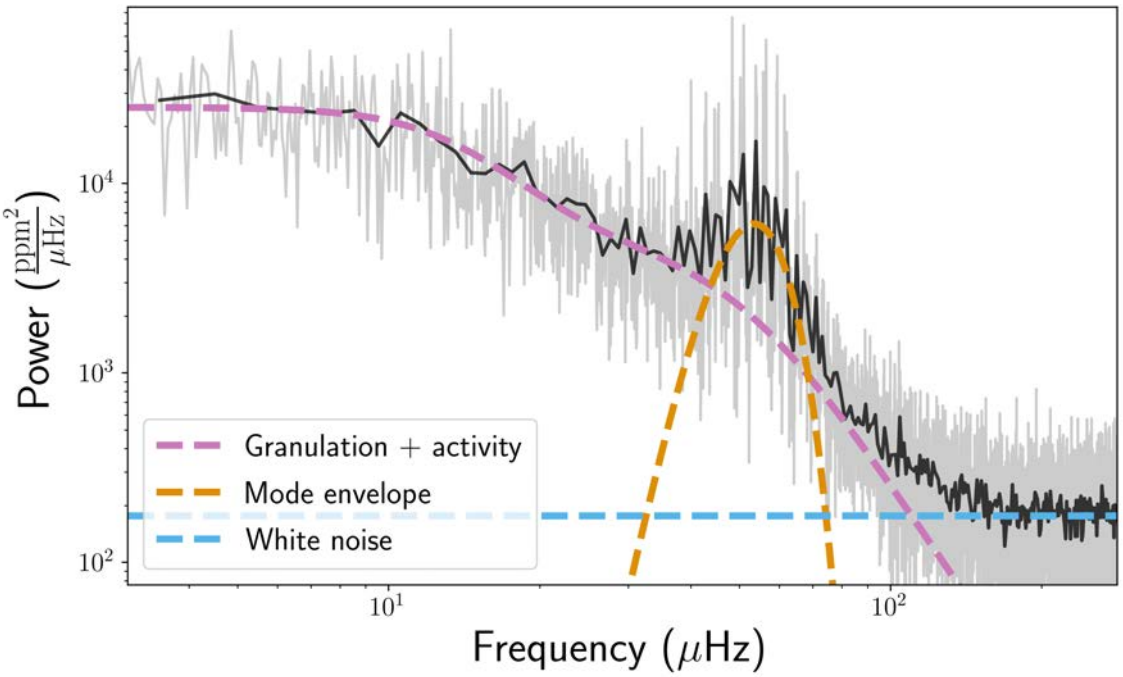
normally distributed noise (Appourchaux et al., 1998), however throughout this thesis we assume it to be gamma distributed. When choosing a limit spectrum model that is a good description of the data, the observed power divided by the limit model should give a flat residual spectrum with no frequency dependence.

2.3.2 Global spectrum properties

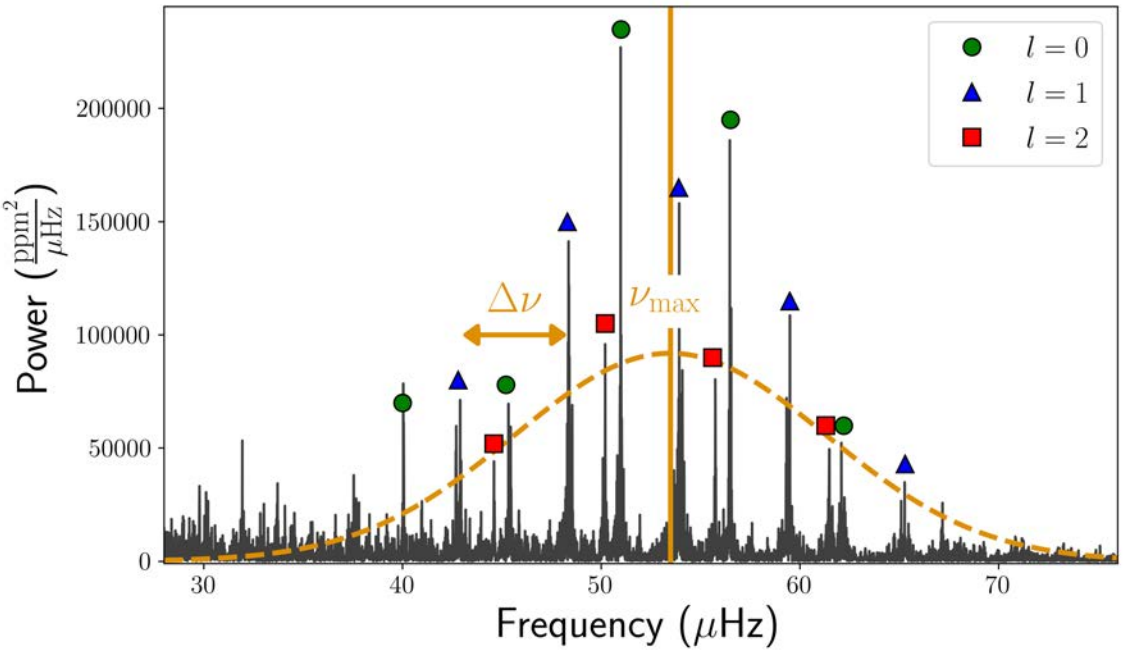
Within a power spectrum there are multiple sources of signal, besides the oscillations we are most interested in. Power due to granulation (from convective motion producing a surface pattern on the star) exists most prominently at low frequencies. Granulation has a wide range of evolving timescales which are dependent on stellar properties but typically the dominant characteristic timescale is on the order of a few to tens of minutes (e.g. Aigrain et al., 2004; Mathur et al., 2011b; Kallinger et al., 2014; Cranmer et al., 2014).

At even lower frequencies we see stochastic signatures of magnetic activity both from the growth and decay of star spots, and periodic variations reflecting the stellar rotation as spots traverse the visible disc. At similar frequencies we have long term instrumental noise contributions, for example from drifts or jitter. An additional instrumental effect from shot noise exists across all frequencies due to the discrete nature in which photons are collected. This noise follows Poisson statistics but at high enough photon collection rates it can be approximated as being normally distributed (i.e. Gaussian) in the time domain. Shot noise therefore appears as a flat offset in the power spectrum (or 'white noise'). Figure 2.4a shows the approximate regions of each source of signal for an example red giant star excluding activity signatures from rotational modulation (since this is a red giant star). Detectability of oscillations depends on how much greater mode power is than noise sources.

Most important to the asteroseismic community are the signatures of solar-like oscillations which appear as rich forest of modes at different frequencies in a power spectrum. The frequencies at which modes are excited to observable amplitudes



(a) An example power spectrum (for the red giant KIC6974745) where observations are shown by the faded grey line and a smoothed version of the data overlaid in black. Different components have been identified including the Gaussian envelope of modes, signatures of granulation at low frequencies, and the flat white noise background.



(b) A crop of the above to the modes of oscillation with no smoothing. The same Gaussian envelope is highlighted in orange. ν_{max} and $\Delta\nu$ have also been indicated. Modes corresponding to $l = 0, 1, 2$ are identified.

Figure 2.4: Example power spectra (for the red giant KIC6974745).

above the background noise offset are particular to the star. For main sequence stars and subgiants modes with high radial orders ($n \gtrsim 20$) have the highest amplitude since these stars resonate at high frequencies. Conversely, stars which have evolved onto the red giant branch have oscillations at lower frequencies and radial orders ($n \lesssim 20$).

Modes appear in a power spectrum at intermediate frequencies as a prominent hump of peaks which is roughly Gaussian-like in shape (e.g. see the dashed orange curve in Figure 2.4). The frequency at which there is maximum power, i.e. the peak of the mode envelope, is called $\nu_{\rm max}$. This is determined by which n are excited to observable amplitudes. The physical origins of $\nu_{\rm max}$ are not yet fully understood but it is found to scale with the acoustic cut-off frequency (Brown et al., 1991; Belkacem et al., 2011). This is expected since both $\nu_{\rm max}$ and $\nu_{\rm ac}$ are dependent on the near-surface layers of stars where modes are excited. Based on its connection with $\nu_{\rm ac}$, a relation for $\nu_{\rm max}$ is:

$$\nu_{\rm max} \propto g\, T_{\rm eff}^{-\frac{1}{2}}, \tag{2.6}$$

where g is the surface gravity of the star and $T_{\rm eff}$ is the effective temperature. Therefore $\nu_{\rm max}$ is a good diagnostic of the conditions in near-surface regions.

Another global asteroseismic variable is the separation between modes of the same angular degree but consecutive radial order known as the 'large frequency separation' and denoted as $\Delta\nu$. This quantity scales to good approximation as:

$$\Delta\nu \propto \sqrt{\langle\rho\rangle}, \tag{2.7}$$

where $\langle\rho\rangle$ is the mean density (Ulrich, 1986).

A relation between $\nu_{\rm max}$ and $\Delta\nu$ can be estimated as

$$\Delta\nu \propto \nu_{\rm max}^{b}, \tag{2.8}$$

where by combining the scaling relations 2.6 and 2.7 it is possible to show that $b = 0.75$ if one disregards the weak mass and effective temperature dependence. Observationally it has been determined that $b = 0.77$ by Stello et al. (2009a). $\Delta\nu$ can be measured from the oscillation patterns in a power spectrum, by taking the Fourier transform of the spectrum (Mathur et al., 2010), or from computing the autocorrelation function of the power spectrum (Mosser & Appourchaux, 2009).

As mentioned previously, as a star evolves the convective envelope expands and the acoustic cavities grow. Since surface gravity reduces, oscillation frequencies and hence $\nu_{\rm max}$ will move to lower frequencies, from over $10^3\mu$Hz for main sequence stars to around a few hundred down to a few μHz for red giants (following Equation 2.6). Similarly, as the average density reduces, so too does the large frequency separation. As a star evolves we begin to see mixed modes where oscillations take on both p and g mode-like behaviour. For more discussion on changes in a power spectrum with age see Chaplin & Miglio (2013).

The maximum frequency in a power spectrum at which one can unambiguously determine the frequency of a signal is known as the 'Nyquist', $\nu_{\rm Nyq}$ (Grenander, 1959). It is dependent on how regularly observations are taken such that a shorter cadence produces a higher Nyquist frequency. The *Kepler* long cadence of just under 30 minutes gives a $\nu_{\rm Nyq} \sim 283\mu$Hz, which is high enough to observe the oscillation modes of evolved stars. However, its short cadence of around 1 minute produces a $\nu_{\rm Nyq} \sim 8496\mu$Hz enabling asteroseismic measurements of main sequence stars. One must be aware of the Nyquist level since it can lead to aliasing effects as any power appearing at frequencies greater than the Nyquist frequency will be reflected back at $\nu_{\rm Nyq}$ adding onto the power already existing there. Multiple studies have investigated the impacts of Super-Nyquist sampling of seismic data including for *Kepler* (e.g. Murphy et al., 2013; Chaplin et al., 2014) and TESS (e.g. Murphy, 2015; Bell, 2020).

2.3.3 Detailed spectrum properties

As explained above, we can build a description of an oscillating star using a set of wave equations where the solutions are the eigenfrequencies at which the star pulsates. The detailed structure of modes of oscillation are defined by the asymptotic expression such that modes have frequencies given by:

$$\nu_{nl} \simeq (n + \frac{l}{2} + \epsilon)\,\Delta\nu - D\,l\,(l+1), \tag{2.9}$$

where ϵ is a phase factor which is dependent on boundary conditions, and

$$\Delta\nu = \left[2\int_0^R \frac{dr}{c_s}\right]^{-1}, \tag{2.10}$$

where the sound speed c_s is a function of the stellar radius r, and

$$D = \frac{\Delta\nu}{4\pi^2\,\nu_{nl}}\int_0^R \frac{dc_s}{dr}\,\frac{dr}{r}, \tag{2.11}$$

i.e. D depends on the sound speed gradient with respect to r (Vandakurov, 1967; Tassoul, 1980, 1990; Vorontsov, 1991; Lopes & Turck-Chieze, 1994). The integral is taken between the centre of the star to its radius, R. The first term on the right-hand side of Equation 2.9 is on the order of $\nu_{\max}$ ($10^3\mu$Hz for main sequence stars and a few hundred to a few μHz for red giants), and assumes that the star is uniform and non-rotating. The second term is on the order of a few μHz for main sequence stars or smaller for red giants and accounts for the non-uniform nature of stars since it is dependent on the sound speed gradient. The asymptotic expression implies that modes of certain degree and consecutive radial orders are near-uniformly spaced in frequency producing a comb of modes. Observationally however this does not quite hold and additional terms need to be added to the asymptotic expression, for example the radial order dependence of large frequency spacings (e.g. Mosser et al.,

2013) or the more complex pattern of mixed modes (e.g. Vrard et al., 2016).

From Equation 2.9 we can derive a useful quantity known as the 'small frequency separation' and calculated as the difference between modes of consecutive radial order and different angular degrees, i.e.

$$\delta\nu_{l,l+2}(n) = \nu_{n\,l} - \nu_{n-1\,l+2} \simeq (4l+6)D. \tag{2.12}$$

Two of the most common examples are $\delta\nu_{0,2}$ and $\delta\nu_{1,3}$. These quantities scale with the sound speed gradient which changes as a star evolves as hydrogen turns into helium. Therefore small separations are useful to examine the core conditions of stars and hence give an indication of stellar age for stars on the main sequence, particularly if the metallicity is known (e.g. Christensen-Dalsgaard, 1988; White et al., 2011). This is expected since small separations are calculated using two modes which propagate to very similar heights near the surface but travel to different depths in the deep stellar interior.

In addition to the small frequency separations another set of asteroseismic quantities often used are the frequency separation ratios. These are calculated by taking a ratio between small and large frequency separations (Roxburgh & Vorontsov, 2003). Separation ratios are useful quantities since they are almost independent of near-surface conditions as both numerator and denominator quantities have similar sensitivities to these regions. These are also the regions that are most poorly described by stellar models and cause the well known offset between observations and theory called the 'surface effect'. Stellar models approximate oscillations as linear and adiabatic, which is reasonable for most of the stellar interior. However, for near-surface layers, where the heat-transport timescales are of a similar length to oscillations, adiabaticity breaks down. Here the usual expression for modes is unable to reproduce observations and, if using frequencies rather than separation ratios, a 'surface term' is typically needed to counteract the offset. Chapter 5 discusses this

further and presents a study which makes use of separation ratios to infer stellar properties.

For a spherically symmetric star all modes of the same l and n lie at identical frequencies hence modes are degenerate in m. With the introduction of asphericities such as rotation or magnetic activity this degeneracy is lifted giving rise to frequency splittings (e.g. Gough & Thompson, 1990). For a single l mode the different azimuthal components, where $-l \leq m \leq l$, are separated out from all occupying the same frequency and form multiplets with peaks for each m. They are shifted in frequency by an amount that depends on their sensitivity and spatial response to the phenomenon.

One example of frequency splitting is the effect of stellar rotation on non-radial modes. For slowly rotating stars (well below break-up velocities) this can be treated as a perturbation to the symmetric case (e.g. Reese et al., 2006). The $m \neq 0$ components of a specific l mode are subject to an additional restoring force from rotation which pushes their oscillations to new frequencies, i.e. the components have a frequency

$$\nu_{nlm} = \nu_{nl} + \delta\nu_{nlm}, \tag{2.13}$$

where $\delta\nu_{nlm}$ is the rotational splitting term to offset from the non-rotating frequency ν_{nl}. The direction of this shift depends on whether the mode is travelling around the star prograde or retrograde to rotation. By treating the cavity that traps the mode as a solidly rotating region, the splitting can be expressed as

$$\delta\nu_{nlm} \approx \frac{1}{2\pi}\, m\Omega, \tag{2.14}$$

where Ω is the rotation rate of the cavity. Hence the frequency shifts are dependent on rotation rate allowing it to be measured using asteroseismology.

Moreover, since m denotes the number of node lines crossing the star's equator and therefore splitting the star up into sections, each azimuthal component is sen-

sitive to different regions of the star. This enables measures of rotation rate that varies with latitude known as 'differential rotation' which is an important quantity in the generation and evolution of solar and stellar magnetic fields (Gizon & Solanki, 2004). The responses of different m components can be used to study the spatial distribution of these acoustic asphericities.

We cannot directly spatially resolve the surfaces of stars other than the Sun since the light received from them will be integrated over the visible disc. Therefore, for high-l oscillations brightness increases in some regions will balance out brightness reductions in others and so there will be net geometric cancellation of high-l signal when observing stars as a single point. This would be the case for the $l = 16, m = 8$ mode oscillation shown in Figure 2.3. The integration across the stellar disc filters out these spatially small-scale variations so that observations are only sensitive to low l modes. It is for this reason that asteroseismology has so far involved the study of $l \leq 3$.

When it is possible to observe individual m components their relative visibilities within a particular l multiplet are known to depend on the stellar inclination angle, i, following the relation:

$$\varepsilon_{lm}(i) = \frac{(l - |m|)!}{(l + |m|)!} \left(P_l^{|m|}(\cos i) \right)^2 \tag{2.15}$$

where ε_{lm} is the mode visibility and $P_l^{|m|}(\cos i)$ the Legendre Polynomials (Gizon & Solanki, 2003). This relation assumes energy equipartition between the m components. By measuring the relative visibilities of modes the stellar inclination angle can be determined (e.g. Davies et al., 2015). This is a crucial quantity because an inclination angle dependence is often inherent in observations, e.g. measuring $v \sin i$ using spectroscopy, and it can also be a degenerate quantity in some measurements, e.g. estimating stellar rotation from lightcurves, or measuring exoplanet obliquities.

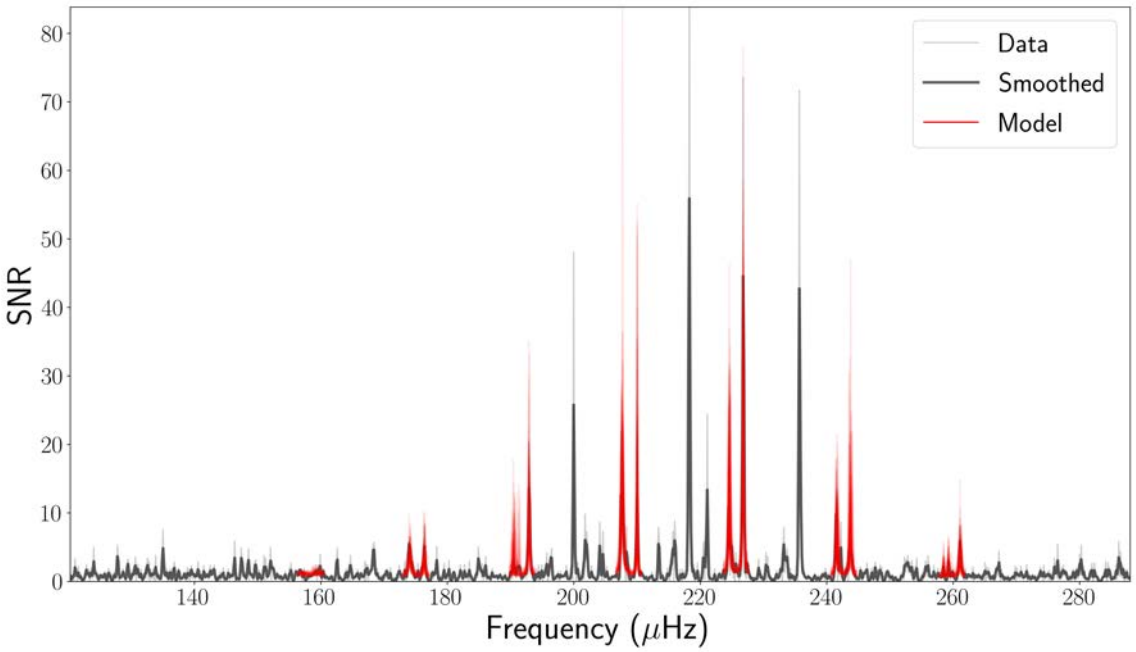


Figure 2.5: An example peak-bagged spectrum for the red giant star KIC4448777. Fitting was done using `PBJam` for just the $l = 0$ and $l = 2$ modes. In black is the observed data and model realisations from the fitted posterior are overlaid in red.

2.4 Stellar properties from the frequency spectrum

2.4.1 Peak-bagging modes

If there is sufficient frequency resolution in a power spectrum, individual modes of oscillation can be identified. In order to extract frequency measurements the modes must be fitted with a model, a process is known as 'peak-bagging'. Multi-parameter models describing the positions and profiles of modes are fitted in the frequency domain to observed oscillations and typically consist of a series of Lorentzian peaks, one for each resonant mode (Anderson et al., 1990). An example peak-bagged spectrum is shown in Figure 2.5. Peak-bagging analysis is used throughout asteroseismology as a means to extract frequency information from which stellar properties are then inferred.

The power spectrum model is a superposition of background trends and individual modes, typically only modelled for $l \leq 3$. Modes follow a repeating pattern of $l = 0$ and $l = 2$ modes in a close pair and $l = 1$ and $l = 3$ in a separate pair. Free

parameters of the model reveal information about a star's internal behaviour. They can include the underlying oscillation amplitudes and widths of each mode, which inform us about convective processes, as well as central frequencies of modes, which help to constrain the internal density and sound speed profile (for more detailed descriptions see reviews including Appourchaux 2014; García 2015; Basu & Chaplin 2017 and references therein).

With sufficient observational signal-to-noise more complex models can be fitted to include terms for frequency splittings due to rotation or magnetic activity. In these models, each l mode is a multiplet comprised of a Lorentzian peak for each m component with visibilities dependent on the inclination angle (Equation 2.15). In some cases it is possible to model $l=1$ mode as a triplet (for $m = -1, 0, 1$) however it is rare for data to be of good enough quality to see distinct m peaks making up a $l=2$ mode quintuplet. Instead they are typically modelled as a peak that is widened due to split m components. Additional complexities when peak bagging can include accounting for correlations between modes and granulation (e.g. Severino et al., 2001) or fitting to mixed modes in the spectra of evolved stars (e.g. Mosser et al., 2012b; Benomar et al., 2013; Mosser et al., 2014; Davies & Miglio, 2016).

Fitting models to the data can be difficult and typically requires good initial guesses to guide the procedure. Models must also be fit to the data using the correct statistics as detailed in Section 2.3.1. The process of fitting the multi-parameter model to the observed oscillation spectrum can be done using Maximum Likelihood Estimation (e.g. Anderson et al., 1990; Appourchaux et al., 1998; Stahn & Gizon, 2008) or Bayesian approaches which incorporate prior information (e.g. Appourchaux, 2008; Handberg & Campante, 2011; Appourchaux et al., 2012; Corsaro & De Ridder, 2014; Davies et al., 2016). The advantage of the latter is that the full posterior can be mapped out without the assumption that it is a multivariate Gaussian. However, it is a more computationally intensive procedure.

Fitting an entire complex model to data, including terms for several modes as

well as background signals, is challenging and time consuming. In order to save computational time, the number of free model parameters can be reduced by dividing up the power spectrum and fitting the modes a few radial orders at a time or more often pairwise. However, this can reduce the accuracy of the best fitting solution and introduce biases on model parameters since the new model will not including tails of power from neighbouring mode peaks. Caution is advised for this local approach to mode fitting particularly when including the effects of correlation between granulation background and mode excitation which give rise to an asymmetry in mode peaks (Jiménez-Reyes et al., 2008). Instead a balance can be achieved by using a 'pseudo-global' fitting approach where a model for the entire spectrum is implemented but free parameters are varied one at a time (e.g. Fletcher et al., 2009).

Two of the biggest challenges in peak-bagging are identifying how many modes are present in the power spectrum to determine the complexity of the model, and obtaining good initial guesses to guide the model fitting. These estimates are time consuming to obtain by human visual inspection and are subjective, particularly when signal-to-noise is very low. We already have large volumes of data from CoRoT, *Kepler* and more recently TESS, and this is soon to be added to by PLATO. In order to fully exploit this science potential we need automated methods to extract seismic parameters. Efforts to automate identifying the presence of modes have included searching for peaks with a signal-to-noise ratio above a threshold (e.g. Appourchaux et al., 2012), and implementing a continuous wavelet transform to highlight peaks in a spectrum (García Saravia Ortiz de Montellano et al., 2018). Following identification, modes can then be peak-bagged.

ABBA (Kallinger, 2019), `FAMED` (Corsaro et al., 2020) and `PBjam` (Nielsen et al., 2021) are three tools which automate both the mode identification and model fitting processes to extract mode information. All three use Bayesian fitting techniques (either Markov Chain Monte Carlo (MCMC) or nested sampling) and begin by

identifying $l = 0/2$ modes pairwise in the spectrum. Whereas AAB is applicable to red giant stars, `FAMED` and `PBjam` can be applied to main sequence through to core-helium burning stars. All three pipelines require $\nu_{\rm max}$, $\Delta\nu$ and $T_{\rm eff}$ as user inputs along with the observed power spectrum, however, AAB and `FAMED` both also require the background signal in the spectrum to be characterised beforehand. Hence, even automated methods require some data preprocessing. In Chapter 6 we present one automated method for characterising the background and providing estimates for $\nu_{\rm max}$ and $T_{\rm eff}$ which could be implemented before applying the above pipelines.

2.4.2 Direct method

The global power spectrum contains a lot of information about stellar characteristics, whether that be from granulation signals, modes of solar-like oscillations, or activity signatures. Even when the signal-to-noise ratio of oscillations is too low to measure individual frequencies it is still possible to extract bulk stellar properties from fitting a model to the global power spectrum. What's more, these estimates are largely independent of stellar evolutionary models.

Pairing measurements of $\nu_{\rm max}$ and $\Delta\nu$ with estimates of effective temperature from spectroscopy we can use empirical scaling relations (Kjeldsen & Bedding, 1995; Huber et al., 2011b) to calculate global properties of the star such as its radius, R, mass, M, and surface gravity, g, as given by

$$\left(\frac{R}{R_\odot}\right) \simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max\odot}}\right)\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{1/2}, \tag{2.16}$$

$$\left(\frac{M}{M_\odot}\right) \simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max\odot}}\right)^{3}\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-4}\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{3/2}, \tag{2.17}$$

$$\left(\frac{g}{g_\odot}\right) \simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max\odot}}\right)\left(\frac{T_{\rm eff}}{T_{\rm eff\odot}}\right)^{1/2} \tag{2.18}$$

where the $\odot$ subscript denotes solar values taken here as $\nu_{\max\odot} = 3090 \pm 30\mu$Hz, $\Delta\nu_{\odot} = 135.1 \pm 0.1\mu$Hz, and $T_{\rm eff\odot} = 5777$K (Huber et al., 2011a, 2013). The accuracy of these relations to predict stellar radius and mass is thought to be approximately 5% and 10-15% respectively (Gaulme et al., 2016; Brogaard et al., 2018). The scaling relations are an essential tool to study bulk properties for a variety of stars, including ensemble studies or simply to obtain first order estimates of properties.

However, there are limitations to using the scaling relations in that stellar properties estimated from these equations can differ to those from independent measurements (e.g. Miglio et al., 2013; Mosser et al., 2013; Rodrigues et al., 2017; Brogaard et al., 2018; Kallinger et al., 2018; Hekker, 2020). For example, the scaling relations require a correction to $\Delta\nu$ since it does not scale perfectly as $\rho^{\frac{1}{2}}$ (e.g. White et al., 2011). To quantify this, Sharma et al. (2016) proposed inserting a correction factor into Equations 2.16, 2.17 and 2.18 which scales with metallicity, mass and age, and which is calibrated using stellar models. Similarly a correction factor is also needed for $\nu_{\max}$ particularly for red giant stars since they are structurally different to the Sun. Viani et al. (2017) suggested that this should scale with mean molecular weight and adiabatic exponent.

2.4.3 Grid-based modelling

An alternative to directly estimating stellar properties using global seismic observables is 'grid-based modelling' (Stello et al., 2009b; Basu et al., 2010; Quirion et al., 2010; Gai et al., 2011). Robust estimates can be made by combining classical constraints with asteroseismic measurements. Grid-based modelling uses asteroseismic inputs such as $\Delta\nu$ and $\nu_{\max}$, as well as spectroscopic or photometric observables which can include effective temperature, metallicity, and luminosity.

A set of theoretical observables are computed for each stellar model within a predetermined grid. Models are created with user-defined input interior physics to thoroughly cover a certain region of parameter space and to follow stellar evolution

theory. Theoretical observables are then compared to the data. By searching the grid of models a posterior distribution is formed based on the set of solutions that best fit the data with a certain amount of spread. Some pipelines are able to interpolate between grid points, both along and in-between evolutionary tracks (e.g. Rendle et al., 2019a; Zinn et al., 2020). This posterior is then used to predict stellar properties. Even more accurate and precise estimates can be made with the inclusion of individual frequencies, or instead separation ratios, as inputs to allow for detailed modelling of main sequence stars and red giants (e.g. Lebreton & Goupil, 2014; Reese et al., 2016; Pérez Hernández et al., 2016; Li et al., 2018).

The main advantages of grid-based asteroseismic analysis over direct methods are that predictions are constrained by equations of stellar structure and evolution. However, one drawback of stellar modelling is that they involve intensive calculations and so the input stellar physics must be simplified. The choice of physics used to create the models will impact parameter estimation, particularly mass, age and chemical composition (Chaplin & Miglio, 2013, and references therein). Consequently, pipelines that include slightly different physics or statistical analysis techniques will likely have differing predictions. This has been investigated by multiple projects, for example KAGES (Silva Aguirre et al., 2015; Davies et al., 2016) and LEGACY (Lund et al., 2017; Silva Aguirre et al., 2017). This simplification also means there will be some missing physics such as that describing the surface effect, although there is some effort to account for this (for a discussion see Basu & Kinnane 2018). Magnetic activity effects are also often not included in models but can impact the inferred stellar properties (see Chapter 5).

2.5 Observing the oscillations of stars

2.5.1 Photometry and Doppler velocity

We observe temporal variations in the brightness of the Sun or other stars due to resonant oscillations, among other background signals. In the fields of helio- and

asteroseismology, the two most typical measurements are in photometry or Doppler velocity. As described above, most measurements are an integration over the visible stellar disc therefore limiting asteroseismic or Sun-as-a-star observations to only low-degree modes. Only those instruments which resolve the solar disc provide information for higher degree oscillations.

Photometry produces lightcurves of intensity observations where fluctuations in flux are primarily attributed to variations in temperature. The sensitivity of observations to temperature fluctuations depends on the wavelength response of the instrument. Hence, when comparing measurements from different sources the wavelength dependence should be considered. For example, TESS has a redder response than *Kepler* resulting in amplitudes being $\sim 15\%$ lower for TESS (Lund, 2019).

Doppler velocity observations measure the line-of-sight velocity of a stellar surface due to Doppler shifts as it pulsates. Different spectral lines form at different heights in a stellar atmosphere and the amplitudes of oscillations also vary with height. Therefore the relative Doppler shifts of multiple spectral lines, or parts of a single line, reveal the pulsations. For example, BiSON instruments measure the relative intensities of the red and blue wings of the Potassium Fraunhofer line. If the spectral line is Doppler shifted the intensities in the red and blue wings will be unequal. From this the line-of-sight velocity is calculable.

There are many differences when comparing photometric to Doppler velocity measurements. Relative to oscillations, the granulation signal is much weaker in velocity observations than photometry so modes have higher signal-to-noise ratios. Additionally, we see that $\nu_{\rm max}$ exists at slightly higher frequency in Doppler observations than in photometric although the position of $\nu_{\rm max}$ in Doppler velocity measures depends on the spectral lines being used (e.g. Howe et al., 2020).

2.5.2 Ground-based observations

Ground-based observatories can in theory continuously measure oscillations of the Sun when they are combined to form networks. BiSON (Chaplin et al., 1996; Hale et al., 2016) is a series of six telescopes located around the world providing continuous Doppler velocity measurements of the Sun as a star (i.e. unresolved observations). Operating since 1976, it has surveyed the Sun for over four solar activity cycles making it the longest running network for helioseismology. Another group of detectors is GONG (Harvey et al., 1996), also comprised of six telescopes but making resolved-disc observations of the solar surface and therefore providing oscillations data on small spatial scales. Finally, the Stellar Observations Network Group (SONG) spectrograph (Grundahl et al., 2008) is also contributing to asteroseismology from the ground by observing the brightest stars in the sky in radial velocity (e.g. Beck et al., 2020).

There are however some problems associated with ground-based telescopes. The diurnal gaps produce aliasing in a power spectrum. The additional peaks, with separations of $(\text{day})^{-1} = 11.57\mu\text{Hz}$, can be confused with the characteristic pattern of solar-like oscillations making interpretation of the power spectrum challenging. However, with a higher duty cycle this aliasing effect can be reduced. In addition ground-based instruments are limited by weather problems which produce broad-band noise in a power spectrum and also affect the duty cycle, on average 82% for BiSON (Hale et al., 2016).

2.5.3 Space-based observations

Although much more expensive, space-based missions suffer fewer of these problems. The Canadian space telescope Microvariablity and Oscillations of STars (MOST) (Walker et al., 2003) was the first space-based mission with a dedicated asteroseismology programme, studying variable stars including several Delta Scuti variables as well as Sun-like stars such as Procyon and α CenA. Convection, Rotation and

planetary Transits (CoRoT) (Baglin et al., 2006) followed, with the primary objective to observe exoplanets but it also supplied the asteroseismic community with photometry surpassing the targetted ground-based observations available prior to its launch.

Its successor, *Kepler* (Borucki et al., 2010) opened up the field, simultaneously observing more than 197,000 stars with a long cadence of 29.4 minutes and 512 stars at a short cadence of 58.85 seconds. This short cadence was necessary to access the short oscillation periods of main sequence stars. However, long cadence observations could capture oscillations in red giant stars which pulsate more slowly. The *Kepler* telescope was continuously pointed at the same field for four years producing long datasets with exquisite frequency resolution enabling extraction of individual modes of oscillation. With *Kepler* it was not only possible to detect exoplanets with longer orbital periods than with CoRoT but the seismic data enabled main sequence stars to be observed with revolutionary precision.

After loss of two reaction wheels in 2012 and 2013, and hence without the ability to maintain stable pointing in the original field, the *Kepler* mission evolved into K2 (Howell et al., 2014). The K2 observation plan comprised of 20 observing campaigns surveying various fields across the ecliptic plane. Although unable to offer the same precision and multi-year duration lightcurves as *Kepler*, one advantage of K2 was enabling asteroseismic exploration in different regions of the Milky Way (e.g. Stello et al., 2015; Miglio et al., 2016; Rendle et al., 2019b).

The most recent contribution to space-based asteroseismic observations was Transiting Exoplanet Survey Satellite (TESS) (Ricker et al., 2015). Launched in 2018, it aimed to study 85% of the sky. For the two year nominal mission TESS conducted long cadence observations of 30 minutes and short cadences of 120 seconds, as well as providing full-frame images (FFI) at a cadence of 30 minutes. Its initial observational strategy was a step and stare approach, pointing at same sector in the sky for 27 days before shifting. Some parts of sectors overlapped providing

a continuous viewing zone of 351 days at the ecliptic poles. The TESS mission will detect approximately 3×10^5 stars with solar-like oscillations (Huber, 2018). Having observed the brightest stars in the sky (30-100 times brighter than *Kepler* targets Ricker 2019) it has produced fewer oscillation detections of main sequence stars but vast quantities for red giants.

Having completed its two year nominal mission as of summer 2020, TESS has begun its extended mission. This will be dedicated to revisiting the southern followed by the northern ecliptic hemispheres, although in slightly different orientation to the first observations (Barclay, 2020). An additional cadence of 20 seconds has also been introduced and the FFI collection will be at a shorter cadence of 10 minutes. Once launched, NASA's James Webb Space Telescope (JWST) field of view should overlap with the already-observed continuous viewing zone of TESS.

Finally, with the upcoming ESA mission PLAnetary Transits and Oscillations of stars (PLATO) (Rauer et al., 2014) on the horizon, we can expect higher quality data and long continuous pointing phases, one of which will overlap with the *Kepler* field of view. This will enable higher precision measures of frequencies (aiming down to 0.1μHz). With PLATO we expect to achieve precision of 1-2% in stellar radii and down to 10% in age with asteroseismology (ESA, 2017). This new mission will extend our knowledge beyond its primary objective of studying exoplanetary systems, but also has the potential to revolutionise our understanding of stellar structure, evolution and its relation to the Milky Way as a whole.

3 Stellar magnetic activity

Magnetic activity is present across the entire Hertzsprung-Russell (HR) diagram and is a typical occurrence in all solar-type stars (e.g. Baliunas et al., 1995; Radick et al., 1998; Linsky, 1999). Since we see similar characteristics of magnetic activity to the Sun in other stars, they are expected to share the same underlying mechanisms of generation and evolution. In order to develop our understanding of stellar magnetic activity it is essential to collect observational evidence. Studying stellar activity cycles and placing the Sun in context will help to reinforce the solar-stellar connection. We can investigate the dependence of activity on properties such as stellar mass, metallicity and age which is not possible by studying the Sun alone.

This chapter starts by outlining the motivations for studying solar and stellar magnetic activity and the fields of research which benefit from expanding our understanding. Section 3.2 will cover the ways in which high levels of magnetic activity manifests, the associated temporal variations, and a brief summary of how we observe activity with reference to the most influential instruments currently being used. We will start with observations of the Sun and move on to discoveries from other stars. Finally, Section 3.3 briefly reviews the signatures of stellar activity in asteroseismic measurements. This chapter should serve to cover the concepts needed for the analysis and results presented in Chapters 4 and 5.

3.1 Why study stellar magnetic activity?

There are many open questions about the magnetic nature of the Sun but one principal focus is to develop our theories on the solar dynamo. This refers to the

generation of magnetic field via dynamo action which theorises how an initially weak and poloidal magnetic field is amplified to the observed level. Figure 3.1 shows a schematic of the process. A large scale poloidal field is stretched and wrapped around the Sun due to differential rotation since the equatorial regions rotate more rapidly than nearer the poles. Over time this will form a toroidal field from the poloidal one in a process known as the Ω-effect (Panels a-c). Concentrations of magnetic field lines form bands of stronger field strength along lines of latitude appearing at $\pm 30^{\circ}$ either side of the equator.

Since plasma is an electrically conducting fluid it couples with magnetic flux to form flux tubes. The flux tubes contain magnetic field making them less dense than their surroundings. Due to this magnetic buoyancy, flux tubes rise towards the surface but since the Sun is rotating they are also acted upon by the Coriolis force. This causes the field lines to twist as they move radially outwards (see Panel d). This process produces small kinks in the toroidal field lines which can break through the photosphere and form manifestations of concentrated field strength such as sunspots. As the field lines continue to tighten, the concentration builds up and the bands of stronger magnetic field migrate towards the equator. We see evidence of this as sunspots become more numerous and move to lower latitudes.

The second half of the dynamo cycle is the mechanism by which the toroidal field returns to the poloidal configuration. The two main competing theories for this process are the α-effect (e.g. Bushby & Mason, 2004, and references therein) and the flux-transport dynamo, also known as the Babcock-Leighton mechanism, (e.g. Dikpati & Gilman, 2009, and references therein). To complete the dynamo cycle in the α-Ω model it has been theorised that the small twists of field lines form small-scale poloidal magnetic fields. These coalesce with a net effect that gives rise to the large-scale poloidal field with opposite polarity to the initial field (see Panels e and f). Alternatively, in the flux-transport dynamo magnetic flux is transported from the equator towards the poles by meridional flows in the near-surface layers

of the Sun (see Panels g and h). Both mechanisms form the reversed large-scale poloidal field (Panel i).

The entire process – from poloidal, to toroidal, back to poloidal magnetic field topology – is known as the Schwabe cycle and takes around 11 years for the Sun. This makes up half of the full 22-year magnetic cycle after which the polarity returns to the original orientation. Both convection and differential rotation play significant roles in field generation and activity cycles.

Improving our understanding of why and how activity changes on a temporal cycle is key to explaining observations. For example, why are solar cycles different lengths as is evident in the extra long cycle 23 and the unusually deep and wide minimum that followed (Sheeley, 2010; Broomhall, 2017)? Or why in addition to the periodic rise and fall of activity do we occasionally have very long periods of weak activity, such as that during the Maunder Minimum which occurred between 1645 and 1715? To improve our understanding of the solar dynamo, it is important to study other stars in order to test theories under different physical conditions.

Studying magnetic activity will help constrain our theories of stellar evolution and of the physical processes within stars. As described above, the generation and evolution of magnetic field is closely related to convection. For example, the decay of magnetic flux can be linked to convective diffusion and therefore affect the stellar diffusion coefficients needed in stellar modelling. Understanding the timescales associated with the growth and decay of activity is also important when determining stellar rotation periods since active regions must be present over multiple rotations (e.g. Nielsen et al., 2013; McQuillan et al., 2014; García et al., 2014; Santos et al., 2019a).

Rotation, and particularly differential rotation, is a key part of the generation of stellar magnetic field. Recent efforts have been to measure differential rotation for ensembles of stars to test theoretical models and to determine how it scales with stellar properties (Reinhold et al., 2013; McQuillan et al., 2014, e.g.). Asteroseismol-

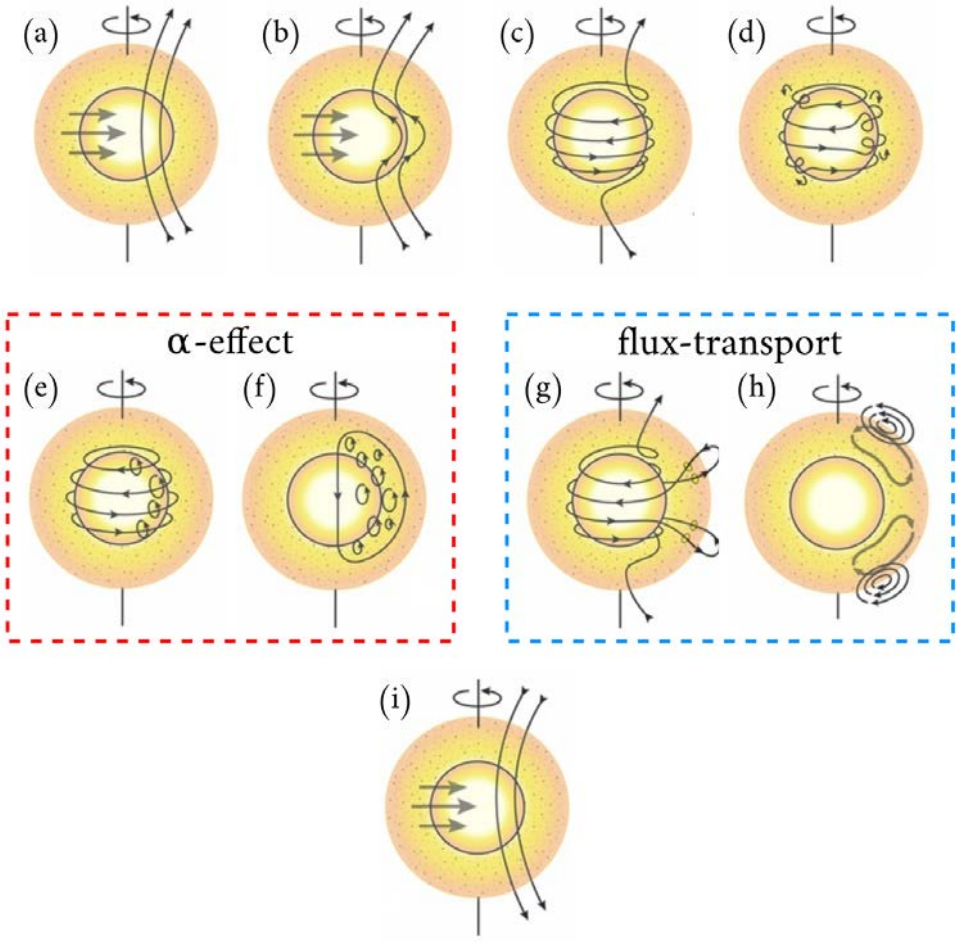


Figure 3.1: A schematic of the main processes within the solar dynamo model. The Ω-effect, shown in Panels a–c, describes how the initial poloidal field transforms into a toroidal field due to differential rotation. As flux tubes rise through the solar interior they are acted on by the Coriolis force causing them to twist (Panel d). Regeneration of a poloidal field with opposite polarity is accomplished by either the α-effect (Panels e and f) or the flux-transport mechanism (Panels g and h). This diagram depicts one half of the magnetic cycle. Sketches adapted from Sanchez et al. (2014)

ogy provides constraints on differential rotation (e.g. Reinhold et al., 2013; Benomar et al., 2018; Bazot et al., 2019) and can probe its dependence on evolutionary state (e.g. Nielsen et al., 2013). These findings impact dynamo models and can be used to calibrate them from the Sun to other stars.

Using asteroseismology to study activity provides a deeper understanding of the connection between age, rotation, and magnetic activity, both theoretically and observationally. By observing the magnetic nature of stars at different points in their evolution we can look beyond the Sun in its current state and make better predictions for how the Sun will change over time. As stars age they lose angular momentum via the stellar wind, known as magnetic braking. The rotation period can therefore be used to estimate age using empirical relations based on observations, a concept known as "Gyrochronology" (e.g. Barnes, 2007; Meibom et al., 2009; García et al., 2014; Davies et al., 2015). However, from asteroseismic age determinations, there is some evidence of a departure from this scaling relation for older stars (Metcalfe et al., 2014; Meibom et al., 2015; van Saders et al., 2016; Hall et al., 2021). Some stars older than the Sun rotate more rapidly than predicted by gyrochronology which van Saders et al. (2016) attributed to a change in magnetic braking efficiency occuring when the Rossby number (the ratio of rotation period to the timescale of convective turnover) extends beyond a critical value of ~ 2.

There has also been a large focus on investigating two branches of stars, separated by the number of rotations that occur within an activity cycle, as initially identified by Brandenburg et al. (1998) and further explored by Böhm-Vitense (2007). A version of this is shown in Figure 3.2. Stars appear to fall onto one of two sequences: either the 'active' branch with younger and more rapidly rotating stars, or the 'inactive' branch containing cooler, older and more slowly rotating stars. Later confirmed using *Kepler* data (e.g. Metcalfe et al., 2016), it is commonly thought that middle-aged stars transition from the active to inactive sequences. This is thought to be due to a sudden change in stellar characteristics at some point in a star's life,

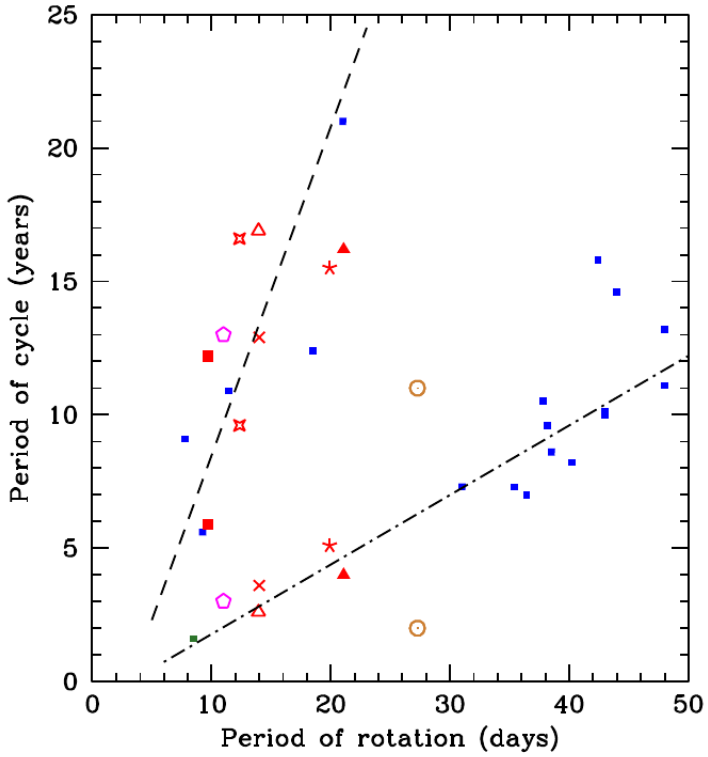


Figure 3.2: Graph showing the relationship between rotation rate and the period of activity cycle, including data from Böhm-Vitense (2007), Metcalfe et al. (2010) and Metcalfe et al. (2013). Blue points indicate stars with a single known cycle period and red points for stars with secondary cycles, where each star has its own unique marker. Stars appear to occupy either the active branch (upper) or inactive branch (lower) aside from the Sun which occupies the region in between as indicated by the orange solar symbols. Image taken from Chaplin & Basu (2015).

resulting in a reduction in magnetic braking efficiency and impacting how efficiently the rotation rate of stars slows down as they age. It has been suggested that the Sun, which occupies neither sequence, is currently undergoing this transition (see the orange points in Figure 3.2) (van Saders et al., 2016; Metcalfe & van Saders, 2017). More investigation is needed into how the stellar dynamos of these two branches differ and the fundamental changes that occur when a star undergoes this transition.

Since the level of activity is related to photometric variability (e.g. Karoff et al., 2016), it is important to be aware of the effects activity has on our observations. As discussed later in this chapter, we know that the modes of oscillation are affected

by activity which may impact whether they will be detectable for some solar-type pulsators. Campante et al. (2016) cautioned the need to consider the magnetic activity effect on mode amplitude when calculating the detectability of oscillations with the Transiting Exoplanet Survey Satellite (TESS). More recently, Mathur et al. (2019) investigated the cause for the non-detection of acoustic modes in a sample of 1,014 main sequence solar-type *Kepler* stars and found that magnetic activity could be responsible for amplitude suppression in 38% of stars. However, for the remainder of the sample they suggested the possible cause of low metallicity which leads to less energy being available for excitation of oscillations and hence modes have diminished amplitudes. It is important to recognise and know when to expect potential magnetic activity contamination when analysing observations or selecting targets, e.g. stellar flares, or jitter in radial velocity measurements.

Magnetic activity is also of interest to exoplanet researchers. When detecting exoplanets using radial velocity methods the signal induced by low mass or long period planets is of a similar amplitude if not smaller than that from stellar activity which could make disentangling detections difficult. Short timescale photometric variability due to activity is also a limiting factor when searching for Earth-analogues via the transit method, as are star spots which can mimic exoplanet signals (e.g. Pont et al., 2013). Typically the signatures of magnetic activity must first be removed or accounted for in models (e.g. using Gaussian Processes) before characterising exoplanets (e.g. Makarov et al., 2009; Haywood et al., 2014), particularly for those systems with orbital periods similar to the rotation period of the star (e.g. Newton et al., 2016; Vanderburg et al., 2016). Stellar activity is also expected to impact the habitability of exoplanets. Variations in stellar radiation and planet-star magnetosphere relationships have consequences for planet atmosphere evolution (e.g. Lammer et al., 2012). High levels of activity are often accompanied by strong flares and coronal mass ejections which will inevitably have a significant impact on protoplanetary disks and planet formation processes (Micela, 2017). Accounting for

stellar activity effects is therefore essential when characterising exoplanet systems and testing models of planet formation and evolution (for a review see Aigrain et al. 2016).

Finally, as an distinct field of its own, magnetic activity is studied for its implications in space weather and the Earth-Sun relationship. It is crucial to build a better picture of solar activity in order to reliably monitor and forecast its impact on the near-Earth environment, and practical issues such as space and air travel, the electricity grid, and communication systems.

3.2 Observing activity

3.2.1 On the Sun

From observations of the Sun – the only star for which we can resolve the surface in detail – there is evidence of magnetic activity in many forms. It is possible to observe manifestations of concentrated field strength and huge releases of energy include sunspots, solar flares, coronal mass ejections (CMEs), faculae, and plages. The timescales on which activity varies range from minutes for flares and CMEs, to thousands of years for long term changes in activity cycles. This means we need instruments with both high cadence measurements for rapid events, and long duration records for historical trends.

The ability to resolve the solar surface enables us to investigate certain activity manifestations at high resolution and to inspect their intricate structures in more detail than on other stars. One such feature are sunspots where magnetic field has broken through the photosphere and has inhibited convection, forming a dark patch of cooler material. They consist of a central 'umbra' surrounded by a hotter, and therefore slightly brighter, annulus known as the 'penumbra'. Spots typically appear in pairs of opposite polarity with one spot leading as they traverse the solar disk in the direction of surface rotation. The polarity of the leading spot is opposite in each hemisphere. Examining the magnetic nature of spots reveals that the polarity of

the leading spot swaps every 11 years providing evidence for the 22-year magnetic cycle. With such specific manifestations and the necessity to ensure the return of the poloidal field to its original orientation, sunspots act as tight constraints for dynamo models. The timescale for spot growth and decay is of the order of months (Solanki, 2003) which is long enough to be useful for stellar rotation determination. However, as alluded to above, if unaccounted for spots can prove an obstacle for exoplanets studies (e.g. Oshagh et al., 2013; Ioannidis et al., 2016; Morris et al., 2018; Tregloan-Reed & Unda-Sanzana, 2019).

Star spots are a very useful proxy for the level of activity since they increase in number and the fractional area of the disc they cover as the magnetic field lines become more concentrated and activity increases. The bottom panel of Figure 3.3 shows the proportion of the solar disk covered by sunspots as a function of time which clearly follows the solar cycle. Spots tend to exist in groups and are observed to occupy active bands of latitude that migrate; both are processes that constrain solar dynamo models. Spots appear at latitudes of $\sim 30^{\circ}$ in a band either side of the equator and move to lower latitudes as the cycle progresses. This is well illustrated in the "Butterfly Diagram" shown in the top panel of Figure 3.3. Towards activity minimum the fractional area covered by spots reduces and they begin to disappear from lower latitudes and appear higher on the solar disk as the new cycle begins.

Records of sunspots date back thousands of years with the earliest report from the Chinese Book of Changes (circa 800 BC, authors unknown). Counts of sunspots date back to 1610 (Usoskin, 2017) enabling studies which show that spots closely map out the 11- and 22-year cycles (e.g. Usoskin, 2013; Clette et al., 2014, amongst many other references). Variations in sunspot measures correlate with other activity proxies for which we can date further back in time. Even older records of solar activity are possible to reconstruct from isotope measurements covering the last ~ 9000 years (Wu et al., 2018). From these measurements, amongst other records, we know there have been rare and peculiar events such as the Maunder Minimum

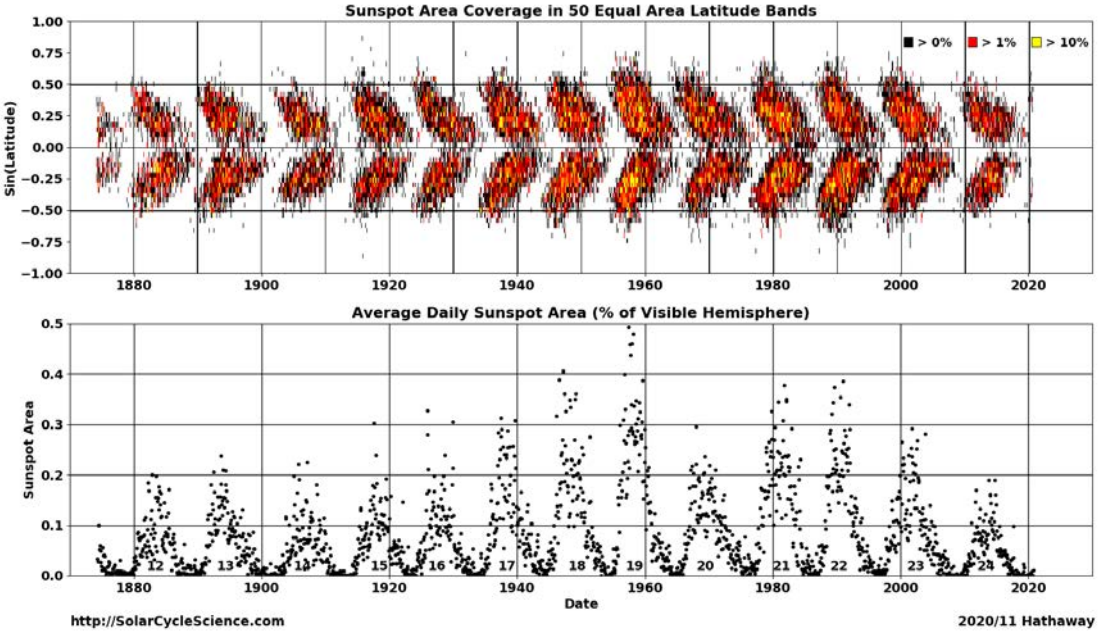


Figure 3.3: Solar activity over the last 145 years (1874-2020) showing the distribution of sunspot area as a function of latitude and time. The bottom panel shows the temporal variations in area as a proportion of the solar disk. The top panel shows the spatial distribution, with colour indicating the relative area which depends on latitude. The 11-year oscillation between activity maximum and minimum can be seen clearly, as can the migration of sunspots, beginning at $\sim 30°$ (sin(0.50)) and travelling towards the equator as solar activity grows. This figure is an updated version of Figure 9 from Hathaway (2015).

and the more recent Dalton Minimum. (For comments on what the helioseismic response of the Sun may have been in the Maunder Minimum see Chaplin et al. (2019).)

Beyond sunspots there are many ways we observe the Sun to probe its magnetic nature. Many space-based instruments investigate the outer layers of the Sun in extreme ultraviolet (EUV) and visible wavelength bands, to detect sunspots, flares, CMEs and other observable magnetic field features. Most relevant to this thesis are those that make observations useful for helioseismology. The Solar and Heliospheric Observatory (SOHO) had multiple instruments: Global Oscillations at Low Frequencies (GOLF) (Gabriel et al., 1995) and Variability of solar IRradiance and Gravity Oscillations (VIRGO) sunphotometer (SPM) (Fröhlich et al., 1995), which take Sun-as-a-star measurements of p mode oscillations, VIRGO's low-resolution imager (LOI) which measures solar radiance distribution across the solar disk, and the Michelson Doppler Imager (MDI) (Scherrer et al., 1995) for studying oscillations, and magnetic fields in the photosphere. Whereas GOLF and MDI take measurements in Doppler velocity, VIRGO uses photometry. The successor to MDI is the Helioseismic and Magnetic Imager (HMI) on board the Solar Dynamics Observatory (SDO) (Scherrer et al., 2012), which also takes Dopplergram and magnetogram observations. In addition, the Birmingham Solar Oscillations Network (BiSON) (Chaplin et al., 1996; Hale et al., 2016) and the Global Oscillation Network Group (GONG) (Harvey et al., 1996) are networks of ground-based observatories that observe the Sun in Doppler velocity and have frequently been used for magnetic activity studies. Operating since 1976, BiSON has utilised the activity-induced impacts on helioseismic measurements (discussed in Section 3.3) to investigate the solar activity cycle for over 40 years, thereby covering multiple Schwabe cycles and over a full magnetic cycle (e.g. Howe et al., 2018; Jain et al., 2018).

Helioseismology has shown that not all activity cycles have the same shape, amplitude, or duration which is in agreement with sunspot observations (e.g. Broomhall

& Nakariakov, 2015, and references therein). In addition to the 11-year solar cycle there is a quasi-biennial oscillation in the seismic data with a period of roughly two years which implies that there may be more than one type of solar dynamo operating at different depths (e.g. Broomhall et al., 2009; Fletcher et al., 2010; Broomhall et al., 2012). Since BiSON observes the Sun as a star it is only able to measure low degree modes of oscillation. One advantage is that this enables techniques developed through helioseismology to be verified on BiSON data before being applied to asteroseismic observations. GONG, MDI and HMI, however, resolve the solar surface allowing frequencies to be measured for a range of degrees ($l = 1 - 150$) and enabling seismic studies of activity on a finer spatial scale. Helioseismic observations allow investigation of both strong and weak components of magnetic flux in contrast to sunspots which only show regions of high magnetic flux. This is useful because weak component flux is spread more widely across the solar surface (e.g. Broomhall & Nakariakov, 2015; Thomas et al., 2019).

As produced by MDI and HMI, magnetograms are incredibly useful tools for studying magnetic fields on the surface of the Sun. Figure 3.4 shows example patterns of active regions on the solar surface from HMI in the middle and right-most panels and a simultaneous image at 171nm wavelength from SDO's the Atmospheric Imaging Assembly (AIA) instrument on the left. The magnetogram shows line-of-sight magnetic flux on the photosphere with positive and negative polarities shown in white and black respectively. Visible in all three images are the clear active regions lying along lines of constant latitude. Magnetograms and other global activity proxies, such as the 10.7cm radio flux (one of the longest records of solar activity) and Mg II H and K emission lines cores, correlate well with each other and reveal the solar cycle clearly (e.g. Jain et al., 2000; Tripathy et al., 2007; Broomhall & Nakariakov, 2015; Salabert et al., 2017).

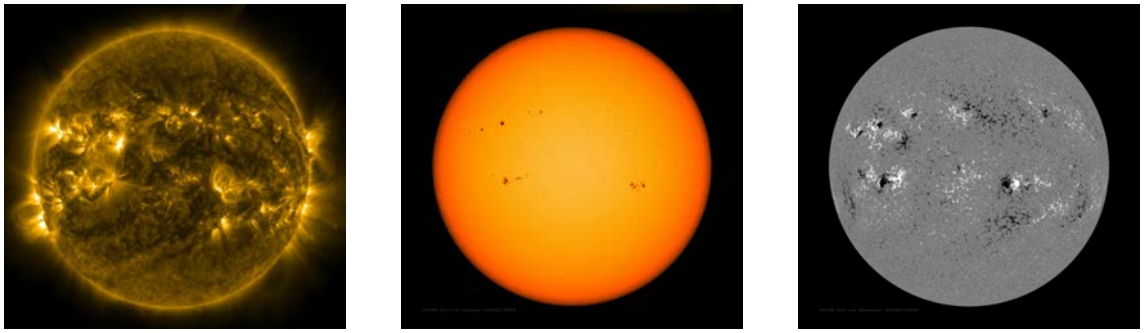


Figure 3.4: Simultaneous SDO views of the solar surface taken on 03/04/2014 (near solar maximum). Clear bands of active latitude are visible either side of the equator in all images. Left: EUV AIA image at 170nm wavelength. Magnetic field lines are highlighted by bright regions. Middle: HMI intensitygram showing sunspots. Right: HMI magnetogram. White and black indicate positive and negative polarity respectively. Images Courtesy of NASA/SDO and the AIA, EVE, and HMI science teams.

3.2.2 On other stars

We see the same type of variations in the magnetic activity of other stars as we see on the Sun. There is, however, controversy about whether the Sun is more or less active than other Sun-like stars (e.g. Basri et al., 2010, 2013; Reinhold et al., 2020). Some have been found to be much more active but some show no detectable variability at all (much like low activity epochs experienced by the Sun in the past). Gilliland et al. (2011) analysed *Kepler* observations and found that more than half of stars in the *Kepler* field were more active than the Sun. One difficulty faced with investigating stellar as opposed to solar activity cycles is achieving good enough frequency resolution to resolve power in oscillations spectra. In addition, although we do see clear evidence of stellar cycles with *Kepler*, the timespan of observations is not long enough to capture full cycles of a similar length to the Sun (*Kepler* was only observing the same patch of the sky for 4 years). Nevertheless, there are still many activity indicators with which investigating stellar cycles is possible.

A common proxy of activity used for stars is Ca II H & K lines, with one of the most complete catalogues being the Mount Wilson HK project (Duncan et al., 1991). These absorption lines, arising due to heating in chromospheric regions of

stars, have cores which exhibit emission in response to the presence of magnetic flux. From measures of emission lines, a chromospheric activity proxy known as the S-index can be calculated. The correlation between S-index and activity level has been shown for the solar cycle (Harvey & White, 1999; Pevtsov et al., 2016) as well as numerous studies for activity variations on other stars (e.g. Baliunas et al., 1995; Henry et al., 1996; Wright et al., 2004; Jenkins et al., 2006; Oláh et al., 2009; Isaacson & Fischer, 2010; Metcalfe et al., 2013; Boro Saikia et al., 2018a). An alternative proxy, known as R'_{HK}, also uses Ca II H & K emission lines but removes the temperature dependence and photometric contributions (Noyes et al., 1984).

An alternative technique for studying stellar magnetic activity is to use changes in stellar brightness over time. Modulations in a lightcurve can be caused by dark active regions or bright faculae transiting the stellar disk with changes in brightness on timescales of the solar rotation rate. Activity indicators can be obtained either by measuring the absolute stellar brightness (discussed later) or by using the modulation of this variability in the lightcurve over long timescales (of the order of activity cycle periods).

One photometric counterpart to the chromospheric indicator above is the $S_{\rm ph}$-index. Developed by García et al. (2010), it uses this concept of variability of a lightcurve as an activity proxy and is observable for stars whose photometric variability is dominated by either faculae or star spots. Mathur et al. (2014b) derived the $S_{\rm ph}$-index by computing the standard deviation of lightcurve flux over the entire time series to act as an indicator of the general level activity of a star. More commonly used is the $\langle S_{\rm ph} \rangle$-index, calculated as the average of several standard deviation measurements, each from independent subseries of the data with length 5 times the rotation period of the star. This proxy can be used to study any temporal variation of activity. The period of rotation is required and is best calculated using a combination of the autocorrelation function of the lightcurve and wavelet transform analysis as suggested by Aigrain et al. (2015). Evidence has been found of the $S_{\rm ph}$-

index correlating with other solar activity proxies, such as Ca H&K line emission and sunspots (e.g. Salabert et al., 2017). However, this is more challenging for other stars since, like the S-index, the $S_{\rm ph}$- and $\langle S_{\rm ph} \rangle$-indices are sensitive to inclination angle, in that at low inclinations the modulation signal is not as variable and not all of the activity regions can be observed. In these cases $\langle S_{\rm ph} \rangle$ must be treated as a lower limit of activity. The $S_{\rm ph}$ activity indicator has been used to study the activity of M dwarfs (Mathur et al., 2014a) and F-type stars (Mathur et al., 2014b), discovering evidence of long-lived features alluding to these stars having active longitudes similar to those seen on the Sun. It has also been used to study solar-like main sequence stars to place the Sun's activity levels into context (e.g. García et al., 2014). Similar methods, such as measurements of variations in the total solar irradiance data, have also been used to discover photometric activity cycles in *Kepler* stars and compare activity levels to the Sun (Reinhold et al., 2017, 2020).

An alternative activity indicator is to use the long-term variability in the absolute brightness of stars. Where the $S_{\rm ph}$-index probes the short-term variations in lightcurves caused by spot modulation and therefore uses observations taken with relatively short cadence (approximately every 30 minutes), this proxy uses longer timescale data to investigate variations associated with magnetic cycles thereby averaging out this spot modulation. Montet et al. (2017) used *Kepler* full-frame images (FFI) to study long-term trends in brightness variations which are usually removed for asteroseismic and exoplanet analysis. Systematic long-term trends in brightness that were not associated with magnetic activity were accounted for using reference stars. The authors report that for the majority of Sun-like stars in their sample, those exhibiting large changes in absolute brightness also showed variability in their $S_{\rm ph}$-index. These could be split into two groups. For stars with photometric variations that were faculae dominated, at activity maximum the $S_{\rm ph}$-index was greater as was the absolute brightness of the star, i.e. they were correlated. This was found to be the case for the more slowly rotating stars in their sample. Conversely, for

spot dominated variability the indicators were anticorrelated since the $S_{\rm ph}$-index was again larger at activity maximum but the absolute brightness was lower (because of the greater prevalence of dark spots).

Another common technique for activity studies involving mapping the magnetic fields of stars is known as Zeeman-Doppler imaging (ZDI) (Semel, 1989; Donati & Brown, 1997). This method maps large-scale magnetic field orientation, topology and strength by analysing the polarised light of spectral lines. A magnetic map is reconstructed by modelling the variations in the Zeeman signal as the star rotates. ZDI has been successfully used to investigate large-scale spatial and temporal variations in activity for Sun-like stars (Boro Saikia et al., 2016, 2018b; Jeffers et al., 2018). Lehmann et al. (2020) used ZDI to study which magnetic field characteristics were the best tracers of stellar activity cycles in Sun-like stars, finding that the temporal evolution of the axi- and non-axi-symmetric components of the magnetic energy were most closely related. Although this is one of the only methods by which we can crudely resolve the surfaces of stars there are some limits to the technique in that it is only applicable to slowly rotating stars. More rapid rotation limits the spatial resolution of the image. In addition, ZDI typically underestimates magnetic field strength and complexity (e.g. Rosén & Kochukhov, 2012; Lehmann et al., 2019), and is unable to produce magnetic field maps for stars observed pole-on.

Stellar flares (e.g. Schaefer et al., 2000; Shibayama et al., 2013; Wu et al., 2015; Maehara et al., 2015) have been discovered on other Sun-like stars, although these flares are typically 1–4 orders of magnitude more energetic than those usually found on the Sun. A considerable number of studies have also investigated the presence of star spots on solar-type stars (e.g. Berdyugina, 2005; Bradshaw & Hartigan, 2014; Davenport, 2015; Giles et al., 2017; Nielsen et al., 2019), however those that are detectable are observed to cover a much larger surface area than sunspots due to lack of spatial resolution. Techniques such as spot modelling use photometric variations from star spots crossing the stellar surface to estimate their number and size.

Through modelling this signal in lightcurves it is possible to measure surface rotation rates and in some cases differential rotation (e.g. Nielsen et al., 2013; McQuillan et al., 2013; Reinhold et al., 2013). It is also possible to reconstruct the distribution of active regions on the stellar surface but caution must be taken due to degeneracies in star spot models between differential rotation and the evolution and numbers of spots (e.g. see discussion by Walkowicz et al. 2013 and Aigrain et al. 2015).

Techniques to constrain active latitudes and investigate the migration of active bands on other stars are sparse and have been focused primarily on mapping spots. As mentioned above, Doppler imaging methods can provide coarse maps of active regions but have significantly lower spatial fidelity than what is possible for the Sun. Using interferometry Roettenbacher et al. (2016) were able to map star spots on the nearby, old star ζ Andromedae, however the distribution of activity was found to be different to that on the Sun. Morris et al. (2017) used photometric data of the K-dwarf HAT-P-11 utilizing signals within a transit curve of one of its planets to measure the latitudinal distribution of star spots. They found two distinct bands of active latitude, one in each hemisphere, much like that on the Sun. Netto & Valio (2020) also employed this technique on a young solar-type star discovering that within each hemisphere it exhibits two bands of active latitude, one at low and one at high latitudes with a gap at $\sim 35°$. These studies were only possible due to the high obliquity of the planets' orbits allowing them to traverse a range of latitudes during their transits.

3.3 Asteroseismic signatures of activity

Asteroseismology allows us to probe the interiors of stars, particularly their convective regions and those where we observe differential rotation. It is an essential tool to reveal information about stellar activity cycles since the mechanisms occurring within convective regions are key drivers for magnetic activity. The activity behaviour of stars is not only revealed by rotational modulation signals in a lightcurve

(from star spots) but the oscillations of a star are also affected by the presence of magnetic fields. Acoustic modes are sensitive to changes in the magnetic field and so their characteristics in power spectra follow activity cycles. There have been several reviews on these topics (e.g. Chaplin & Basu, 2015; Basu, 2016; Brun & Browning, 2017; García & Ballot, 2019) covering the capabilities and limitations of the asteroseismic study of magnetic activity, but the principal effects will be summarised below. The changes caused by activity on oscillation modes allows particular features of a star's magnetic nature to be studied and can affect inference of bulk stellar properties. These two issues are the focus of my studies reported in Chapters 4 and 5.

Asteroseismology as an activity proxy can be used to investigate near-surface magnetic fields for populations of stars and is one that is sensitive to both weak and strong component flux. Progressively, asteroseismic activity proxies have been used to study the cycles of more and more stars, for example to build our understanding of the relationship between activity and some stellar properties (e.g Chaplin et al., 2007a; Metcalfe et al., 2007; Santos et al., 2019b). With the advent of space-based missions like Convection, Rotation and planetary Transits (CoRoT), *Kepler* and now TESS, photometric variability is high enough quality to be useful to activity cycle studies, whereas before the reliance was on Ca II HK emission line data.

3.3.1 Frequency shifts

One asteroseismic signature of magnetic activity, and the one most thoroughly studied, is a frequency shift experienced by acoustic modes. In the case of the Sun we observe low-l modes nearest $\nu_{\rm max}$ to experience a shift of the order of $\sim 0.4\mu$Hz between solar minimum and maximum (e.g. Broomhall et al., 2014). Typically, p modes are shifted to higher frequency with increased activity (e.g. Howe et al., 2008; Rhodes et al., 2011). The cause of these frequency shifts can be due to two sources. Modes can be directly affected by being subject to the Lorentz force which acts as

an additional restoring force driving the waves to higher frequencies. The indirect effect is where the magnetic field alters the physical characteristics of the cavities in which the waves form and are trapped, e.g. the sound speed, or location of cavity boundaries, thereby affecting the way they propagate. This can increase or decrease the oscillation frequencies.

The size of shift experienced by modes is known to vary with the solar cycle. Starting in the 1980s, Woodard & Noyes (1985) found evidence in data from the Solar Maximum Mission satellite of low-l frequencies reducing over time when solar activity was decreasing during the epoch 1980-1984. From there followed a plethora of solar studies into cyclic variations in the mode parameters of low degree acoustic modes (e.g. Pallé et al., 1990; Elsworth et al., 1990; Libbrecht & Woodard, 1990; Chaplin et al., 2001; Jiménez et al., 2002; Chaplin et al., 2004; Salabert et al., 2015; Howe et al., 2020) with several studies providing evidence for the correlation between activity proxies, including frequency shifts (e.g. Salabert et al., 2009; Broomhall & Nakariakov, 2015; Howe et al., 2018). However, differences between asteroseismic and other activity proxies have been found for the Sun during the minimum between solar cycles 23 and 24, showing that even when variations are not visible in standard activity proxies there still exists magnetic perturbations just below the surface that could affect oscillations (Broomhall et al., 2009; Salabert et al., 2009, 2015). For more discussion and a more full collection of past studies see reviews on this topic by Broomhall et al. (2014) and Chaplin (2014).

The first observation of stellar-cycle induced asteroseismic variability in a Sun-like star was presented by García et al. (2010) who measured modes which had experienced temporal variations that tracked an activity proxy derived from photometric data. This evidence indicated a much shorter cycle than that of the Sun, being between 1 to 2 years. Since then, several studies have used data from the *Kepler* telescope to show similar asteroseismic signatures of stellar cycles (e.g. Salabert et al., 2011; Mathur et al., 2013; Mathur, 2014; Régulo et al., 2016; Salabert

et al., 2016a,b; Kiefer et al., 2017; Santos et al., 2018, 2019b; Mathur et al., 2019). Santos et al. (2018) analysed 87 solar-type *Kepler* stars finding that more than half of their sample exhibited time varying frequency shifts, with most tracking the variations of other activity proxies. However, recently Karoff et al. (2019) compared asteroseismic proxies to chromospheric and photometric activity indicators and concluded that they rarely show correlation. They did however suggest that this can be explained by the sparsity with which ground-based chromospheric measurements used in the study had been taken and that they were unable to observe enough of a stellar cycle using the 4-year nominal *Kepler* mission data. We may soon obtain more samples to populate these comparison studies with the addition of TESS observations and measurements from the new Stellar Observations Network Group (SONG) spectrograph.

In the case of the Sun we now have collections of data spanning multiple decades and hence covering multiple solar cycles. Chaplin et al. (2007b) used BiSON Sun-as-a-star data from 1976-2007 spanning three solar cycles to investigate the correlations between frequency shifts of low degree p modes and six different activity proxies. They found that those proxies that were also sensitive to weak component magnetic flux were more closely related to frequency variations. Over the next few years we can begin to build these long time span data sets for other stars and study stellar activity cycles more similar in length to the Sun.

Knowing the spatial range in latitude over which magnetic field strength is stronger on stars is a key constraint in dynamo theory. On the Sun this can be observed through sunspots, however, it is not possible for other stars since we cannot directly image their surfaces at sufficient resolution to discriminate individual features so it is unreliable to use star spot data. As mentioned above, there have been a handful of studies which explore the spatial distribution of activity via the spread of star spots. However, this is just one indicator of activity and one which is only sensitive to high magnetic flux. This is where asteroseismology comes in, since

we can dissect our asteroseismic measurements into spherical harmonics and use this to identify regions where activity is higher. By having more detailed maps of the surface activity distribution we can improve models, including our descriptions of stellar dynamos and our understanding of stellar spin-down (Garraffo et al., 2016).

The size of frequency shift experienced by a mode depends on its angular degree and azimuthal order, and on the magnetic field strength of the component which occupies the same spherical harmonic projection. Modes are most sensitive to near-surface perturbations. Since for stars exhibiting a Sun-like dynamo the distribution of activity on a stellar surface is not spherically symmetric, modes which occupy different regions on the stellar surface will experience different magnitudes of shift. In Chapter 4 we present how the relative responses of low-l modes to non-homogeneously spread magnetic activity can be used to tell us something about the distribution of activity. GONG measurements allow isolation of higher degree modes and therefore high spatial resolution solar observations. The resulting distribution of frequency shifts as a function of latitude and time look very similar to butterfly diagrams constructed using sunspots. This can be seen in Figure 3.5 where we can depict regions of higher activity moving towards the equator during the solar cycle due to the responses of modes of different degree (averaged over $l = 40 - 80$ and $n = 9 - 11$) and their spatial representation. Unfortunately, for other stars we are observationally limited to low-l modes due to geometric cancellation since their surfaces cannot be resolved. Therefore many studies focus on developing techniques using Sun-as-a-star data which can then be applied to data from *Kepler*, CoRoT, or TESS.

As will be described in Chapter 4, the visibility of azimuthal mode components depends on the inclination of the star (e.g. Gizon & Solanki, 2003; Chaplin, 2014). This means that when comparing the relative frequency shifts of modes made up of different m, the inclination angle must also be considered. Figure 3.6 shows the relative visibilities of the components of $l = 0, 1, 2$ modes as a function of inclination

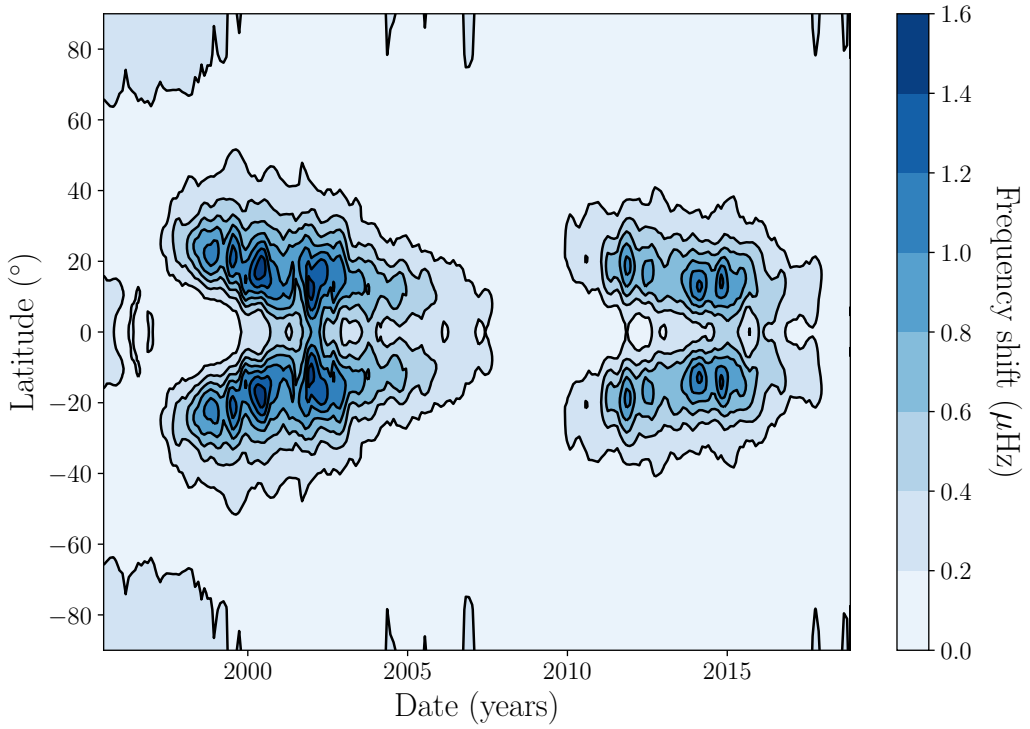


Figure 3.5: Mode frequency shifts as a function of time and latitude. The values are averaged over $l = 40 - 80$ and $n = 9 - 11$ from analysis of GONG observations. Data courtesy of Rachel Howe.

angle. In the solar case, where the axis of rotation is perpendicular to the line of sight and hence the inclination angle $i \simeq 90°$, the $m = \pm 1$ components of a dipole mode are visible however the $m = 0$ component has zero amplitude, as indicated by the green line in Figure 3.7. This will mean that the overall shift of a $l = 1$ mode averaged over all m will be controlled solely by the $m = \pm 1$ shift. However, for a quadrupole mode the visible components are $m = \pm 2$ and $m = 0$ and it is the $m = \pm 1$ contribution which is zero (also in Figure 3.7). Therefore the overall shift of a $l = 2$ mode will be made up of contributions from different m. For stars other than the Sun with inclination angles away from $i = 90°$ the visibilities of each m component will be different again (see Figure 3.7 for examples of multiplets at low, high and intermediate inclination angles). The relative contributions of each azimuthal mode can be modelled using Equation 2.15 as long as the inclination angle

is known. Chapter 4 discusses this in more detail and how it can be used to infer surface activity distribution.

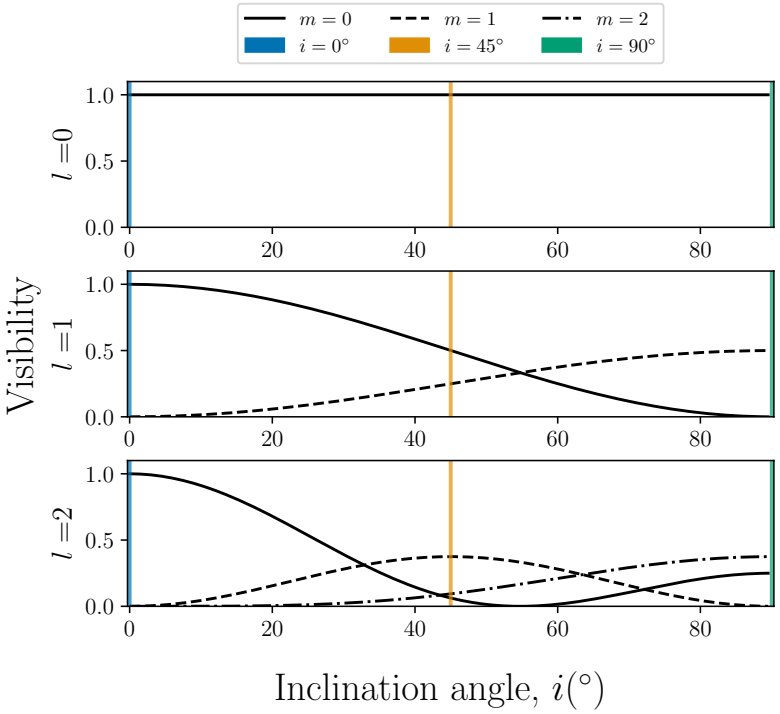


Figure 3.6: Relative visibilities of the components of $l = 0, 1, 2$ modes as a function of inclination angle. The coloured vertical lines refer to the inclination angles used in Figure 3.7.

Asymmetry in the distribution of oscillation frequencies within a multiplet may also be an activity indicator. As mentioned above, modes of different m have different shift responses in a non-spherically symmetric magnetic field resulting in the multiplet for a particular degree mode having an asymmetric profile in frequency. If we attempt to make inference on data taken over a longer timescale than activity effects, the Lorentzian profiles of individual mode peaks may be affected and biases may incur in mode heights and widths comparable to their uncertainties. For example, if the modes experienced activity-induced shifts over the duration of the data then their Lorentzian peaks may be distorted and have a flattened appearance. Chaplin et al. (2008) explored this issue for solar p modes and advised caution when analysing multiple years worth of data on stars with high levels of activity.

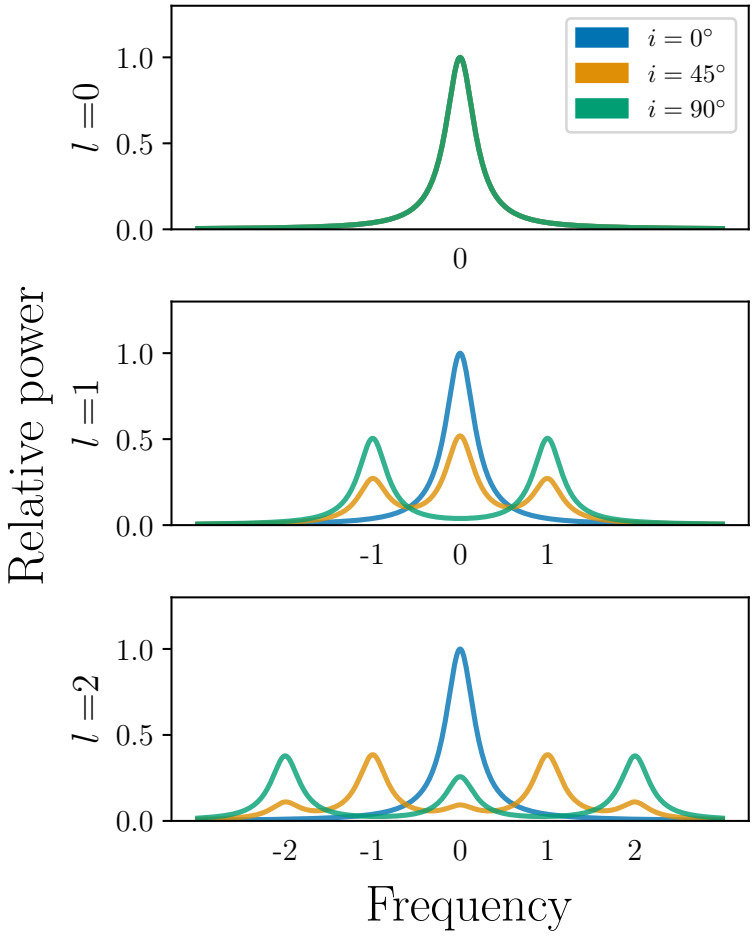


Figure 3.7: The shape of multiplets for $l = 0, 1, 2$ modes at inclination angles of 0°, 45° and 90°.

It is widely accepted that there is a frequency dependence to the shifts such that in general they are larger for higher frequencies. We know that for a given l, higher frequency modes have more shallow upper turning points. In addition, at a fixed frequency the higher l modes do not penetrate as deeply as lower l. Therefore modes of higher frequency and lower mode inertia (higher l) are more sensitive to the surface conditions and hence will undergo larger shifts. Since the mode shifts do not scale like $\propto \frac{\nu}{\sqrt{l(l+1)}}$, i.e. they are largely independent of angular degree, the implication is that the mode changes are due to near surface activity and are not spread through a larger proportion of the star (Libbrecht & Woodard, 1990).

Evidence of this frequency dependence has long existed for the Sun (e.g. Chaplin et al., 2001; Salabert et al., 2015, 2018) and has also been seen in some stars with asteroseismic shifts (e.g. Salabert et al., 2011; Kiefer et al., 2017; Salabert et al., 2018).

Finally, it is worth noting that since modes experience frequency shifts it is unsurprising that global asteroseismic quantities will also show signatures of magnetic activity. Temporal variations in the large frequency separations have been observed in the Sun (e.g. Broomhall et al., 2011) and more recently, using BiSON data, also discovered in $\nu_{\rm max}$ by Howe et al. (2020). These are particularly useful activity proxies since they can be measured for many stars where it may not be possible to identify individual modes.

3.3.2 Mode amplitude and mode damping

The amplitudes of modes are known to diminish with increased activity (e.g. Chaplin et al., 2011c; Mathur et al., 2019). The presence of a stronger magnetic field inhibits motion in convective regions resulting in less energy being available to excite modes. Helioseismic analyses on local scales with GONG and MDI data have revealed that the amplitudes of oscillations are inversely proportional to the level of activity (e.g. Komm et al., 2000a; Howe et al., 2004; Tripathy et al., 2006). Braun et al. (1988) compared the amplitudes of high-degree p mode oscillations in multiple regions around sunspots and found absorption of the mode power to occur not just in the umbrae but also the penumbrae and surrounding area containing extended magnetic fields. On more global scales, low-degree mode amplitudes show similar behaviour both on the Sun (e.g. Pallé et al., 1990; Anguera Gubau et al., 1992; Chaplin et al., 2000) and for other solar-like pulsators (e.g. Mathur et al., 2011a; Chaplin et al., 2011c).

With higher activity we also see a broadening in mode width suggesting an increased damping or reduced mode lifetime. One of the first reports of this effect in

solar data was by Toutain & Wehrli (1997) using early measurements from VIRGO and the space experiment IPHIR (the solar irradiance experiments on the Soviet PHOBOS satellites). Both instruments observed p modes using Sun photometers to suggest that differences in measured mode parameters were due to different levels of activity during their respective observing periods. More recently, evidence of anticorrelation between mode amplitude and activity level has been found in high-degree mode observations (e.g. Komm et al., 2000b; Burtseva et al., 2009) as well as low-degree oscillations on the Sun (e.g. Jiménez et al., 2002; Jiménez-Reyes et al., 2003; Salabert et al., 2007; Broomhall et al., 2015) and other stars (e.g. García et al., 2010).

A consequence of depressed mode amplitudes may be that oscillations are not detectable in some stars where we would expect them. For example, Chaplin et al. (2011c) conducted a statistical study of *Kepler* stars finding that higher levels of magnetic activity inhibit the amplitudes of acoustic modes resulting in the oscillation signatures of those stars being undetectable. As mentioned previously, the $S_{\rm ph}$ index activity proxy used by Mathur et al. (2019) indicated that the reduction in mode amplitude was partially accountable to magnetic activity, as well as metallicity.

4 Constraining magnetic activity distribution with asteroseismology

This chapter is taken verbatim from Thomas et al. (2019), published in Monthly Notices of the Royal Astronomical Society, Volume 485, Issue 3, pages 3857-3868, for which I was first author. As such, some of the text may repeat the concepts covered in earlier chapters. All of the work in this chapter was completed by myself with the exception of extracting the frequency data. Dr Ângela Santos provided the Kepler *frequency shifts of HD 173701 and the inclination angle posterior was supplied by Dr Guy Davies. The BiSON, VIRGO and GONG frequency shifts were all provided by Dr Rachel Howe and the the National Solar Observatory (NSO) magnetic field strength data were given to us by Dr Rudolf Komm.*

This paper reports a new method which uses measurements of the frequency positions of oscillation modes over time to constrain the distribution in latitude of near-surface activity on stars like the Sun. The principle of our method was first proved using the Sun and then applied to the solar-type star HD 173701, a well studied solar-analogue (e.g. Karoff et al., 2013, 2018) with almost identical mass, radius and age to the Sun but with a higher metallicity (Buchhave & Latham, 2015; Creevey et al., 2017). It is known to have unusually active characteristics with frequency shifts of up to $1\mu Hz$ *(Santos et al., 2018) and a cycle amplitude observed in a chromospheric activity index that is 2.7 times larger than the Sun (Karoff et al., 2018).*

A natural follow-on from this work would be to extend our method to other stars that have been demonstrated to show stellar activity cycles (e.g. Salabert et al., 2011; Régulo et al., 2016; Salabert et al., 2016b; Kiefer et al., 2017; Santos et al., 2018). The method presented here relies on high quality frequencies being measurable from short duration lightcurves in order that a temporal evolution in shifts is observable. Frequency shifts from short timescale lightcurves as well constrained as those of HD 173701 may not be observable for other Kepler *Sun-like stars. Similarly, in order to investigate any latitudinal migration of active bands we require either that the frequency measurement uncertainties are small or the shifts are very large. At the resolution available to us it was not possible to observe any migration of the activity bands for HD 173701. This kind of investigation may be possible with PLATO which should achieve frequency precisions of* $0.04 - 0.09 \mu Hz$ *from 150-day long lightcurves (Miglio et al., 2017).*

4.1 Introduction

Spatially resolved observations of the Sun reveal that regions of strong surface magnetic activity are not distributed uniformly over the solar surface, but instead lie in so-called active bands of latitude either side of the equator (Carrington, 1858). These bands migrate during the solar cycle and this can be observed by mapping surface magnetic fields over time (e.g. Howard & Labonte, 1981) or by resolving the solar surface and studying the positions of sunspots giving rise to the classic butterfly diagram (e.g. Maunder, 1904). The distribution of sunspots is a manifestation of the action of the Sun's internal dynamo, which converts poloidal into toroidal field during the rising phase of the activity cycle. Concentrated field lines form in bands parallel to the equator and, as the cycle progresses, move towards lower latitudes. It is reasonable to assume that other solar-type stars are likely to show active bands of latitude.

Some techniques exist to map the topology of magnetic activity on stellar surfaces. One method uses spot-induced modulations in stellar lightcurves to measure surface differential rotation and, if the inclination angle of the star is known, to constrain the latitudes of the spots (e.g. Eaton & Hall, 1979; Lanza et al., 2003, 2007; Silva, 2003; Santos et al., 2017). Tests on artificial data (Aigrain et al., 2015) suggest that results from such analyses may need to be treated with caution. Other methods include the Zeeman Doppler imaging technique (Fares, 2014, and references therein) and spectro-polarimetric observations (Piskunov & Kochukhov, 2002; Kochukhov et al., 2014). These have been applied to Ap and Bp stars (e.g. Rusomarov et al., 2015; Kochukhov et al., 2011) as well as to Sun-like exoplanet hosts (e.g. Fares et al., 2013; Alvarado-Gómez et al., 2015).

Chaplin et al. (2007a) and Chaplin (2014) suggested that it should be possible to use stellar-cycle-induced frequency shifts of solar-like oscillations in Sun-like stars to place constraints on the latitudinal dependence of the near-surface magnetic activity.

That frequencies of solar acoustic (p) modes vary with the solar cycle is now a well established result (e.g. Libbrecht & Woodard, 1990; Elsworth et al., 1990; Jimenez-Reyes et al., 1998; Howe et al., 2002; Chaplin et al., 2004, 2007b). Temporal changes of p mode oscillation frequencies have now been observed in several Sun-like stars (e.g. García et al., 2010; Salabert et al., 2016b; Régulo et al., 2016; Kiefer et al., 2017; Salabert et al., 2018; Santos et al., 2018) thanks to the availability of long, high-cadence, high-quality lightcurves from the French-led CoRoT satellite, and in particular the NASA *Kepler* Mission. This includes the solar analogue HD 173701 (HIP 91949, KIC 8006161), which is the focus of this paper. With near-solar radius, mass and age, it exhibits solar-like differential rotation and a 7.4-year activity cycle, making it a particularly interesting object for comparison with the Sun (Karoff et al., 2018). Its p mode frequency shifts have been found to be larger than those seen in the Sun (e.g. Kiefer et al., 2017), which Karoff et al. attributed to a higher metallicity and a deeper outer convection zone.

The layout of the rest of the paper is as follows. We begin in Section 4.2 by explaining the basic principles of using frequency shifts of solar-like oscillations to constrain the location of the active latitudes on a star. As we shall see, the technique has the potential to provide particularly useful constraints for dynamo modellers, since it responds to the latitude range over which significant magnetic flux is present over and above basal, ephemeral levels.

In order to validate our approach we used Sun-as-a-star helioseismic data and artificial asteroseismic data, before applying the methodology to over 3 years of *Kepler* data on HD 173701. The data we used are described in Section 4.3. We then present in detail in Section 4.4 the two complementary ways in which we used the frequency shift data to infer the active latitudes. The results from verification using solar and artificial data are presented in Section 4.5.1. We then apply the same methods to the frequencies of HD 173701 (Section 4.5.2), which was observed during the nominal *Kepler* Mission. We conclude our paper with a summary of the

key findings in Section 4.6.

4.2 Basic principles of the method

Sun-like stars show a rich spectrum of overtones of solar-like oscillations, acoustic pulsations that are stochastically excited and intrinsically damped by near-surface convection. Since gradients of pressure act to restore the modes, they are often colloquially referred to as p modes. The spatial properties of the modes may be described by spherical harmonics of angular degree l and azimuthal order m. Any spherical asymmetry, such as rotation or magnetic fields, can lift the degeneracy in l, so that each non-radial mode is split into $2l + 1$ azimuthal components, where $-l \leq m \leq l$. Overtones of each l are labelled by a radial order, n.

The frequencies of the modes shift in the presence of magnetic activity, either due to the direct effects of the Lorentz force or indirectly from magnetic-field-induced changes in the sound speed within the mode cavities. With a non-homogeneous distribution of magnetic activity, modes of different angular degree and azimuthal order will experience different frequency shifts (Moreno-Insertis & Solanki, 2000). The various combinations of l and m have different spatial distributions on a sphere and therefore the responses, in amplitude and phase, of these components depend on the magnetic field strength in the regions where the acoustic waves forming them propagate. For this work we assume that magnetic and spin axes are aligned.

Under the assumption that the source driving the sound-speed perturbations, and therefore the shifts in frequency, is located close to the stellar surface, those shifts may be described as being proportional to:

$$\delta\nu_{lm} \propto \left(l + \frac{1}{2}\right) \frac{(l - |m|)!}{(l + |m|)!} \int_{\theta_{\min}}^{\theta_{\max}} |P_l^{|m|}(\cos\theta)|^2 \, B(\theta) \, \sin\theta \, d\theta, \tag{4.1}$$

where the co-latitude $\theta = (\frac{\pi}{2} - \lambda)$ and λ is the latitude, $B(\theta)$ is the magnetic field strength, and $P_l^{|m|}(\cos\theta)$ are the associated Legendre polynomials (Moreno-

Insertis & Solanki, 2000). Note how the magnitude of the shift depends on (l,$|m|$). The relative sizes of the frequency shifts therefore give information on the spatial distribution of the near-surface activity. This relation can be calibrated in order to get absolute frequency shifts (for example calibrated to the Sun).

From highly spatially resolved observations of the Sun over time it is possible to measure the frequency shifts of modes of individual (l,m) up to degrees that provide high-fidelity constraints on the spatial properties of the near-surface activity. For example, Howe et al. (2002) used frequency shifts of modes covering the range $l = 0$ to 150 observed by the Global Oscillations Network Group (GONG) – with effectively a separate frequency available for each azimuthal order m – to produce a map of levels of shift distributed over latitude and time. This map looks very similar to butterfly diagrams that show spatial and temporal changes in levels of magnetic field.

For unresolved observations of stars, geometric cancellation means we are unable to obtain data over such a wide range in l. Using intensity measurements from *Kepler* it is possible to detect only overtones of $l = 0$, 1 and 2, and in a few cases $l = 3$, and so we must accept that any latitudinal localisation from asteroseismology will be less refined than is possible in the resolved-Sun case. Nevertheless, shifts of varying size have been observed for different low-degree solar p modes by Chaplin et al. (2004), Jiménez-Reyes et al. (2004), Chaplin et al. (2007b) and Salabert et al. (2015) that are entirely consistent with the non-spherically-symmetric distribution of near-surface activity on the Sun. In what follows we make use of data on $l = 0$ and $l = 1$ modes, since $l = 2$ frequency shifts inferred from *Kepler* data are noticeably noisier (owing to their lower visibility in unresolved observations compared to the radial and dipole modes).

For simplicity, and given the above constraints, we model the magnetic field distribution as a top-hat function assuming bands of activity wrapped around a star

parallel to its equator, i.e.,

$$B(\theta = \frac{\pi}{2} - \lambda) = \begin{cases} b, & \text{if } \lambda_{\min} \leq \lambda \leq \lambda_{\max}, \\ 0, & \text{otherwise,} \end{cases} \tag{4.2}$$

where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum latitudes that define the band in each hemisphere – the southern hemisphere is assumed to be a reflection of the northern hemisphere since globally coherent modes provide an averaged measure of the properties above and below the stellar equator – and b is a constant which is independent of latitude. Using butterfly diagrams built from observations of the Sun we find evidence of solar activity distributions lying along such bands. We therefore use Equation 4.2 as an approximation of activity distribution which can be applied to other Sun-like stars.

In what follows we attempt to constrain distributions that are integrated over long epochs (entire rising or falling phases, or over an entire cycle). We therefore ignore any cyclic variations where the latitudinal dependence of activity may change across the period of time observed (within the resolution of observation) and instead seek a epoch-averaged distribution. Consequently, $\lambda_{\min}$ and $\lambda_{\max}$ are time-independent; the precision in the measurements of the frequency shifts does not allow us to constrain more complex models that include time variation of the latitudes. Although we acknowledge that this assumption is not entirely true since bands of activity migrate during the solar cycle, a justification is discussed in a later section. However, since b is representative of the time-varying field strength it is of course time-dependent.

By measuring the frequency shifts of modes of different (l,m) over time we may place constraints on $\lambda_{\min}$ and $\lambda_{\max}$. The frequency shifts needed for the analysis are calculated from estimates of mode frequencies over different epochs. We need observations of a star that span at least a significant fraction of the ascending or descending phase of its activity cycle. The full lightcurve is divided temporally into

shorter lightcurves, and frequencies extracted from each to enable time-dependent frequency shifts to be calculated.

The extraction of individual frequencies proceeds by fitting multi-parameter models to the resonant mode peaks in the frequency-power spectrum of a lightcurve, a process colloquially referred to as peak-bagging (e.g. Anderson et al., 1990; Handberg & Campante, 2011; Davies et al., 2015; Lund et al., 2017). Because the modes are damped, they manifest as Lorentzian-like peaks in the spectrum. The ability to resolve the individual m components of a given (n,l) multiplet depends on how the widths of peaks compare to the splittings between them, which components are detectable, and the frequency resolution of the data. The visibility of modes in the power spectrum (assuming photometric observations) is well described by

$$\varepsilon_{lm}(i) = \frac{(l-|m|)!}{(l+|m|)!}\left(P_l^{|m|}(\cos i)\right)^2, \tag{4.3}$$

where $\varepsilon_{lm}(i)$ is the mode visibility and $P_l^{|m|}(\cos i)$ are Legendre Polynomials (Gizon & Solanki, 2003). Depending on the angle of inclination i offered by the star, the above implies that due to geometric cancellation when averaging over the visible stellar disc, the non-spherically symmetric perturbations for some of the modes may have zero visibility.

Long, multi-year lightcurves give the exquisite frequency resolution needed to resolve individual m components of the non-radial modes. Equation 4.3 implies that if one can then measure the relative power in the detectable m, it is possible to constrain the angle of inclination of the star (for a detailed description see Davies et al. (2015)). However, multi-year lightcurves are too long to track stellar cycle changes and their frequency shifts. Instead we must divide the data into subsets of a few month's duration. The reduced frequency resolution then makes it extremely hard to disentangle the frequencies of the individual m components of the non-radial modes. Rather than having the luxury of working with shifts for individual

m components, this means that in practice we must use shifts for each non-radial (n,l) multiplet that are *weighted averages* of the shifts shown by the constituent m. The respective weighting depends on the relative visibilities of the different components, as described by Equation 4.3.

The usual practice is to fit a single frequency to each non-radial multiplet – in effect the weighted centroid of the contributing components (see Figure 4.10) – and a single frequency splitting to describe the separations between adjacent m components. Since the visibility of each m component is dependent on the angle of inclination of the star it is necessary to understand the relative contribution of each of the components to the centroid of the multiplet at different angles, i. For dipole ($l = 1$) modes, this balance may be described by

$$\nu_{nl} = \alpha\nu_{nl0} + (1 - \alpha)\nu_{nl|1|}, \tag{4.4}$$

where ν_{nl} is the central frequency of the multiplet, ν_{nl0} is the frequency of its constituent $m = 0$ component, and $\nu_{nl|1|}$ is the average frequency of the $m = -1$ and 1 components. The coefficient α describes the relative contributions of each azimuthal component which depend on the underlying mode parameters and the inclination angle of the star. These were calculated as detailed in Appendix A (placed at the end of this Chapter).

In summary, provided we have a good estimate from the full dataset of the angle of inclination of the star, we can work with m-averaged frequency shifts since we know the relative contribution of the different m to these estimates.

4.3 The data

We used just over 3 years of short-cadence *Kepler* data on HD 173701, covering observing quarters Q5.1 through Q17.2 inclusive from the nominal mission. This period spans most of a rising phase of the 7.4-year stellar activity cycle shown by the star (Santos et al., 2018). Lightcurves were corrected using the KASOC

filter (Handberg & Lund, 2014) and prepared as described in Lund et al. (2017). Frequencies were estimated for 90-day subsets of the full lightcurve (each overlapping by 45 days), as described in Santos et al. (2018). For reference later in the paper, the inclination angle of HD 173701 was estimated using the full lightcurve to be $i = 37.2^{+4.0}_{-4.0}$ degree, using the approach described in Davies et al. (2015); this is consistent with the value quoted in Karoff et al. (2018).

Figure 4.1 shows frequency shifts from these 90-day data, averaged over six orders of, respectively, $l = 0$ modes (top panel) and $l = 1$ modes (bottom panel), centered on the order showing the highest radial-mode amplitudes. The reasons for choosing these six orders are discussed in Section 4.4 below. The selected modes span the frequency range ≈ 3100 to $3900\,\mu$Hz. Results on the overlapping sets are shown in red, whilst the dark blue line (and its lighter $1\,\sigma$ uncertainty envelope) shows results on the independent data used in the analyses.

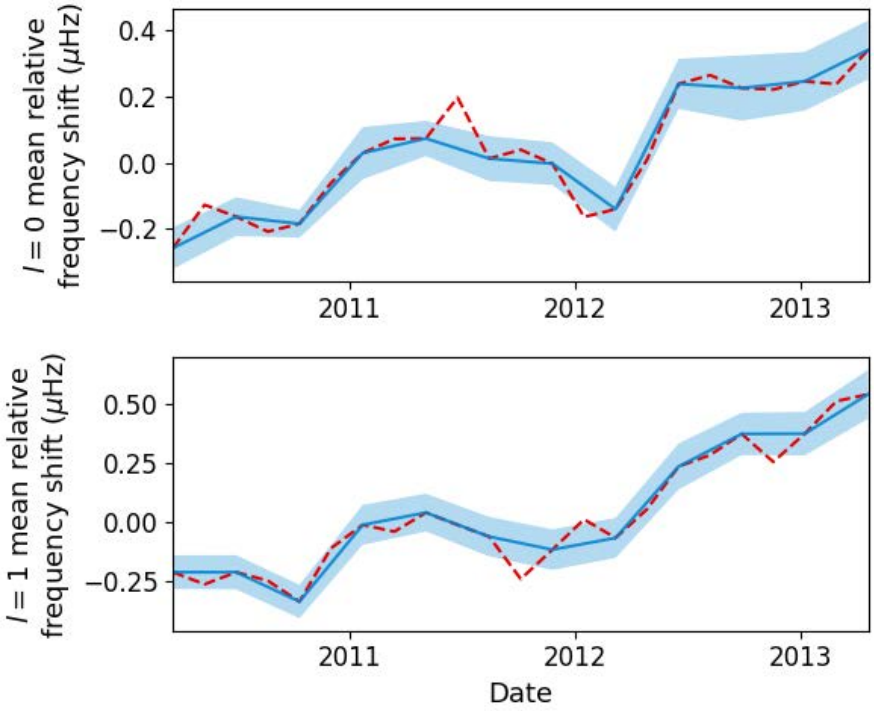


Figure 4.1: Mean frequency shifts from the 90-day data on HD 173701, averaged over the six strongest orders at $l = 0$ and $l = 1$, respectively. All available (and overlapping) data are shown in dashed red and overlaid in solid blue are the independent data sets with no overlap.

Our methods (described in Section 4.4) were validated by application to low-l solar p mode frequencies measured by the Birmingham Solar-Oscillations Network (BiSON) (Chaplin et al., 1996; Hale et al., 2016). BiSON is comprised of six ground based telescopes that observe the Sun as if it were a star in Doppler velocity. Mode frequencies were extracted from 108-day timeseries covering solar cycle 23 using the peak-bagging approach outlined in Howe et al. (2015). This length of timeseries covers four solar rotations and so any signal arising from sun-spot modulation is smoothed over. The timeseries start at 36-day intervals, giving each timeseries a 72-day overlap with adjoining series. Our analysis to infer the active latitudes was carried out on a full solar cycle as well as on individual rising and falling sections. The chosen epochs are shown in Figure 4.2, which displays weighted mean frequency shifts averaged over the six orders at $l = 0$ and $l = 1$, respectively, centred on the frequency of highest amplitude (which for the Sun spans the frequency range ≈ 2750 to $3500\,\mu$Hz).

For additional method validation, low-l solar data from the VIRGO (Variability of solar IRradiance and Gravity Oscillations) SPM (sun photometers) on board the Solar and Heliospheric Observatory (SoHO) (Fröhlich et al., 1995; Frohlich et al., 1997) were also analysed. Frequencies were estimated from peak-bagging applied to the power spectra of 180-day lightcurves, each overlapping in time by 90 days (Santos et al., 2018). Timeseries of 180 days were used since frequencies from 90-day datasets were notably noisier. The VIRGO data cover approximately 12 years, and hence we were able to analyse one entire 11-year cycle in addition to the constituent rising and falling phases.

4.4 Methodology

We used two methods to constrain the active latitudes, both using estimates of individual mode frequencies from different epochs as inputs. The first method works directly with the ratios of shifts shown by non-radial and radial modes.

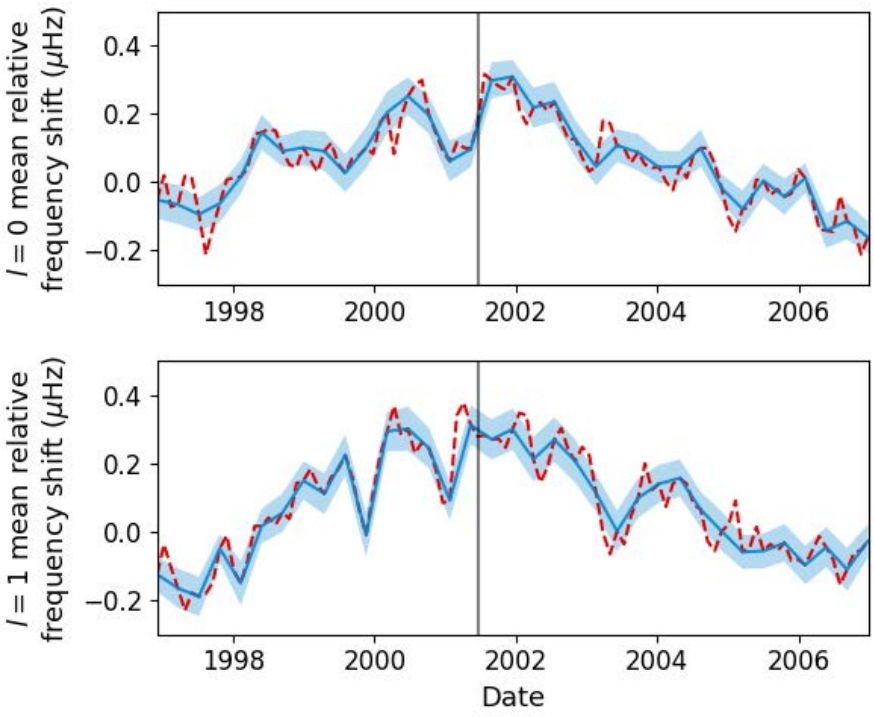


Figure 4.2: Mean frequency shifts from the 108-day BiSON data on the Sun, averaged over the six strongest orders at $l = 0$ and $l = 1$, respectively. Data for solar cycle 23 are shown in dashed red and overlaid in solid blue are the independent data sets used for the analyses. The vertical lines indicate the separate rising and falling phases used.

4.4.1 Method 1: Ratio of frequency shifts

From Equation 4.1 we know that observed frequencies will experience a shift due to the presence of a magnetic field. This can be described as a shift with respect to a field-free epoch such that at some epoch, t_j, the measured frequency of a mode having radial order n and degree l is:

$$\nu_{nl}(t_j) = \nu_{nl}(t_0) + \delta\nu_{nl}(t_j), \tag{4.5}$$

where frequencies for non-radial modes are the aforementioned weighted averages over the constituent m components (Equation 4.4), and $\nu_{nl}(t_0)$ are reference, field-free frequencies. Once extracted, the frequency shifts $\delta\nu_{nl}(t_j)$, can be used to constrain the latitudinal dependence. A credible and obvious concern with the above

is how one would in practice go about estimating a field-free frequency, i.e., even at solar minimum, significant amounts of widely distributed magnetic flux are still present in the near-surface layers, implying the mode frequencies will be offset from those expected for a notional field-free Sun. Provided the surface field architecture (i.e., the latitudinal dependence) does not change significantly over time, or instead the quality of the observations precludes us from detecting such changes (so in effect we see the same distribution, irrespective of epoch), the above concern is irrelevant and we can recover the sought-for latitudinal dependence. There is no need to measure shifts with respect to a field-free epoch and instead we may use any ratio of frequency differences, as we now go on to show.

By extracting the frequency shifts, the ratios at a given epoch of the shifts shown by modes of different (l, m) should provide constraints on the near-surface latitudinal dependence of activity. The assumption of a constant temporal latitudinal dependence means that the ratio of contemporaneous shifts at two different degrees is always the same, regardless of the epoch. Taking the example of the ratio between $l = 1$ and $l = 0$ modes of the same radial order at a particular epoch, t_j, we have

$$\eta = \frac{\delta\nu_{n1}(t_j)}{\delta\nu_{n0}(t_j)} = \frac{\delta\nu_{n1}(t_{j+1})}{\delta\nu_{n0}(t_{j+1})} = \frac{\delta\nu_{n1}(t_{j+j'})}{\delta\nu_{n0}(t_{j+j'})}, \tag{4.6}$$

where $t_{j+j'}$ is some epoch falling after t_j, and η is a constant describing the ratio since it is n-independent.

Now, the difference in frequency of a mode at two different epochs is:

$$\nu_{nl}(t_{j+1}) - \nu_{nl}(t_j) = \delta\nu_{nl}(t_{j+1}) - \delta\nu_{nl}(t_j), \tag{4.7}$$

The ratio of differences in the measured frequencies, and also in the corresponding shifts, is:

$$\Phi_{10}(n, t_j) = \frac{\nu_{n1}(t_{j+1}) - \nu_{n1}(t_j)}{\nu_{n0}(t_{j+1}) - \nu_{n0}(t_j)} = \frac{\delta\nu_{n1}(t_{j+1}) - \delta\nu_{n1}(t_j)}{\delta\nu_{n0}(t_{j+1}) - \delta\nu_{n0}(t_j)}. \tag{4.8}$$

We can rearrange Equation 4.8 as follows to obtain η:

$$\begin{aligned} \Phi_{10}(n,t_j)\delta\nu_{n0}(t_{j+1}) - \Phi_{10}(n,t_j)\delta\nu_{n0}(t_j) \\ = \delta\nu_{n1}(t_{j+1}) - \delta\nu_{n1}(t_j), \\ = \eta(\delta\nu_{n0}(t_{j+1}) - \delta\nu_{n0}(t_j)). \\ \Rightarrow \eta = \frac{\Phi_{10}(n,t_j)(\delta\nu_{n0}(t_{j+1}) - \delta\nu_{n0}(t_j))}{\delta\nu_{n0}(t_{j+1}) - \delta\nu_{n0}(t_j)} \equiv \Phi_{10}(n,t_j). \end{aligned} \tag{4.9}$$

This shows that by taking any ratio of frequency differences we need not know the shift with respects to a field-free frequency.

Another advantage of taking ratios of modes of a particular n at any epoch can be seen by re-writing the basic equation for η with the contributing shifts of the different $l = 1$ mode components $\delta\nu_{nlm}$ included explicitly, i.e.,

$$\eta = \frac{\delta\nu_{n1}(t_j)}{\delta\nu_{n0}(t_j)} = \frac{\alpha\delta\nu_{n10}(t_j) + (1-\alpha)\nu_{nl|1|}}{\delta\nu_{n00}(t_j)}, \tag{4.10}$$

where α has the same meaning as in Equation 4.4. Using Equation 4.1 and by assuming a given model for the magnetic field (Equation 4.2) we again see that the field strength cancels leaving the ratio to depend only on the geometry of the near-surface activity. Hence the ratios can be used to place constraints on $\lambda_{\min}$ and $\lambda_{\max}$.

Finally, we also assume that the shifts do not change with radial order (frequency). Solar frequency shifts do increase in size at higher orders. However, provided we limit ourselves to using only a few orders, the typical precision in the data means the above is a reasonable approximation. The frequency dependence of the shifts for HD 173701 has been shown to be comparable to the Sun (Salabert et al., 2018). We tested the impact of including in our model a power law relation in frequency to describe the shifts, as given by Chaplin et al. (2001). We found that adopting a range of reasonable values for the exponent made little impact to the resulting latitude estimation.

Here, we use data on six radial orders for our analysis, straddling the order showing the largest measured mode amplitudes.

Calculating observed η

Frequency shifts were calculated by taking differences between observed frequencies separated by k epochs, i.e.,

$$\delta\nu_{nl}(t_j) = \nu_{nl}(t_{j+k}) - \nu_{nl}(t_j). \tag{4.11}$$

To optimize the signal-to-noise ratio in the computed shifts, we chose $k = 9$ for HD 173701, which is equivalent to epochs separated by 9×45 days. We note that three years of data on HD 173701, from the total 7.4 year activity cycle (Karoff et al., 2018) is equivalent to 4.4 years from the 11 year solar cycle of BiSON data. For BiSON's 108-day data this equated to $k = 17$ epochs taken every 36 days. For the 180-day observations of VIRGO, data sets overlapped by 90 days, therefore the corresponding number of epochs was $k = 4$.

The ratio between $l = 1$ and $l = 0$ shifts of radial order n, calculated over an epoch difference j is given by

$$\eta_{\mathrm{obs,nj}} = \frac{\delta\nu_{n1,j}}{\delta\nu_{n0,j}}, \tag{4.12}$$

and is evaluated between each pair of modes of the same radial order for all overtones and over all pairs of epochs. Since ratios of small numbers can be noisy with the available computational precision we perform calculations in logarithmic space, i.e.

$$\ln(\eta_{\mathrm{obs,nj}}) = \ln(\delta\nu_{n1,j}) - \ln(\delta\nu_{n0,j}). \tag{4.13}$$

Uncertainties on this logarithmic ratio, $\sigma\ln(\eta_{\mathrm{obs,nj}})$, were propagated from observational uncertainties on the individual frequencies used to construct each shift.

To find a time- and order- averaged observed ratio $\langle\eta_{\mathrm{obs}}\rangle$, we compute the weighted average over all ratios and epochs. The first sum is over all ratios at

a given epoch, i.e.,

$$\ln(\eta_{\rm obs,j}) = \left(\sum_{n=1}^{N_{\rm ord}} \frac{\ln(\eta_{\rm obs,nj})}{\sigma \ln(\eta_{\rm obs,nj})^2} \right) / \left(\sum_{n=1}^{N_{\rm ord}} \frac{1}{\sigma \ln(\eta_{\rm obs,nj})^2} \right), \tag{4.14}$$

with an uncertainty given by

$$\sigma \ln(\eta_{\rm obs,j}) = \left(\sum_{n=1}^{N_{\rm ord}} \frac{1}{\sigma \ln(\eta_{\rm obs,nj})^2} \right)^{-1/2}, \tag{4.15}$$

where $N_{\rm ord} = 6$ is the total number of orders used. The second sum is then made over all epochs, $N_{\rm e}$, i.e.,

$$\ln(\langle \eta_{\rm obs} \rangle) = \left(\sum_{j=1}^{N_{\rm e}} \frac{\ln(\eta_{\rm obs,j})}{\sigma \ln(\eta_{\rm obs,j})^2} \right) / \left(\sum_{j=1}^{N_{\rm e}} \frac{1}{\sigma \ln(\eta_{\rm obs,j})^2} \right), \tag{4.16}$$

with an uncertainty given by

$$\sigma \ln(\langle \eta_{\rm obs} \rangle) = \left(\sum_{j=1}^{N_{\rm e}} \frac{1}{\sigma \ln(\eta_{\rm obs,j})^2} \right)^{-1/2}. \tag{4.17}$$

The resulting $\langle \eta_{\rm obs} \rangle$ depends on the latitudinal profile of the near-surface magnetic activity.

Constraining the minimum and maximum latitudes

In order to interpret the observed $\langle \eta_{\rm obs} \rangle$ we need to identify the model that best matches the observations. From our description of the frequency shifts given by Equation 4.1, it is clear that $\langle \eta_{\rm obs} \rangle$ will depend on the magnetic field geometry and hence on $\lambda_{\rm min}$ and $\lambda_{\rm max}$, but not on the field strength since it cancels. A grid of models was constructed using combinations of the parameters $\lambda_{\rm min}$ and $\lambda_{\rm max}$ each in the range zero to 90 degrees in increments of 0.5 degrees. Modelled ratios, $\eta_{\rm mod}$, were evaluated using Equation 4.10 where α was deduced from the known inclination angle and shifts were calculated using Equation 4.1 and $\lambda_{\rm min}$ and $\lambda_{\rm max}$. The best-

fitting model ratio $\eta_{\rm mod}$ to the observed value $\langle\eta_{\rm obs}\rangle$ was identified from the grid of modelled ratios as that which gave the lowest chi-squared statistic.

To build a probability distribution of inferred latitudes we sampled from a distribution that is representative of the observations and iterated multiple times. The measurement of $\langle\eta_{\rm obs}\rangle$ and its associated uncertainty, $\langle\sigma\eta_{\rm obs}\rangle$, were used to randomly sample from a Gaussian distribution of $\mathcal{N}(\langle\eta_{\rm obs}\rangle, \sigma\langle\eta_{\rm obs}\rangle)$. From this we produced sample ratios $\eta'_{\rm obs}$ given by

$$\eta'_{\rm obs} = \langle\eta_{\rm obs}\rangle + n_i\,\sigma\langle\eta_{\rm obs}\rangle\,, \tag{4.18}$$

where n_i is a random number from a Gaussian distribution with $\mathcal{N}(0,1)$. To incorporate the uncertainty on inclination angle, σi, we again sampled from a representative distribution, $\mathcal{N}(i,\sigma i)$, to produce samples i' which were used to calculate each new grid of ratios, $\eta'_{\rm mod}$, i.e.

$$\eta'_{\rm mod} = \frac{\alpha'\delta\nu_{n10} + (1-\alpha')\delta\nu_{n1|1|}}{\delta\nu_{n00}}, \tag{4.19}$$

where α' is calculated using i' (see Appendix A).

For each iteration, where a new sample of $\eta'_{\rm obs}$ was taken and a new grid, $\eta'_{\rm mod}$, was calculated, we minimised χ^2 to find the best fitting $\lambda_{\rm min}$ and $\lambda_{\rm max}$, where

$$\chi^2 = \frac{(\eta'_{\rm obs} - \eta'_{\rm mod})^2}{\eta'_{\rm mod}}. \tag{4.20}$$

Posterior distributions in $\lambda_{\rm min}$ and $\lambda_{\rm max}$ were made after performing 15,000 iterations.

4.4.2 Method 2: Forward modelling approach

To ensure the robustness of our results we also applied a second method comprising a forward modelling approach.

The model

We directly modelled the frequencies of $l = 0$ and $l = 1$ modes observed over different epochs. The field-free frequency was therefore included in the model but handled as a nuisance parameter, as described below. We assumed that there was no frequency dependence in the shifts, as discussed in Section 4.3.

The frequency shifts in the model are defined by

$$\delta\nu_{lm}(t_j) = \beta(t_j) \left(l + \frac{1}{2}\right) K_{lm}, \tag{4.21}$$

where the $\beta(t_j)$ capture the time dependence of the shifts (see below), and

$$K_{lm} = \frac{(l - |m|)!}{(l + |m|)!} \int_{\theta_{\min}}^{\theta_{\max}} |P_l^{|m|}(\cos\theta)|^2 \sin\theta d\theta. \tag{4.22}$$

As a reminder, θ is the co-latitude given by $\theta = \left(\frac{\pi}{2} - \lambda\right)$.

As discussed in Section 4.2, we extract what are in effect centroid frequencies for each non-radial mode rather than frequencies for the individual m components. The above equations, when combined with Equation 4.5, then become

$$\nu_{nl}(t_j) = \nu_{nl}(t_0) + \beta(t_j) \left(l + \frac{1}{2}\right) K_l, \tag{4.23}$$

and

$$K_l = \alpha K_{l0} + (1 - \alpha) K_{l|1|}, \tag{4.24}$$

describes the contributions of the different m-components to the centroid frequency (using α described in Appendix A).

We re-parametrise so that $\beta(t_j) = \beta_j$ is used to describe the shifts of the $l = 0$ and $l = 1$ modes from their field-free frequencies and $I(\theta_{\min}, \theta_{\max}) = K_1/K_0$ acts as a correction factor to incorporate the latitudinal dependence of magnetic field which will only affect non-radial modes. We define $\theta_{\min} = \left(\frac{\pi}{2} - \lambda_{\min}\right)$ and $\theta_{\max} =$

$\left(\frac{\pi}{2} - \lambda_{\max}\right)$. The result is that we model radial and dipole frequencies separately, but with shared parameters, as described by

$$\begin{aligned} \nu_{n0,j} &= \nu_{n0,0} + \frac{1}{2}\beta_j, \\ \nu_{n1,j} &= \nu_{n1,0} + \frac{3}{2}\beta_j I(\theta_{\min}, \theta_{\max}), \end{aligned} \tag{4.25}$$

and

$$I(\theta_{\min}, \theta_{\max}) = \frac{\alpha A_0 + (1-\alpha) A_1}{\cos(\theta_{\min}) - \cos(\theta_{\max})}, \tag{4.26}$$

where

$$\begin{aligned} A_0 &= \frac{1}{3}\left[\cos^3(\theta_{\min}) - \cos^3(\theta_{\max})\right], \\ A_1 &= \frac{1}{24}\left[\cos^3(\theta_{\max}) - 9\cos(\theta_{\max}) - \cos^3(\theta_{\min}) + 9\cos(\theta_{\min})\right]. \end{aligned} \tag{4.27}$$

Fitting procedure

The model was fitted simultaneously to $N_{\rm ord}$ radial and dipole orders over $N_{\rm e}$ epochs, with the same two latitudes for all frequencies resulting in $(2N_{\rm ord} + 2 + N_{\rm e})$ free parameters. Parameters were estimated using a sampling approach where the posterior probability distributions for parameters $\bar{\theta}$, are mapped assuming prior information, J, and given data, D, using Bayes' theorem:

$$p(\bar{\theta}|D, J) \propto p(D|\bar{\theta}, J)\, p(\bar{\theta}|J). \tag{4.28}$$

Here $p(\bar{\theta}|J)$ is the prior, and $p(D|\bar{\theta}, J)$ is the likelihood given by

$$p(D|\bar{\theta}, J) = \prod_{nl}\left[\left(2\pi\sigma^2_{{\rm obs},nl}\right)^{-\frac{1}{2}} \exp\left(\frac{-(\nu_{{\rm pred},nl} - \nu_{{\rm obs},nl})^2}{2\sigma^2_{{\rm obs},nl}}\right)\right], \tag{4.29}$$

where $\nu_{{\rm obs},nl}$ and $\sigma_{{\rm obs},nl}$ are the observed frequency and associated uncertainty for a particular mode of order, n, and angular degree, l, and $\nu_{{\rm pred},nl}$ is the modelled frequency of that mode.

Uniform priors were used, with $-5\% \leq \nu_{nl,0} \leq +5\%$, $0 \leq \beta_j \leq 10$ and $0 \leq \lambda_{\mathrm{min,max}} \leq \frac{\pi}{2}$. The location parameters were constrained so that $\lambda_{\mathrm{min}} < \lambda_{\mathrm{max}}$, and so that consecutive radial field-free frequencies were separated by the average spacing $\Delta\nu$ between consecutive overtones of the same degree, l, with a small tolerance of $\pm 2\,\mu$Hz. Walkers were first made to explore the entire prior space with uniform likelihood so that their starting positions were well spread before fitting our model to the data using the likelihood in Equation 4.29.

Both the β_j and the field-free frequency parameters control the absolute amount of frequency shift, whereas the relative shifts of radial and dipole modes depend on the latitudinal distribution of activity. Since we seek to constrain our determination of the active band latitudes, the actual values of β_j are uninformative and can be marginalised over since we require just the changes in β_j over time. The β_j parameter corresponding to the first epoch could therefore be constrained using a Gaussian prior, $\mathcal{N}(2., 0.01)$, to speed up convergence. Chosen different values for the first β_j made no difference to the resulting latitude predictions.

The fit was made using an affine invariant ensemble sampling algorithm (see Goodman & Weare, 2010) implemented by `emcee` in Python (Foreman-Mackey et al., 2013). Convergence of the chains was checked and the autocorrelation time was calculated to be considerably smaller ($\sim$500 times smaller) than our effective number of samples meaning that the chains were independent of their starting position. Figure 4.3 shows an example corner plot (Foreman-Mackey, 2016) after convergence with both joint and marginalised posterior distributions for a subset of parameters fitted to the rising phase of the BiSON dataset. Note that due to the prior $\lambda_{\mathrm{min}} < \lambda_{\mathrm{max}}$ there is a diagonal cut off for the joint posterior of the latitudes. Also visible in Figure 4.3 is the limit preventing λ_{min} from becoming negative. As discussed above, both the field-free frequencies and β_j parameters were uninformative and were therefore treated as nuisance parameters.

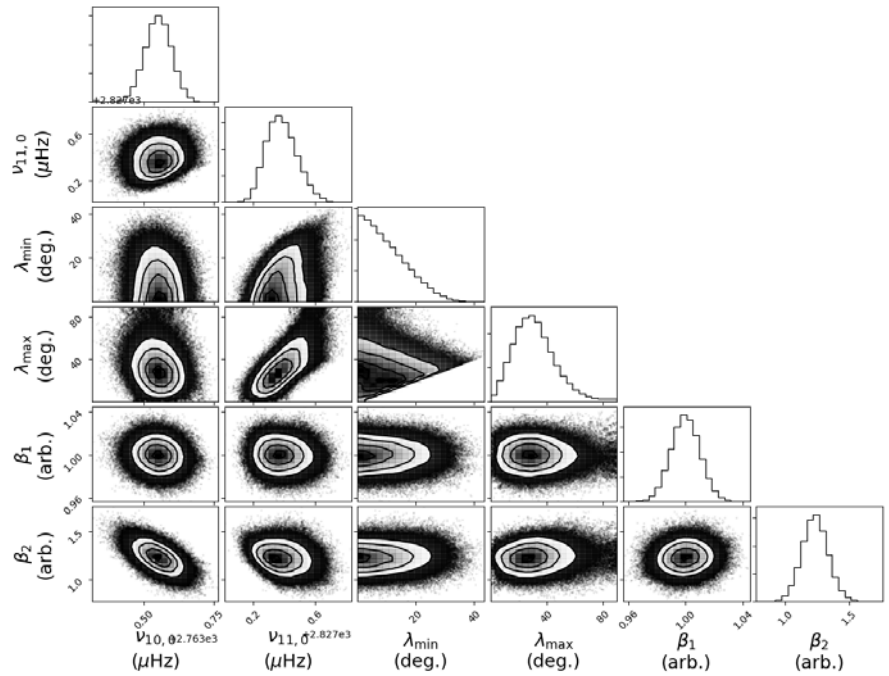


Figure 4.3: A corner plot showing the joint and marginalised posterior probability distributions for a subset of model parameters fitted to the rising section of the BiSON dataset.

4.5 Results

4.5.1 The Sun

We applied both methods to the BiSON and VIRGO data. Best-fitting latitudes were obtained using the statistical modes of the marginalised posterior distributions, since they were found to be the best summary statistics based on tests with artificial data, i.e., by investigating the response of our top hat model to more realistic-shaped activity distributions with softer boundaries (see below). Reported uncertainties on the inferred latitudes correspond to the 68 percent credible intervals either side of the mode.

The first method gave estimated minimum latitudes for individual sections of the solar cycle (either rising or falling) of $\sim$10°and maximum values of $\sim$40°. The latitudes for the BiSON data set covering the entire solar cycle were estimated to be $\lambda_{\rm min} = (5.8^{+13.7}_{-3.8})$ degrees and $\lambda_{\rm max} = (42.9^{+17.1}_{-9.9})$ degrees using Method 1; and

$\lambda_{\rm min} = (3.3^{+16.6}_{-16.6})$ degrees and $\lambda_{\rm max} = (40.6^{+18.3}_{-9.4})$ degrees using Method 2. The resulting posterior distributions from the entire observation using each method are shown in Figures 4.4 and 4.5. We obtained similar results using the VIRGO data and other rising and falling phases from BiSON data in addition to those shown in Figure 4.2. For the statistic on $\lambda_{\rm min}$ from Method 2 we chose to report a symmetric error from the positive-side 68 percent credible interval since the posterior distribution peaks at very low latitudes and is therefore symmetric but $\lambda_{\rm min} \geq 0$. This can be seen in Figure 4.3 where the posteriors for $\lambda_{\rm min}$ for the Sun are half rather than full ellipses.

In order to measure significant frequency shifts with respect to their uncertainties it was necessary to use frequency differences separated over long periods of time (whole rising or falling sections of the activity cycle). Consequently, the resulting latitudes are an integration over that entire duration. For this reason we cannot expect to detect any migration of activity band latitude during the solar cycle. To be able to detect cyclic variations, frequency shifts would need to be measured over shorter stretches of time which would require either smaller frequency uncertainties and therefore better signal-to-noise levels in the individual power spectra, or larger frequency shifts.

Our results, whilst reassuringly consistent, are notable for the fact that they appear to straddle a wider band in latitude than is usually associated with sunspots. Histograms of spot latitudes shown over recent solar cycles do actually extend beyond 40 degrees (e.g. Solanki et al., 2008; Mandal et al., 2017). However, it is also important to note that our method responds to the range in latitude over which there is significant magnetic flux over and above the weak basal ephemeral flux levels that are present throughout the cycle, as we shall now go on to explain. Chaplin et al. (2007b) studied the correlation of frequency shifts to proxies of activity that have a different balance of strong- and weaker-field flux, and found that the results implied the frequencies were responding to a mix of strong and weaker-component flux. In addition, Santos et al. (2016) found that only around 30% of frequency shifts

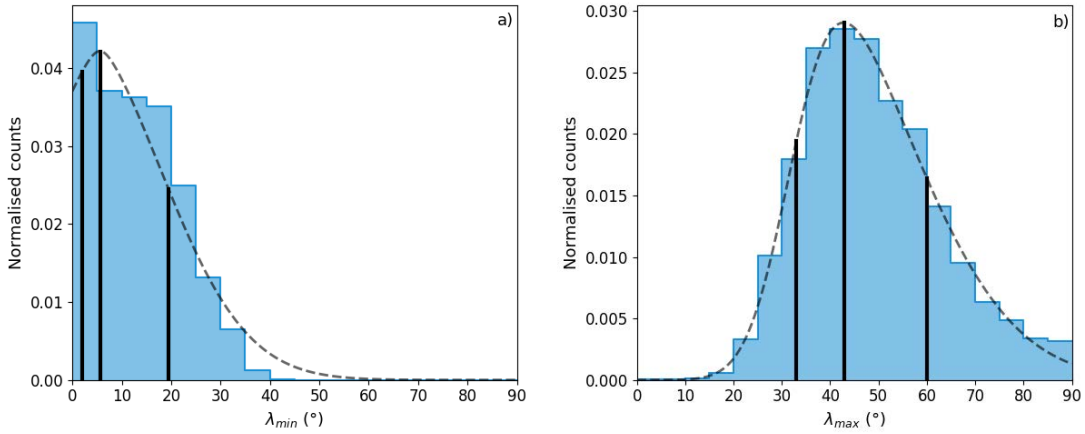


Figure 4.4: Posterior distributions of (a) minimum- and (b) maximum-latitudes for the entire 108-day BiSON dataset from method 1. The curves are fitted skewed Gaussians and the vertical lines indicate the mode and 68 % confidence intervals.

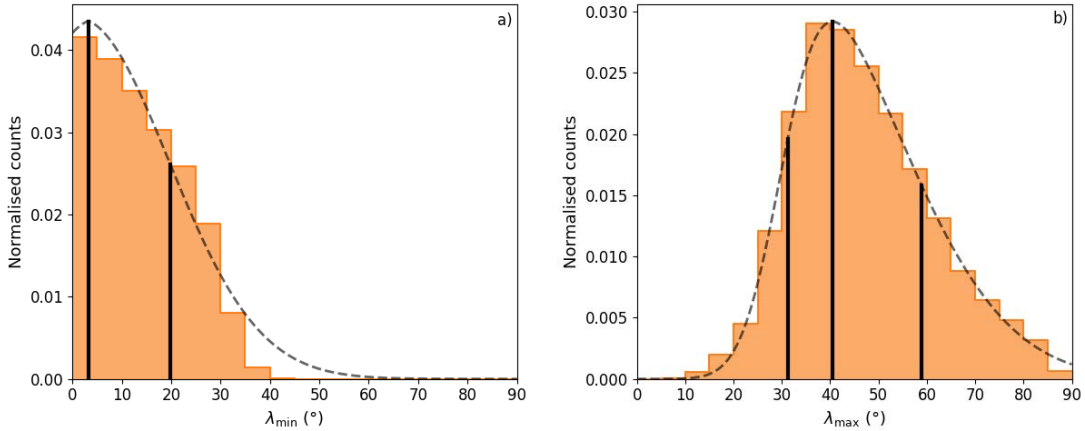


Figure 4.5: Posterior distributions of (a) minimum- and (b) maximum-latitudes for the entire 108-day BiSON data from method 2. The curves are fitted skewed Gaussians and the vertical lines indicate the mode and 68 % confidence intervals.

seem to be directly associated to sunspots, the remaining being likely associated to more global variations, possibly resulting from changes in the overall magnetic field. This supports the idea that the magnetic influence on the frequencies needs not to be restricted to the latitudes where spots are observed.

Figure 4.6 shows unsigned solar magnetic field strength data from synoptic charts generated by the National Solar Observatory (Howe et al., 2017). The solid lines in the top two panels show normalised sums, made over the same two epochs as in our study, of the magnetic flux at different latitudes. These epochs cover the

rising and falling phases of cycle 23. The bottom panel shows the sum over both epochs. It is clear from these plots that significant flux extends beyond a latitude of 40 degrees. As expected, the distributions peak at higher latitudes during the rising phases, since the latitudes at which spots are found moves towards the equator as cycles progress.

Also shown on each plot (dotted lines) is the equivalent latitudinal dependence of frequency shifts shown by medium-l solar p modes during each epoch (Howe et al., 2002). These frequencies come from highly spatially resolved observations of the Sun made by the GONG network and, as noted previously, from them it is possible to make inferences on the latitudinal response of the modes to levels of detail that are not possible with the low-l modes only. As expected, the distributions for the mode frequency shifts follow very closely those of the magnetic flux. We therefore used the flux distributions to generate realistic artificial p mode data to test our latitudinal inference using the low-l frequencies only. Artificial frequencies were made to mimic six orders of BiSON data straddling the centre of the oscillations spectrum, with $i = 90 \pm 4$ degrees and shifts for different mode components generated using the latitudinal flux distribution integrated over cycle 23. The modelled frequencies covered $N_{\mathrm{e}} = 12$ and had frequency uncertainties of $\sigma\nu_{\mathrm{obs}} = 0.15\,\mu$Hz. Frequencies were scattered using a normal distribution with standard deviation equal to the frequency uncertainty. Our top-hat model was then fitted using the previously described methods.

Figure 4.7 shows the response of our model to the artificial solar data, showing the median and modal best-fitting latitudes. It is clear that the boundaries of our top-hat model extend to the tails of the latitudinal distribution where the magnetic flux has fallen to ephemeral levels. These tests also show that the modal values better represent the distributions since this test statistic encompasses a much larger proportion of the significant flux.

We verified that our results are robust to using different initial parameter guesses.

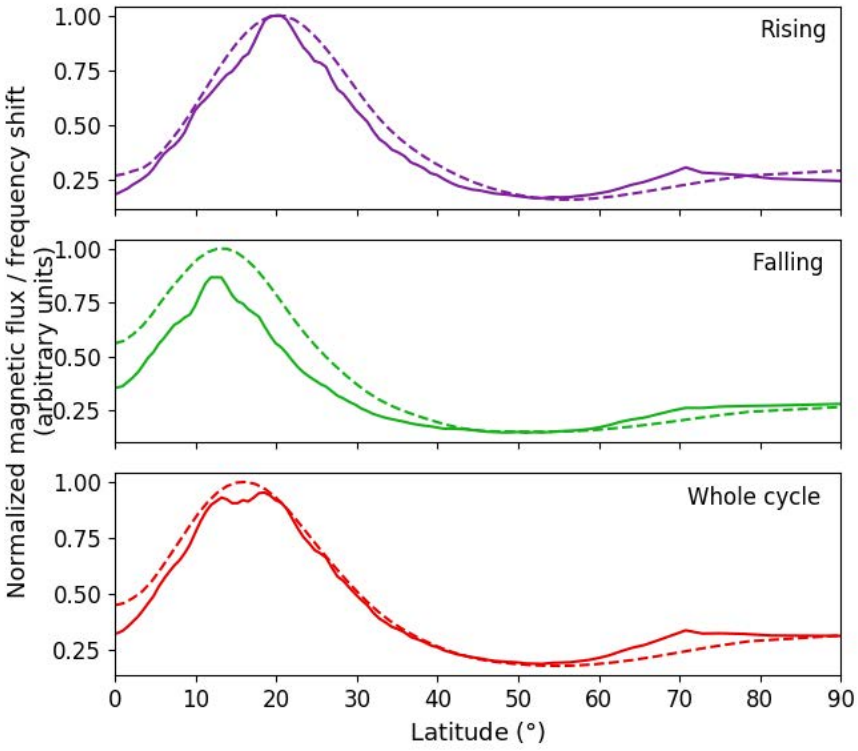


Figure 4.6: Distribution of mean Carrington rotation averages of magnetic field strength from NSO (solid lines) and frequency shifts from GONG (dashed lines), which have been integrated over rising and falling sections of the solar cycle as defined earlier in the text.

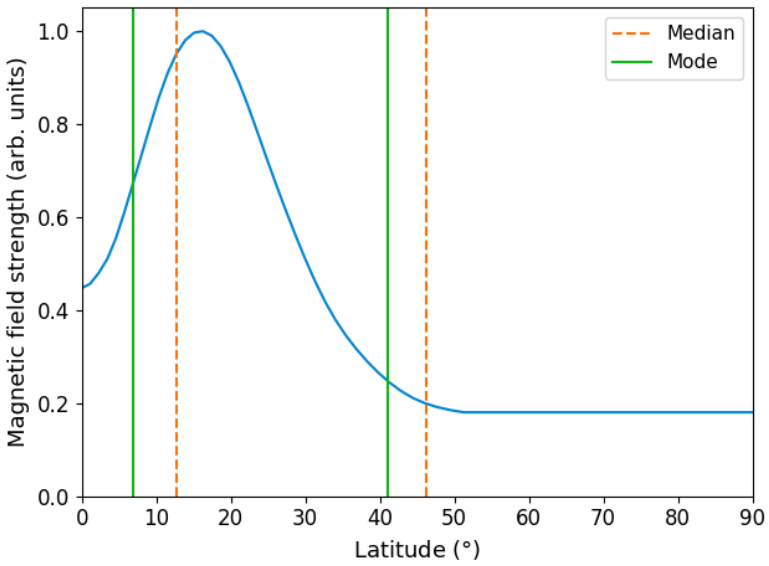


Figure 4.7: In blue shows the magnetic field distribution used to generate artificial frequency data based on GONG measurements of frequency shifts of the Sun integrated over two solar cycles. Dashed red and solid green lines show the median and modal values, respectively, of $\lambda_{\min}$ and $\lambda_{\max}$ from fitting our top hat-model.

We also tested artificial data with underlying combinations of $\lambda_{\rm min}$ and $\lambda_{\rm max}$ which covered a range of angles and widths for the active latitude bands, and found that we could recover the expected latitudes given data of the precision available on the Sun and HD 173701. That the method does not recover the correct *absolute* values of the field-free frequencies and β_j coefficients is not a cause for concern. These parameters are highly anti-correlated and we do recover the correct *change* in the β_j with time, which allows for robust inference on the latitudes.

4.5.2 HD 173701

Both methods were applied to the 90-day *Kepler* datasets on HD 173701. The posterior distributions on the latitudes given by the first and second methods are shown in Figures 4.8 and 4.9 respectively. The inferred latitudes for HD 173701 are $\lambda_{\rm min} = (34.5^{+11.5}_{-16.5})$ degrees and $\lambda_{\rm max} = (69.6^{+13.9}_{-17.1})$ degrees from Method 1; and $\lambda_{\rm min} = (24.9^{+12.0}_{-12.2})$ degrees and $\lambda_{\rm max} = (68.8^{+14.2}_{-17.8})$ degrees from Method 2. As with the solar results, the two separate analyses produced consistent conclusions for HD 173701.

In order to make comparisons with the Sun, a section of BiSON data was chosen to span the same proportion of the star's activity cycle. This corresponded to 4.4 years of BiSON data, since the cycle of HD 173701 is 7.4 years, of which we have observed 3.0 years. We know that the amplitudes of the frequency shifts in HD173701 are larger than those in the Sun and so it should be easier for the model to predict the latitudes of activity given that the power-spectra signal-to-noise ratios are similar.

It is clear from comparison between the latitude posterior distributions using the equivalent duration from BiSON (shown in grey) that the spread of activity on HD 173701 is quite different to that on the Sun and is more widely distributed, i.e. it extends to much higher latitudes. From the application of a Kolmogorov-Smirnov test (Kolmogorov, 1933; Smirnov, 1939), we may conclude that the probability is vanishingly small that the $\lambda_{\rm min,max}$ posteriors for the Sun and HD 173701 are drawn

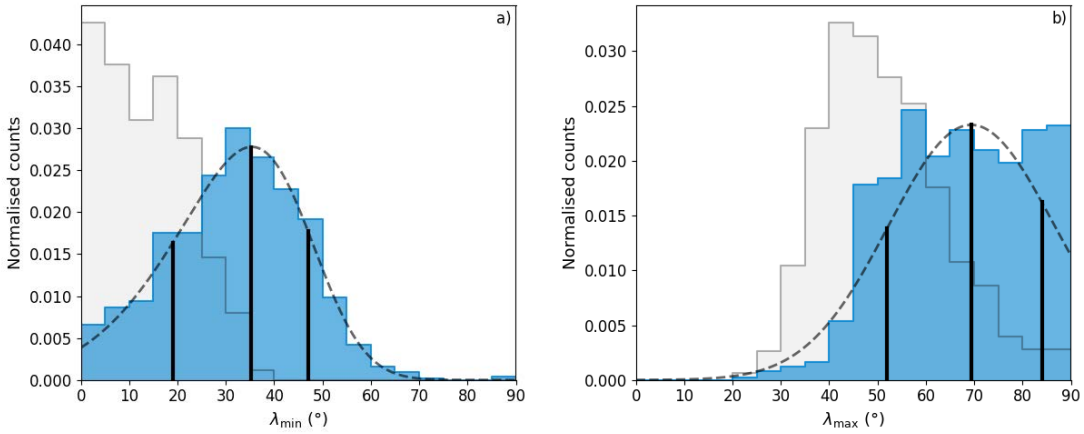


Figure 4.8: Posterior distributions of (a) minimum- and (b) maximum-latitudes for HD 173701 from method 1. In faded grey is the distribution for BiSON analysis over an equivalent duration of the solar cycle (4.4 years) for comparison. The curves are fitted skewed Gaussians and the vertical lines indicate the mode and 68 % confidence intervals.

from the same underlying distribution.

We again performed tests with artificial data, this time with simulated frequencies made from an underlying model of the magnetic flux similar to that shown in Figure 4.7 but peaked at higher degrees (around 55 degrees). We again found that our results had sensitivity extending to the tails of the modelled magnetic field distribution.

4.6 Conclusions

We have presented a new approach to constraining active latitudes on solar-type stars, which uses stellar-cycle-induced frequency shifts of detectable solar-like oscillations. These oscillations may be described by spherical harmonic functions of angular degree l and azimuthal order m. When near-surface activity is distributed in a non-homogeneous manner, modes of different (l,m) will have their frequencies shifted by different amounts. By adopting a simple model for the spatial distribution of the near-surface activity, we show that data on the frequency shifts may be used to infer minimum and maximum latitudes for the active bands, $\lambda_{\rm min}$ and $\lambda_{\rm max}$, respectively.

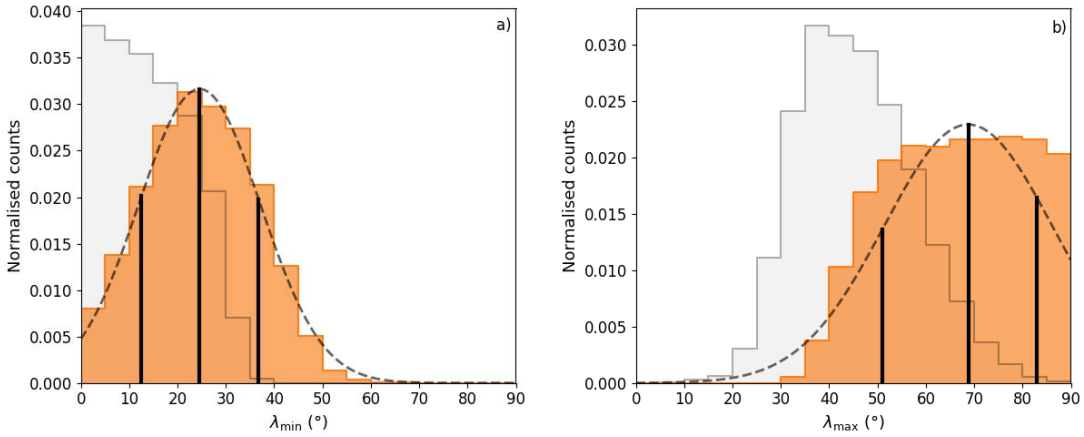


Figure 4.9: Posterior distributions of (a) minimum- and (b) maximum-latitudes for HD 173701 from method 2. In faded grey is the distribution for BiSON analysis over an equivalent duration of the solar cycle (4.4 years) for comparison. The curves are fitted skewed Gaussians and the vertical lines indicate the mode and 68 % confidence intervals.

We presented two complementary methods, which we tested on Sun-as-a-star helioseismic data and artificial data. From BiSON data covering solar cycle 23, we found $\lambda_{\rm min} = (5.8^{+13.7}_{-3.8})$ degrees and $\lambda_{\rm max} = (42.9^{+17.1}_{-9.9})$ degrees using our first method; and $\lambda_{\rm min} = (3.3^{+16.6}_{-16.6})$ degrees and $\lambda_{\rm max} = (40.6^{+18.3}_{-9.4})$ degrees using our second method. We obtained similar results using data from the VIRGO/SPM instrument on board the ESA/NASA SoHO satellite. At first glance, these solar results appear to straddle a wider band in latitude than is usually associated with sunspots. However, histograms of spot latitudes do extend beyond 40 degrees over recent cycles. Moreover, we also showed that our method responds to the range in latitude over which there is significant magnetic flux over and above the weak basal ephemeral flux levels that are present throughout the cycle.

We then applied our methods to just over 3 years of *Kepler* data on the solar analogue HD 173701. This star shows a 7.4-year activity cycle, and stronger frequency shifts than the Sun. We found $\lambda_{\rm min} = (34.5^{+11.5}_{-16.5})$ degrees and $\lambda_{\rm max} = (69.6^{+13.9}_{-17.1})$ degrees from our first method; and $\lambda_{\rm min} = (24.9^{+12.0}_{-12.2})$ degrees and $\lambda_{\rm max} = (68.8^{+14.2}_{-17.8})$ degrees from our second method. Comparing the peaks of the posterior distributions it is clear that the active bands on HD 173701 straddle a much wider

range in latitude than do the bands on the Sun. From our inferred posteriors, we find that there is a vanishingly small chance that the results on HD 173701 and the Sun describe the same underlying distribution.

We may compare our findings with those from spot modulation signatures found in *Kepler* data on this star. Karoff et al. (2018) inferred spot latitudes ranging from a few up to 40 degrees, using the relative strengths of different rotational harmonics of the spot modulation signal. Bazot et al. (2018) combined the spot modulation signal with asteroseismic constraints on latitudinal differential rotation given by the frequency splittings of non-radial p modes, and uncovered similar latitudes to Karoff et al. (2018), but with results plotted over time.

That these two studies find similar results is perhaps not surprising, given that both use the same spot modulation signal. Such results are of course sensitive to the latitudes of discrete, strong-field structures, i.e., spots. A possible explanation for why our inference on the activity extends over a much wider latitude range is that our results are recording the presence of significant amounts of more widely distributed, weaker-component flux, lying outside spots.

4.7 Appendices

Appendix A: Frequency dependence on inclination angle

As discussed in Section 4.2, we must in practice work with non-radial mode frequencies which represent some suitably weighted average of the frequencies of the constituent m components.

The relative visibilities of the individual azimuthal components of a mode depend on the angle of inclination shown by the star. In the case where it is difficult to place tight constraints on the frequencies of the individual m components – e.g., when shorter datasets are needed to follow stellar-cycle changes in the frequencies over time – it is necessary to understand the relative contribution of the constituent m to the central frequency of the multiplet. Here, we place constraints on those

contributions for $l = 1$ modes.

The individual components of a non-radial mode are split by rotation. Moreover, because the activity-induced shift at each (l,$|m|$) is likely to differ in size, the result is a multiplet in which the constituent frequencies are arranged asymmetrically. At $l = 1$, the component frequencies are related by

$$\nu_{10} = \nu_0,$$

$$\nu_{11} = \nu_0 + \nu_{\rm s},$$

$$\nu_{1-1} = \nu_0 - \nu_{\rm s} + 2a.$$

where ν_0 is the frequency all components would have in the absence of rotation or magnetic fields, $\nu_{\rm s}$ is the rotational splitting, and a is the frequency asymmetry.

Figure 4.10 shows a representation of the limit profile in frequency for a multiplet viewed at $i = 45°$, having Lorentzian peak linewidths of 1μHz, a rotational splitting of $\nu_{\rm s} = 0.4\,\mu$Hz and an asymmetry of $a = 0.25\,\mu$Hz. The mode parameters are typical of a prominent mode in the solar oscillations spectrum for high levels of activity; these parameters are also representative of HD 173701, whilst the angle chosen for the plot is similar to that extracted on HD 173701. The two $|m| = 1$ components, shown by the dashed lines, have been shifted from the positions they would have occupied in the absence of the near-surface magnetic field (as shown by the grey vertical lines). The overall result is a slightly distorted multiplet whose centroid has been shifted.

Artificial data were generated for $l = 1$ modes comprising three azimuthal components, plus background, with the combined power spectral density described by

$$M(\nu;\nu_{lm}) = \sum_{m=-l}^{l} \epsilon_{lm}(i)\, M_L(\nu;\nu_{lm}) + B, \tag{4.30}$$

where the component frequencies are ν_{lm}, $\epsilon_{lm}(i)$ is the visibility, B the background,

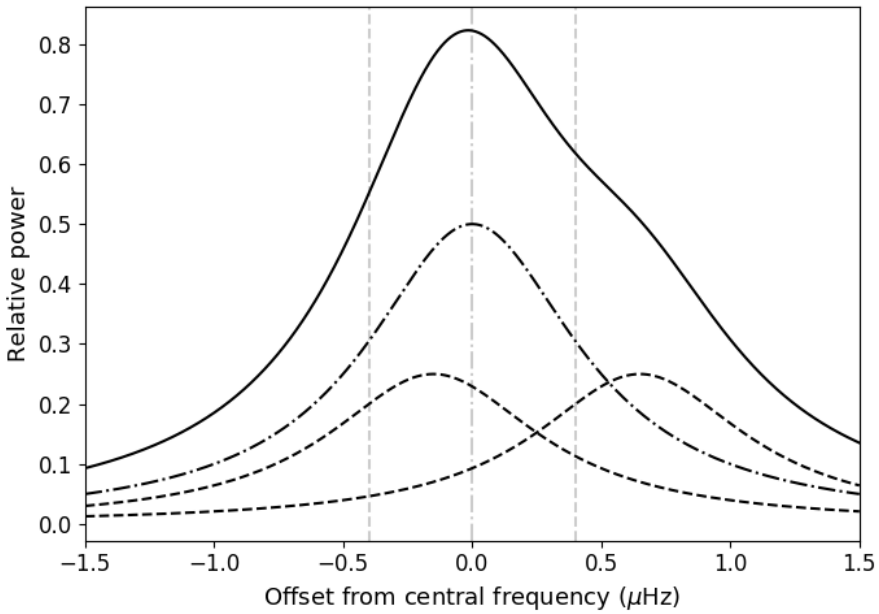


Figure 4.10: Multiplet peak profile of an $l = 1$ mode situated at $i = 45°$. The dot-dashed curve indicates $m = 0$, dashed curve $|m| = 1$, the solid curve the combined power, and vertical lines with corresponding line style indicate the positions of the individual azimuthal components in the absence of the perturbing magnetic field.

and $M_L(\nu_{lm})$ describes a Lorentzian profile with the form

$$M_L(\nu; \nu_{lm}) = \frac{H}{1 + \frac{4}{\Gamma^2}(\nu - \nu_{lm})^2}, \tag{4.31}$$

where Γ is the mode linewidth, and the mode height is H.

We used $\Gamma = 1\,\mu$Hz and $\nu_s = 0.4\,\mu$Hz. The generated data were given a frequency resolution and height-to-background ratio, H/B, representative of a *Kepler* 90-day dataset. We made data for different inclination angles, covering the range zero to 90 degrees in steps of 0.5 degrees, and different asymmetries a in the range $(-1.0 \leq a \leq 1.0)\,\mu$Hz at increments of 0.2μHz. Noise was included by perturbing the limit spectra with negative exponential statistics (Anderson et al., 1990).

A symmetric model was fitted to the generated data using Markov chain Monte Carlo (MCMC) techniques (e.g. Goodman & Weare, 2010). The parameters H, Γ, ν_s, i and B were set as their true values for all fits and the process was run to find ν_0 using 100 simultaneous walkers and 5000 iterations. For a single inclination

angle and asymmetry this was repeated 50 times and a total posterior distribution constructed over all fits and iterations. The median gave the fitted frequency, $\nu_{\rm fit}$, and the uncertainties were taken from 68 % confidence intervals. We found that small changes to the underlying mode parameters produced almost identical results for the fitted centroid frequencies.

By fitting a symmetric model to asymmetric data we tested how much the central frequency of an $l = 1$ mode is affected by fitting an inaccurate symmetric model. This depends on the balance of each azimuthal contribution and therefore on the inclination angle. We define a quantity α which describes the relative contributions from $m = 0$ and $|m| = 1$ components, ν_{10} and $\nu_{1|1|}$ respectively, to the fitted frequency in the equation

$$\nu_{\rm fit} = \alpha\nu_{10} + (1 - \alpha)\nu_{1|1|}, \tag{4.32}$$

where

$$\nu_{1|1|} = \frac{1}{2}(\nu_{11} + \nu_{1-1}) \equiv \nu_0 + a. \tag{4.33}$$

Rearranging Equation 4.32 gives

$$\alpha = \frac{\nu_{\rm fit} - \nu_{1|1|}}{\nu_0 - \nu_{1|1|}} = \frac{\nu_{\rm fit} - \nu_{1|1|}}{-a}. \tag{4.34}$$

Therefore if we plot $(\nu_{\rm fit} - \nu_{1|1|})$ against a, the negative gradient will give α. Values of α along with their uncertainties were calculated for each inclination angle. A spline fit to this relation enabled interpolation for any other angle (see Figure 4.11).

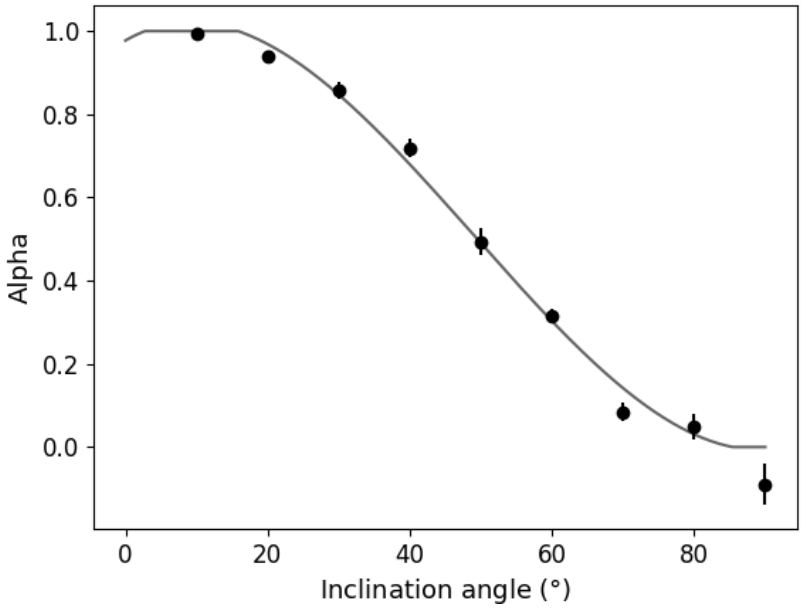


Figure 4.11: Plot of α versus inclination angle found using methods described in the text. The solid line shows the spline fit used to find α at any angle, i.

5 Magnetic activity effects on asteroseismic inference

This chapter is taken verbatim from Thomas et al. (2021), published in Monthly Notices of the Royal Astronomical Society, Volume 502, Issue 4, pages 5808-5820. I was first author on this paper and completed the majority of the work. One exception to this was running one of the two stellar modelling pipelines (YGM) which was performed by Professor Sarbani Basu.

The work presented here builds on the ideas from Chapter 4 of using activity-affected oscillation frequencies to study how this concept will impact predictions we make about the properties of stars. We generated frequencies for several artificial Sun-like stars with various amounts of magnetic activity and determined the properties we would estimate from two stellar modelling pipelines. By comparing the difference in predicted ages, masses and temperatures, amongst other properties, of these stars with that of equivalent stars in the absence of magnetic activity we investigated the bias to expect due to the presence of activity.

This first study focused on the potential issue of how magnetic activity affects stellar property estimation of Sun-like stars but as future work any impact should be investigated for elsewhere in the HR diagram. In order to apply similar analysis to other solar-like oscillators we would need to have a clearer expectation of how the activity changes with stellar characteristics such as mass and luminosity. The activity patterns and field strengths on F and K stars are likely to be very different

from those of the Sun (e.g. Mathur et al., 2014b; Zhang et al., 2020). We would also need a better understanding of changes in activity behaviour with stellar evolution, for example between the active and inactive branches (e.g. Fabbian et al., 2017). Our model for estimating the expected frequency shifts could be adapted to account for these differences.

Asteroseismic evidence of stellar activity cycles (e.g. Salabert et al., 2011; Régulo et al., 2016; Salabert et al., 2016b; Kiefer et al., 2017; Santos et al., 2018; Kiefer et al., 2019) and other magnetic activity proxies (e.g. Mathur et al., 2019) have been found for many solar-type stars in the Kepler *sample. Therefore, a natural next step from the work presented here would be to conduct a statistical survey of any magnetically active* Kepler *and TESS stars for which we would expect the biases on property predictions to be significant. Indeed, quantifying the expected fraction of impacted stars is needed for wider applicability. A plan for future work would first be to analyse the high-quality stars in the* Kepler *archive in order to understand the extent to which targets observed by* Kepler *and TESS will be affected.*

Finally, another extension to this work could be to investigate topologies of surface magnetic activity different to the simple model we use of a single band of active latitude in each hemisphere. Our activity distribution model could be adapted to include more complex shapes or additional bands of activity, however, in order to constrain these new models we would require higher fidelity data.

5.1 Introduction

Sun-like stars with outer convective zones pulsate due to turbulent motion within these layers. These perturbations excite acoustic waves within the star producing a rich spectrum of modes of oscillation which can be studied via asteroseismology to investigate internal stellar physics and determine global properties. Thanks to the high-resolution photometric observations from CoRoT (Baglin et al., 2006), *Kepler* (Borucki et al., 2010; Howell et al., 2014), and more recently TESS (Ricker et al., 2015), we have measured acoustic oscillations for thousands of stars.

Surface magnetic activity is known to affect acoustic modes, changing their frequencies, enlarging damping rates and suppressing amplitudes (e.g. Christensen-Dalsgaard, 2002; Chaplin et al., 2007a; Metcalfe et al., 2007; Fuller et al., 2015; Kiefer et al., 2019). Several studies have found evidence of acoustic modes being shifted in frequency, an effect that varies with the 11-year solar activity cycle (e.g. Woodard & Noyes, 1985; Pallé et al., 1989; Elsworth et al., 1990; Howe et al., 2002; Chaplin et al., 2007a; Broomhall et al., 2014) as well as with a quasi-biennial period (e.g. Fletcher et al., 2010; Broomhall et al., 2012; Simoniello et al., 2012). Similar behaviour has also been found in other solar-type stars (e.g. García et al., 2010; Salabert et al., 2016b; Karoff et al., 2018; Santos et al., 2018). Kiefer & Broomhall (2020) showed that mode frequencies from main sequence and subgiant stars are particularly sensitive to perturbations from magnetic activity. The sizes of frequency shifts tell us about activity related changes in the outer layers of stellar interiors enabling us to probe activity cycles (Kiefer et al., 2017; Santos et al., 2018), and relative sizes of shifts can reveal information about the surface activity distribution (Salabert et al., 2018; Thomas et al., 2019). It is important to account for activity related effects on asteroseismic measurements since they are often used in stellar modelling or to infer global stellar properties such as mass, age or radius (see e.g. Chaplin & Basu, 2015; Kiefer et al., 2019; Pérez Hernández et al., 2019; Santos et al.,

2019b).

When modelling stars the description of the near-surface layers is generally incomplete causing a discrepancy between theory and observation. Different techniques have been developed to deal with this (see Section 5.2), however, generally these assume spherical symmetry within these regions which is not true for magnetic activity. The aim of this work was to determine the impact a non-homogeneous spread of near-surface activity has on the estimates of fundamental stellar properties made by asteroseismic modelling pipelines and in what situations the effect must be considered. Since this is a first attempt to quantify the expected bias due to magnetic activity, we considered Sun-like stars since the activity patterns on the Sun are well known. We generated artificial observations for Sun-like stars with various near-surface field strengths and spatial distributions of surface activity. We studied the difference between estimations of stellar models fitted to shifted and non-shifted sets of frequency separation ratios. Comparisons made between these results would reveal the impact of activity. The justification for using separation ratios is explained in Section 5.2. In Sections 5.3 and 5.4 we outline the model used to calculate activity-induced frequency shifts and our process for generating sets of frequencies. Section 5.5 contains a description of the two stellar modelling pipelines used to infer global properties. Our results are shown in Section 5.6 followed by discussion and conclusions.

5.2 Use of separation ratios

Asteroseismology is a powerful tool to infer fundamental properties of solar-type stars. With long-timebase photometry from, for example, *Kepler* and TESS it is possible to resolve individual modes of oscillation in stellar spectra giving us a window into the inner workings of stars. Asteroseismic modelling pipelines can be used to obtain precise properties of stars using inputs of mode frequencies along with complimentary non-seismic data, typically, but not limited to, effective tempera-

ture, metallicity, and luminosity derived using parallaxes. More robust and higher accuracy constraints on stellar properties, including mass, radius and age, are possible with the inclusion of individual mode frequencies, or combinations of frequencies, as opposed to solely using global seismic quantities (e.g. Gai et al., 2011; Lebreton & Goupil, 2014; Metcalfe et al., 2014; Reese et al., 2016; Silva Aguirre et al., 2017). Using a set of input physics and evolutionary codes stellar models are computed, either on the fly or to build a predefined grid, where each stellar model corresponds to a combination of properties for a model star. For each model pulsation codes are then used to predict theoretical oscillation frequencies. Theoretical observables, including frequencies and additional non-seismic data, are fit to actual observations to obtain the best matching model and the corresponding stellar properties.

The oscillations are acoustic modes where pressure perturbations drive standing waves within a main sequence star. Spherical harmonics are used to describe the appearance of these modes on a sphere with oscillations usually described by three numbers: the radial order, n, the angular degree, l, and the azimuthal order, m. Solar-like oscillators produce a spectrum of modes, the frequencies of which depend on the star's properties and internal structure. However, for solar-type stars observed by *Kepler* and TESS it is only possible to measure modes with $l \leq 3$ due to geometric cancellation for higher degrees.

When modelling oscillation frequencies there is a known systematic difference between models and observations called the surface effect. This frequency bias is caused by the incomplete modelling of near-surface layers of stars, for example by using approximations such as mixing-length theory, or inadequate modelling of the interactions between oscillations and convection (for more description see Basu & Kinnane (2018) and references therein). There are however several methods to allow for this effect. Correction terms have been included to account for the offset (Kjeldsen et al., 2008; Gruberbauer et al., 2013; Ball & Gizon, 2014; Sonoi et al., 2015), or the use of an asteroseismic phase to parameterise the frequency dependent

difference between model and observation (Roxburgh, 2015).

Another method is to use combinations of frequencies when fitting rather than the individual frequencies themselves. These are known as 'separation ratios' and are useful since they are roughly independent of the structure of surface regions of stars thereby mitigating the impact of the surface effect (Roxburgh & Vorontsov, 2003). For this work we use the r_{02} separation ratio defined as

$$r_{02}(n) = \frac{d_{02}(n)}{\Delta\nu_1(n)},$$

where

$$d_{02}(n) = \nu_{n\,0} - \nu_{n-1\,2}\,, \tag{5.1}$$

$$\Delta\nu_1(n) = \nu_{n\,1} - \nu_{n-1\,1}\,.$$

Here $\nu_{n\,l}$ is the frequency of a mode with radial order n and angular degree l, $d_{02}(n)$ is the small frequency separation between $l = 0$ and $l = 2$ modes, and $\Delta\nu_1(n)$ is the large frequency separation for $l = 1$ modes. The sensitivity of low-l modes to near-surface conditions is independent of the degree. Therefore the small separation is already fairly insensitive to surface layers since it calculates the difference between two modes of very similar frequency which both propagate in the near-surface regions. This sensitivity is reduced even further when taking the ratio of small to large separations. Roxburgh & Vorontsov (2003) compared stellar models with the same interior structure but different surface layers to illustrate that frequency ratios are much less sensitive to the poorly-modelled outer layers' conditions. As a result, despite some loss of information when taking ratios, separation ratios can be used to isolate the effects of the deep stellar interior, a key focus for those determining ages and evolutionary states of stars. Due to this sensitivity to central conditions and being almost unaffected by surface regions Silva Aguirre et al. (2013) argued that more reliable stellar properties can be obtain by using separation ratios rather than oscillation frequencies themselves. Basu & Kinnane (2018) showed that, as long as

the surface effect is somehow compensated for when inferring stellar properties from models, then the obtained results are robust.

Nevertheless, one underlying assumption when using frequency ratios is that there are no non-spherically symmetric near-surface perturbations (Otí Floranes et al., 2005) which would induce frequency shifts that depend on the degree of the mode. One example source would be a surface magnetic activity distribution such as that seen on the Sun. The $\Delta\nu_l$ are calculated from modes with the same combination of m and l therefore changes in $\Delta\nu_l$ due to activity will be negligible. However, d_{02} uses frequencies with different combinations of m and l which occupy different spatial regions on a star and hence experience different size shifts in the presence of non-spherically symmetric magnetic activity. For this reason d_{02}, and hence r_{02}, will carry a signature of magnetic field changes; for example the solar cycle. Chaplin et al. (2005) studied this effect using Sun-as-star observations from BiSON (Chaplin et al., 1996) and measured variations in the ratios with solar activity level. They attribute this change to acoustic asphericity from surface activity and advised that care must be taken to account for biases when using ratios from long data sets. For this reason in this study we use separation ratios when fitting stellar models, avoiding the need for a surface term and we know that any remaining effect will be caused by non-spherically symmetric activity.

5.3 Modelling activity-induced frequency shifts

The shift in frequencies due to surface activity will depend not only on the star's magnetic field strength but also the spatial distribution of the activity on the stellar surface. What's more, our ability to observe these shifts will depend on the inclination angle of the star since this dictates the mode visibility and hence our ability to observe mode components of different l and m. For this work we build on the model from Thomas et al. (2019) but summarise the main principles here.

The response of each mode depends on the field strength in the region where

the corresponding acoustic wave propagates so the impact of a non-homogeneous distribution of activity on a mode will depend on the mode's spatial distribution. Assuming that the frequency shifts are caused by a source in the near-surface regions of the star, the shift experienced by modes of different m and l is given by (Moreno-Insertis & Solanki, 2000):

$$\delta\nu_{lm} \propto \left(l + \frac{1}{2}\right) \frac{(l - |m|)!}{(l + |m|)!} \int_{\theta_{\min}}^{\theta_{\max}} |P_l^{|m|}(\cos\theta)|^2 \, B(\theta) \, \sin\theta \, d\theta. \tag{5.2}$$

From this we can see that a mode's sensitivity to activity has a term that depends on the mode m and l whose spatial response is described by the associated Legendre polynomials, $P_l^{|m|}(\cos\theta)$. This is combined with the magnetic field strength, $B(\theta)$, which is a function of the distribution of activity. The above describes this distribution using θ, the co-latitude on the sphere, however, for the rest of this paper we will use the latitude, λ, where $\theta = (\frac{\pi}{2} - \lambda)$. The arrangement of activity on the stellar surface therefore determines the relative magnitudes of shifts for modes of different l and m. To calculate these shifts we assume the same top-hat model as Thomas et al. (2019), defined such that the magnetic activity is distributed in each hemisphere as a band of uniform field strength, B, lying between latitudes $\lambda_{\min}$ and $\lambda_{\max}$, i.e.

$$B(\theta = \frac{\pi}{2} - \lambda) = \begin{cases} B, & \text{if } \lambda_{\min} \leq \lambda \leq \lambda_{\max}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

The southern hemisphere is assumed to be a reflection of the northern hemisphere, since the globally coherent modes have no sensitivity to differences between hemispheres.

Although theoretically we can describe the response of individual m modes to activity, it is not always possible to isolate them in a frequency spectrum. Typically, stellar modelling pipelines use only one frequency per n and l mode as input so the frequencies of the m components must be combined. The relative visibility of each

azimuthal mode depends on the inclination angle of the star, i. By introducing a weighting factor, $\alpha(i)$, which is a function of the inclination angle, Thomas et al. (2019) defined how the contributions combine to give the central frequency of an $l = 1$ multiplet. For this work we used this along with an equivalent weighting for the $l = 2$ multiplet so as to include the mode visibility dependence on i. Therefore the *measured* frequency shift of the combined l peak will depend not only on the magnitudes of the individual m shifts but also on the inclination angle.

It is well known that activity-induced frequency shifts have a dependence on the radial order of the mode (Libbrecht & Woodard, 1990; Chaplin et al., 1998). Higher frequency modes have shallower upper reflection points than their lower-frequency counterparts and are therefore more sensitive to the perturbations in the layers closer to the stellar surface. That they therefore show larger frequency shifts has been observed for the Sun and also other solar-like oscillators (e.g. Salabert et al., 2011, 2016b; Kiefer et al., 2017; Salabert et al., 2018; Santos et al., 2018). To account for this in our artificial data we adjusted the shifts to have a Sun-like frequency dependence as represented by a polynomial relation in frequency (Chaplin et al., 1998; Howe et al., 2017).

Using the same frequency shift model, Thomas et al. (2019) found that the active latitudes required to produce the observed solar shifts (for solar cycle 23) extend between $\lambda_{\rm min} = 3.3^\circ$ and $\lambda_{\rm max} = 40.6^\circ$. This describes the spread of significant, large-scale field on the solar surface. We therefore re-parameterise Equation 5.3 to replace B by a relative field strength $B_{\rm rel}$, normalised to unity for the Sun:

$$B(\theta = \frac{\pi}{2} - \lambda) = \begin{cases} B_{\rm rel}, & \text{if } \lambda_{\rm min} \leq \lambda \leq \lambda_{\rm max}, \\ 0, & \text{otherwise.} \end{cases} \tag{5.4}$$

To enable this we introduced a multiplicative calibration constant $C_{\delta\nu}$ into the frequency shift calculation, turning the proportionality in Equation 5.2 to an equality

giving

$$\delta\nu_{lm} = C_{\delta\nu}\, B_{\rm rel} \left(l + \frac{1}{2}\right) \frac{(l-|m|)!}{(l+|m|)!} \int_{\theta_{\rm min}}^{\theta_{\rm max}} |P_l^{|m|}\,(\cos\theta)|^2\, \sin\theta\, d\theta. \tag{5.5}$$

The value of this calibration constant is fixed to $C_{\delta\nu} = 0.8$ in order to produce modelled shifts that match those from BiSON observations for $B_{\rm rel}$= 1, $i = 90°$, and the above mentioned latitude parameters pertaining to the Sun.

Our model uses Equation 5.5 to calculate the frequency shift of each m mode. As described above we used the weighted contributions of the azimuthal components to calculate the central frequencies of the shifted $l = 1$ and $l = 2$ combined peaks. Finally all frequencies were corrected for the radial order dependence. We used this model to generate artificial activity-affected frequencies sets.

From the above we know that the measured (i.e. observed) frequency shift depends on the degree l, inclination angle i, relative magnetic field strength $B_{\rm rel}$, and distribution of activity $\lambda_{\rm min}$ and $\lambda_{\rm max}$. Increasing the magnetic field strength will induce a larger magnitude frequency shift. At certain inclination angles it will be easier to observe particular m components due to the relative mode visibility. The response of modes and how we measure them is complex. Figure 5.1 shows an example of how the measured frequency of a $l = 1$ mode responds to changes in inclination angle for a few different $\lambda_{\rm min}$ and $\lambda_{\rm max}$ combinations. We can see that the response is not necessarily straightforward. The $l = 2$ case is even less intuitive since there are five m components to combine which each depend on the activity distribution and whose relative contribution to the measured $l = 2$ frequency also depends on the inclination angle.

5.4 Generating artificial frequency sets

We use two stellar modelling pipelines to infer properties for artificial stars from their 'observed' frequencies: Asteroseismic Inference on a Massive Scale (AIMS) and another grid-based approach which we call Yale Grid-based Modelling (YGM)

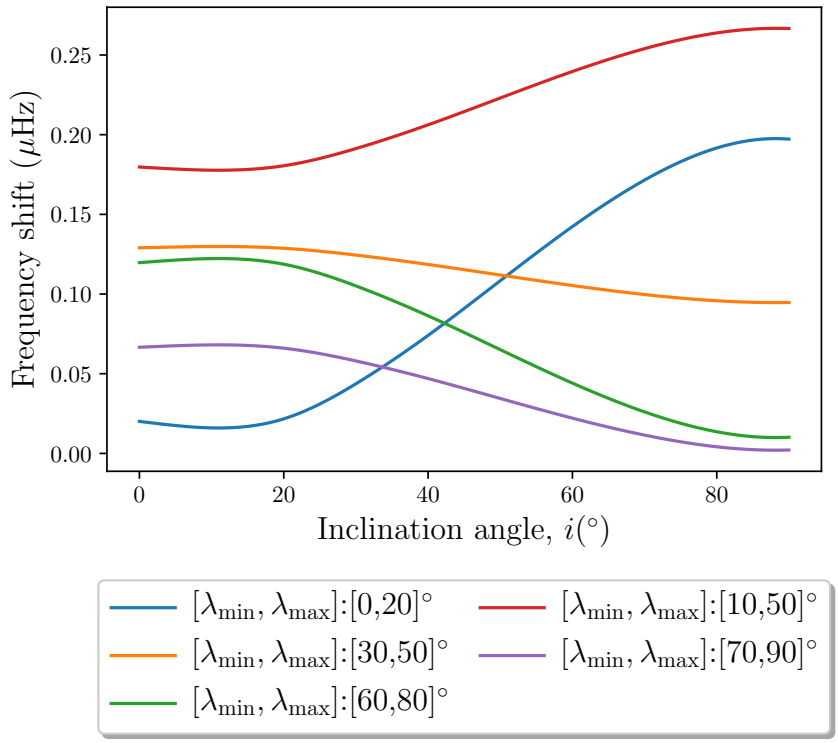


Figure 5.1: Example frequency shifts of $l = 1$ modes at various inclination angles in response to different magnetic activity distributions. The field strength was kept constant.

(see Section 5.5 for more details). Both methods can fit to separation ratios thereby avoiding the need for a surface term.

To build our artificial data sets we started with a set of 'pristine' frequencies, free from any simulated magnetic activity effects. At an activity minimum we expect a more uniform distribution of activity across the stellar surface so all modes experience the same size of shift. This cancels out when taking the ratio of frequencies. Therefore by using ratios when fitting stellar models, our pristine separation ratios are equivalent to what we would observe on a field-free star, and also what we would expect to observe at minimum levels of stellar activity. Chaplin et al. (2019) showed that minimum-epoch solar p modes should have frequencies very close to field-free case.

For the AIMS analysis presented in Section 5.6 the pristine frequencies were taken from the model in the grid which was most similar to the Sun in terms of mass and age (4.61 Gyr). The pristine frequencies for the YGM analysis were taken from

a calibrated Standard Solar Model (SSM). This was created with the same input physics as the grid that was used for fitting, except for the atmospheric model, which was that of Krishna Swamy (1966) (see Section 5.5 for details of the physics of the grid). As is usual in constructing SSMs, we iterated over the mixing length parameter and the initial helium abundance in order to get a $1M_{\odot}$ model that has the correct radius and luminosity at the solar age (4.57 Gyr). The converged model has a mixing-length parameter of 2.1566 and an initial helium abundance of 0.2734. The model has a convection-zone helium abundance of 0.2447, and Z/X of 0.02299. The base of the convection zone is at $0.71317R_{\odot}$. For completeness we repeated the analysis with the pristine frequency sets swapped; i.e. the AIMS pipeline was also run with data sets based on the SSM frequencies, and the YGM analysis using frequency sets based on those from the most solar-like AIMS model. The results were in agreement whichever set of pristine frequencies were used as a base.

Data sets were comprised of the 10 overtones of degrees l=0,1,2 centred on $\nu_{\rm max}$ (to match the procedure of Ball & Gizon 2014). Frequency uncertainties were taken from BiSON 1-year data and are comparable to uncertainties given by *Kepler* data of duration a year or more from high-quality SNR targets.

Using the pristine data as the base, activity-affected frequency sets were generated by shifting the pristine modes according to our model and the chosen combination of $B_{\rm rel}$, i, $\lambda_{\rm min}$ and $\lambda_{\rm max}$. The artificial data were created to represent solar-like oscillators at various inclination angles and with a variety of magnetic activity strengths and distributions.

To choose interesting combinations we first determined those which would produce a set of shifted separation ratios that were, on average, discernibly different from those of the pristine set, i.e beyond the uncertainty of the pristine ratios (Figure 5.2 shows an example set of shifted r_{02} compared to the pristine r_{02} using $B_{\rm rel}$= 1.2, $i = 0^{\circ}$, $\lambda_{\rm min}$= 11° and $\lambda_{\rm max}$= 53°). This was motivated by our goal to find the combinations that would incur a significant bias in stellar property estimates from

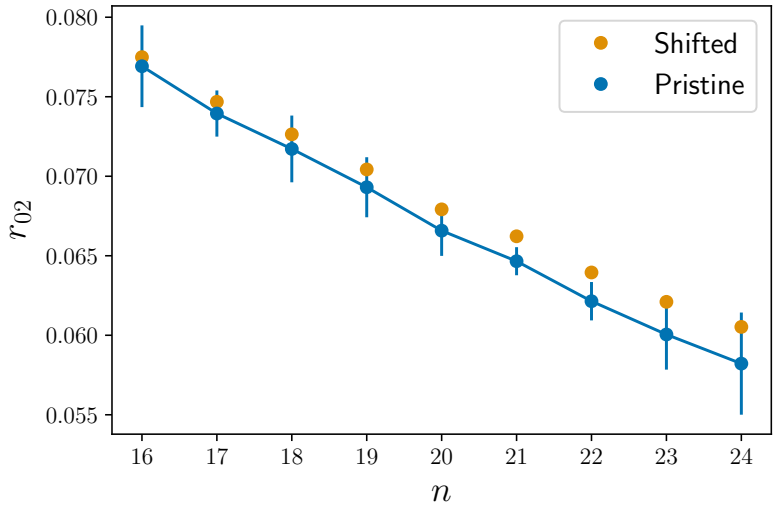


Figure 5.2: Example of artificially shifted r_{02} with respect to the pristine ratios, calculated using $B_{\rm rel}$= 1.2, i = 0°, $\lambda_{\rm min}$= 11° and $\lambda_{\rm max}$= 53°. An effect of this size produces an average shifted r_{02} which lies just outside of uncertainty of the average pristine r_{02}.

modelling pipelines. We constructed a grid of i, $\lambda_{\rm min}$ and $\lambda_{\rm max}$, each in the range $0-90°$ with increments of 1°, and calculated the minimum field strength, $B_{\rm min}$, needed to produce the desired shifted separation ratios. $B_{\rm min}$ was in fact taken to be the weighted average minimum field strength over all of the ratios. Figure 5.3 shows the latitudinal positions of activity bands (shaded regions) used to produce the required separation ratios across the range of inclination angles. For each element in the grid (corresponding to a particular i, $\lambda_{\rm min}$ and $\lambda_{\rm max}$) we shaded the region in latitude and inclination space with a colour intensity that was proportional to $1/B_{\rm min}$ for that element. This was repeated for the entire grid to build up Figure 5.3. Therefore the darker regions indicate where a lower $B_{\rm min}$ was necessary to sufficiently shift separation ratios beyond the pristine frequencies, and lighter areas where a much greater $B_{\rm min}$ was needed.

For each inclination angle we took the [$\lambda_{\rm min}$,$\lambda_{\rm max}$] pair corresponding to the smallest $B_{\rm min}$ thereby compressing a 3D grid to 1D. We chose combinations to study both when the $l=0$ shifts, $\delta\nu_0$, were larger than the $l=2$ shifts, $\delta\nu_2$, and vice versa. The relative magnitudes of these shifts affect the direction of the bias we get in the

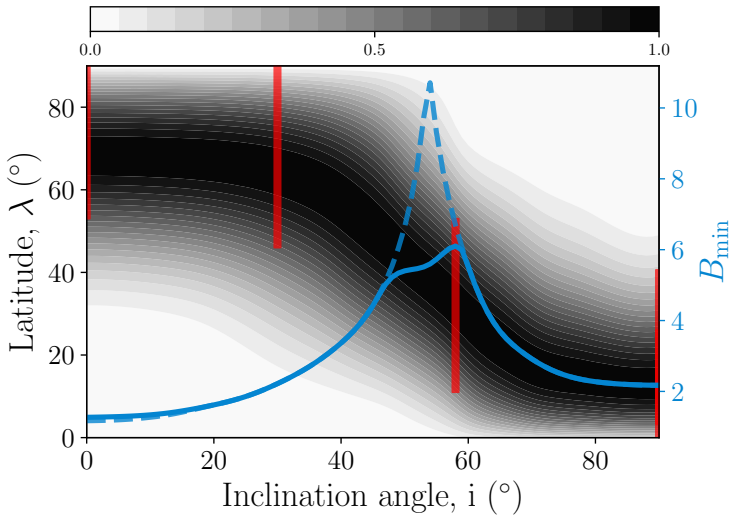
stellar property estimations from models. Figure 5.3a shows the results for $\delta\nu_0<\delta\nu_2$ shifts and Figure 5.3b for $\delta\nu_0>\delta\nu_2$ shifts. We can see that for each case the $B_{\rm min}$ is produced by activity bands at very different latitudes; for example to achieve $B_{\rm min}$ at high inclinations a lower latitude activity band would produce $\delta\nu_0<\delta\nu_2$ whereas a band situated at higher latitudes would cause $\delta\nu_0>\delta\nu_2$. The solid blue curve in each plot shows the smallest $B_{\rm min}$ value at each inclination angle and the dashed blue curve shows the same but for the opposite sign of shifts. We can see that at intermediate inclination angles the magnetic field strength would need to be larger than at low or high angles to induce the same size r_{02} shift. Smaller field strengths are necessary for the lowest and highest inclinations.

A variety of combinations were chosen to generate several sets of shifted frequencies. In addition, we also created artificial data to produce shifts that we would expect from the Sun using $i=90^\circ$ and $[\lambda_{\rm min},\lambda_{\rm max}]= [3.3, 40.6]^\circ$ as found by Thomas et al. (2019) for the Sun's activity distribution. For one set of frequencies we chose a solar-like magnetic field strength (i.e. $B_{\rm rel}= 1$). For another we used the same latitudes and the $B_{\rm min}$ value required to produce shifted ratios discernibly different from the pristine frequencies (which is higher than we see in the Sun). The red bars in Figure 5.3 correspond to the chosen latitudes and inclination angles. The collections of parameters we used to produce each data set are summarised in Table 5.1.

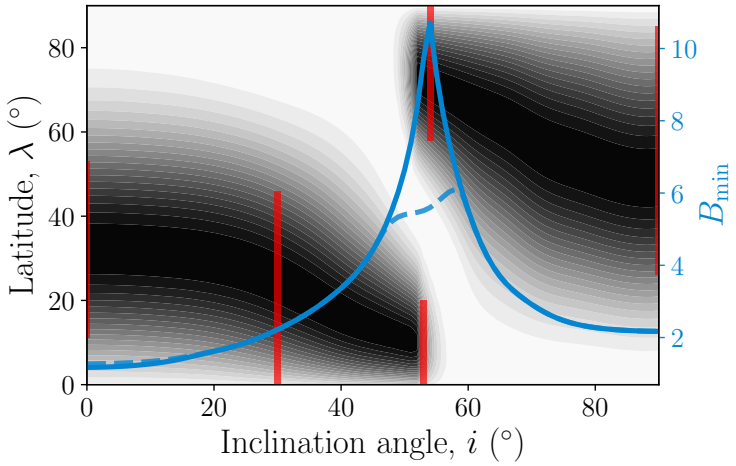
We fitted stellar models to our sets of artificial frequencies and compared deviations between their estimations of stellar properties from application to the pristine frequencies and those from the activity-shifted sets. Any differences must be due to the simulated magnetic effects. If activity does not affect stellar model predictions then all results will be similar to those obtained from pristine frequencies.

Table 5.1: Combinations of parameters used to calculate frequency shifts for artificial data sets. The method for choosing parameters is explained in the main text. The pristine data set is representative of a field-free star. The Sun model uses the parameters necessary to produce solar-like frequency shifts from our model. The Sun 2 model is the same but for a Sun with stronger magnetic field strength in order to make significantly shifted separation ratios (i.e. $B_{\rm rel}$=$B_{\rm min}$).

	Relative magnetic field strength, $B_{\rm rel}$	Inclination angle, $i(^\circ)$	Minimum latitude, $\lambda_{\rm min}(^\circ)$	Maximum latitude, $\lambda_{\rm max}(^\circ)$	Frequency shift at $\nu_{\rm max}(\mu$Hz)		
					$\delta\nu_{l=0}$	$\delta\nu_{l=1}$	$\delta\nu_{l=2}$
pristine	0.0	-	-	-	0	0	0
i	1.2	0	11	53	0.31	0.27	0.14
ii	2.2	30	0	46	0.66	0.45	0.56
iii	9.7	53	0	20	1.38	1.30	1.40
iv	10.7	54	58	90	0.68	0.87	0.58
v	2.2	90	26	85	0.51	0.38	0.38
vi	1.7	0	53	90	0.14	0.37	0.44
vii	2.2	30	46	90	0.26	0.55	0.51
viii	6.1	58	11	53	1.55	1.67	2.01
ix	2.2	90	0	26	0.40	0.61	0.68
Sun	1.0	90	3.3	40.6	0.25	0.34	0.34
Sun 2	4.0	90	3.3	40.6	0.99	1.36	1.36



(a) $\delta\nu_0 < \delta\nu_2$



(b) $\delta\nu_0 > \delta\nu_2$

Figure 5.3: Determining the combination of parameters to produce separation ratios discernible from the pristine set, where $B_{\min}$ is the minimum field strength needed to do this. The shaded areas show the latitudinal distribution of activity with darker regions indicating where a lower $B_{\min}$ was necessary. The colorbar shows how the shading is inversely proportional to the field strength with the maximum value of 1.0 corresponding to $\frac{1}{B_{\min}}$ for that inclination angle. Top: the case for $(\delta\nu_0)<(\delta\nu_2)$. Bottom: the case for $(\delta\nu_0)>(\delta\nu_2)$. The solid blue lines are the smallest $B_{\min}$ needed for a particular inclination angle. The dashed blue line is the same but for the opposite sign of shift. The red vertical bars indicate the latitudes occupied by the active band for each set of parameters we chose to focus on (see Table 5.1). Results were constructed from a grid covering $0 < i, \lambda_{\min}, \lambda_{\max} < 90°$, hence containing 91^3 models.

5.5 The modelling pipelines

We used two different pipelines to fit stellar models to oscillation frequencies. Both used a predefined grid of models. There are many different choices for how to carry out the analysis and the constraints to use when fitting which will impact the uncertainty on estimated stellar properties. In particular, anchoring the lowest frequency modes can reduce error bars as we will show later. Below we detail the input physics and briefly cover the methods of each pipeline, one of which implemented anchoring.

5.5.1 AIMS

The AIMS pipeline (Reese, 2016) uses individual oscillation frequencies, or in this case frequency ratios, along with classical constraints to determine global stellar properties. The grid of models we used was the same as the MS grid from Rendle et al. (2019a). Models were computed using the CLÉS (Code Liégeois d'Évolution Stellaire, Scuflaire et al. (2008a)) stellar evolution code and the grid was parameterised by mass in the range $0.75\mathrm{M}_{\odot}$ to $2.25\mathrm{M}_{\odot}$ with an interval of $0.02\mathrm{M}_{\odot}$, initial metallicity (Z_{init}) from $0.0032 - 0.0300$ and initial hydrogen content (X_{init}) in the range $0.691 - 0.745$. The Z_{init} and X_{init} values used can be found in Table 1 of Rendle et al. (2019a). Microscopic diffusion with a fixed solar-calibrated mixing length of 1.81 was included since Rendle et al. (2019a) found it to produce more closely matching values for the Sun (Thoul et al., 1994). The convective overshoot was 0.05 times the local pressure scale height, Grevesse & Noels (1993) abundances were used to convert [Fe/H] to Z/X, and nuclear reaction rates from taken Adelberger et al. (2011). The models were computed using opacities from Iglesias & Rogers (1996) and the equation of state from FreeEOS (Irwin, 2012). Frequencies were calculated using the LOSC (Liège Oscillation Code, Scuflaire et al. (2008b)) pulsation code.

AIMS combines approximating a set of best fitting models using a Markov Chain Monte Carlo (MCMC) algorithm (`emcee`, Foreman-Mackey et al. (2013)) with in-

terpolation implemented within the grid of models in order to refine constraints on properties. Interpolation is conducted using multidimensional Delaunay tessellation (see e.g. Field (1991)) both linearly along an evolutionary track and between tracks.

Further details may be found in Rendle et al. (2019a).

5.5.2 YGM

For YGM analysis we constructed a uniform grid of models for masses in the range $0.95M_{\odot}$to $1.05M_{\odot}$ with a spacing of $0.01M_{\odot}$. For each mass, models were created with fifteen values of the mixing length parameter spanning $\alpha_{\rm MLT} = 1.4$ to 2.625, initial helium abundances spanning from the primordial helium abundance of 0.248 (Steigman, 2010) to 0.30 in steps of 0.01, and initial [Fe/H] in the range -0.30 to +0.30 in steps of 0.01. We use the Grevesse & Sauval (1998) solar mixture to convert [Fe/H] to Z/X. The stars were modelled using the Yale Stellar Evolution Code, YREC (Demarque et al., 2008). For each of the parameters, the models were evolved from the zero-age main sequence to an age of 8 Gyr. Models were output at intermediate ages.

The models were constructed using the Opacity Project (OP) opacities (Badnell et al., 2005) supplemented with low temperature opacities from Ferguson et al. (2005). The OPAL equation of state (Rogers & Nayfonov, 2002) was used. All nuclear reaction rates are obtained from Adelberger et al. (1998), except for that of the $^{14}N(p,\gamma)^{15}O$ reaction, for which we use the rate of Formicola et al. (2004). All models included gravitational settling of helium and heavy elements using the formulation of Thoul et al. (1994). The frequencies of the models were calculated with the code of Antia & Basu (1994).

To determine stellar properties, we defined a goodness of fit for each model in the grid as follows. For each of the spectroscopic observables, [Fe/H], $T_{\rm eff}$ and luminosity L, we define a likelihood. For instance, the likelihood for effective temperature was

defined as

$$\mathcal{L}(T_{\rm eff}) = C\exp(-\frac{1}{2}\chi^2(T_{\rm eff})), \tag{5.6}$$

with

$$\chi^2(T_{\rm eff}) = \frac{(T_{\rm eff}^{\rm obs} - T_{\rm eff}^{\rm model})^2}{\sigma_T^2}, \tag{5.7}$$

where σ_T is the uncertainty on the effective temperature, and C the constant of normalisation. We define the likelihoods for [Fe/H] and L in a similar manner.

We considered the seismic data using the separation ratio r_{02}. For this we need to take error correlations into account and thus

$$\chi^2(r_{02}) = (\overline{r}_{02}{}^{\rm obs} - \overline{r}_{02}{}^{\rm model})^T \mathbf{C}^{-1} (\overline{r}_{02}{}^{\rm obs} - \overline{r}_{02}{}^{\rm model}), \tag{5.8}$$

where $\overline{r}_{02}{}^{\rm obs}$ is the vector defining the observed r_{02}, $\overline{r}_{02}{}^{\rm model}$ is the vector defining the r_{02} for the model at the observed frequency, and $\mathbf{C}$ is the error-covariance matrix. Thus

$$\mathcal{L}(r_{02}) = D\exp(-\frac{1}{2}\chi^2(r_{02})), \tag{5.9}$$

D being the normalisation constant.

The total likelihood is then

$$\mathcal{L}_{\rm total} = \mathcal{L}(r_{02})\mathcal{L}(T_{\rm eff})\mathcal{L}([{\rm Fe/H}])\mathcal{L}(L). \tag{5.10}$$

The likelihood was normalised by the prior distributions of each property in order to convert to a probability density. The medians of the marginalised likelihoods of the ensemble of models was then used to determine the parameters of the star.

However, the total likelihood defined in Equation 5.10 can result in erroneously high likelihood for some models. The surface term is smaller at low frequencies than at high frequencies, but the seismic likelihood function defined above does not take this into account. Presently, it could be possible to have a model with low $\chi^2(r_{02})$

but where the low frequency modes are badly fit. In order to down weight models for which frequency differences are large, we multiply Equation 5.10 with the term

$$\mathcal{L}_{\rm reg} = E \exp(-\frac{1}{100}\chi^2(\nu_{\rm low})), \tag{5.11}$$

where $\chi^2(\nu_{\rm low})$ is the χ^2 for the two lowest frequency modes of each degree and E is another normalisation constant. Note that Equation 5.11 is not a true likelihood function; the division of the χ^2 by 100 rather than 2 ensures that this term does not dominate the final selection process. One can set this anchoring of the low frequency modes in AIMS, however in order to show the range of results to expect using different analysis approaches here we ran AIMS without this constraint.

5.6 Results

We determined stellar properties for the pristine and activity-affected stars using the pipelines described in the previous section. Both methods were supplied with the artificial sets of frequencies and fitted using the separation ratios r_{02}. Observational constraints of effective temperature $T_{\rm eff} = 5777 \pm 80$K and metallicity [Fe/H] $= 0 \pm 0.1$dex were also provided and were the same for all datasets. Additionally, we refit the same frequency sets, this time with an extra constraint of luminosity $L = 1.00 \pm 0.03 L_{\odot}$. The assumed use of a luminosity uncertainty of 3% was based on *Gaia* (Gaia Collaboration, 2018) parallaxes for Sun-like stars. Figures 5.4 and 5.5 show results from fitting using just frequencies, $T_{\rm eff}$ and [Fe/H]; Figures 5.6 and 5.7 show results from fitting including a luminosity constraint.

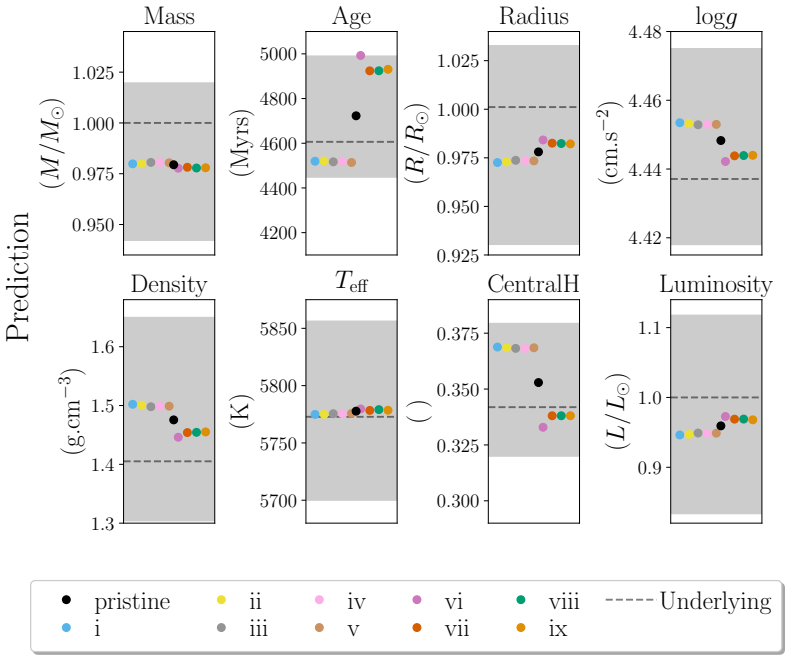
Presented are the median results taken from the posteriors of each property, either from the AIMS (Panels (a)) or the YGM pipeline (Panels (b)). The black circles indicate the median value of stellar properties obtained by fitting to a 'pristine', i.e. field-free, set of frequencies which act as a reference. The results obtained from other data sets have been spread along the x-axis for clarity. The grey band shows the uncertainties from 68% confidence intervals on the estimates from stellar

models applied to pristine frequencies. The horizontal grey dashed lines illustrate the underlying properties from the corresponding model used to generate pristine frequencies. There is a systematic offset between the underlying properties used to compute the pristine data set and the results from fitting to the pristine frequencies, however they are generally well within error. The focus of this work is on how the results from fitting to shifted frequency ratios differ to those from fitting to pristine frequency ratios since this will be due to magnetic activity effects.

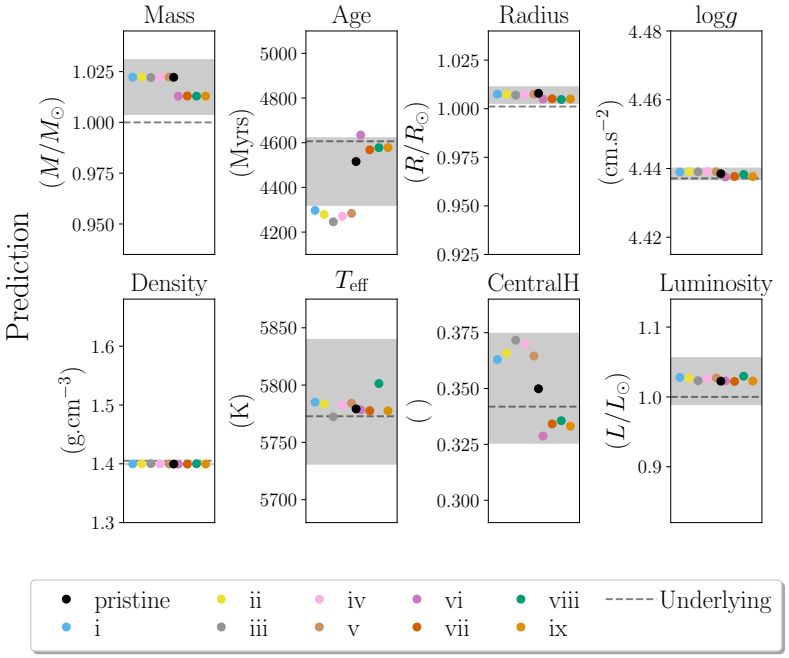
For the majority of properties the median estimates for the activity-affected data lie within the uncertainties of the equivalent pristine values. However, for all runs we can see that the largest differences between estimates from the pristine frequencies and those for different data sets are in age and central hydrogen abundance. For the case of the YGM pipeline we can see there is also a considerable spread in the $T_{\rm eff}$ values. The spread is more significant for the fits that included a luminosity constraint since the uncertainty bars are smaller.

It is clear that for stars experiencing this amount of activity-induced frequency shift, some of the stellar properties we infer will have a notable bias. Focusing on mass and age we find that the bias can be up to 5% in age, but only up to 0.5% in mass. This offset is therefore not a concern since for the analysis carried out in this work, the age parameter typically has a 4.5% uncertainty and we see a $\sim 2.5\%$ error on mass.

As described in Section 5.4, there are four combinations of results: AIMS fitted to AIMS model frequencies, AIMS applied to frequencies built on the SSM data, YGM applied to the AIMS model frequencies, and finally YGM analysis of the SSM frequency sets. In general we see similar results for all combinations. The uncertainties from YGM fitting are consistently smaller than those from AIMS but this can be attributed to the additional constraints placed on the low frequency modes (see Section 5.5). AIMS uncertainties match what we would expect from Rendle et al. (2019a) (Table 3) when fitting using separation ratios and a luminosity

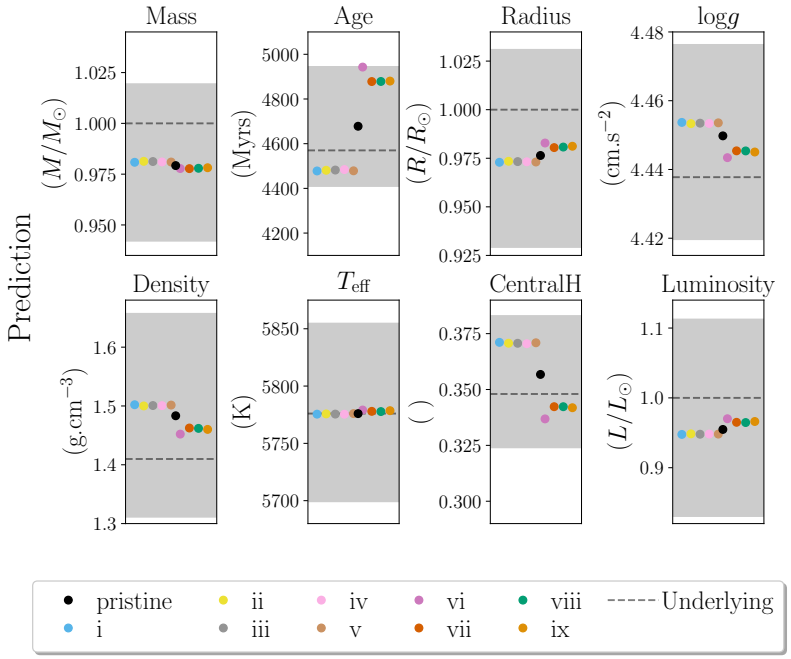


(a) AIMS fitting.

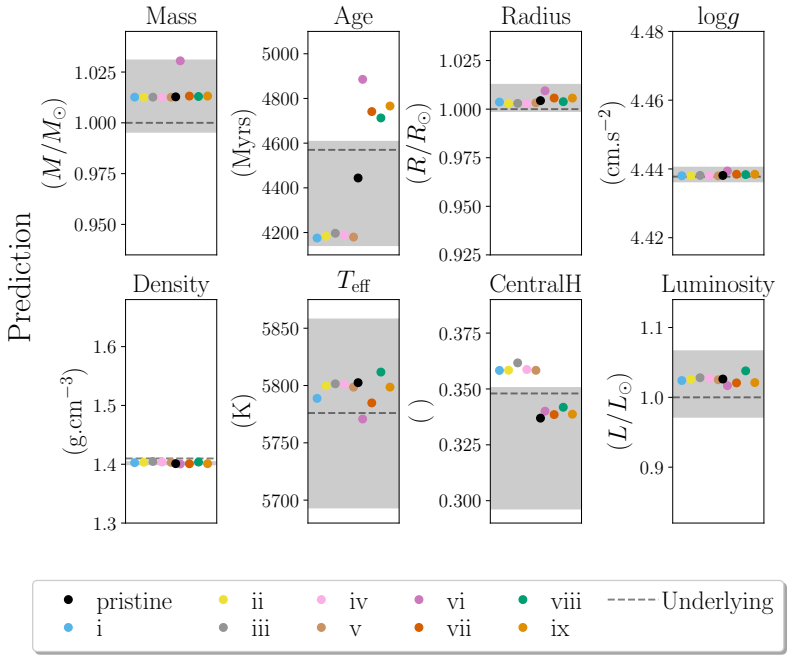


(b) YGM fitting.

Figure 5.4: Predictions from fitting without a luminosity constraint. The pristine data set was based on the most solar-like model from the AIMS grid. Black circles indicate median results from the pristine data set with the grey band showing the uncertainty on the stellar model estimates from the pristine frequencies. The coloured circles correspond to the results from frequency sets shown in Table 5.1, and are spread along the x-direction for clarity; their position along the x-axis has no meaning. The same y-axis ranges have been used in Figures 5.4-5.8 to more easily allow a like-for-like comparison between plots.

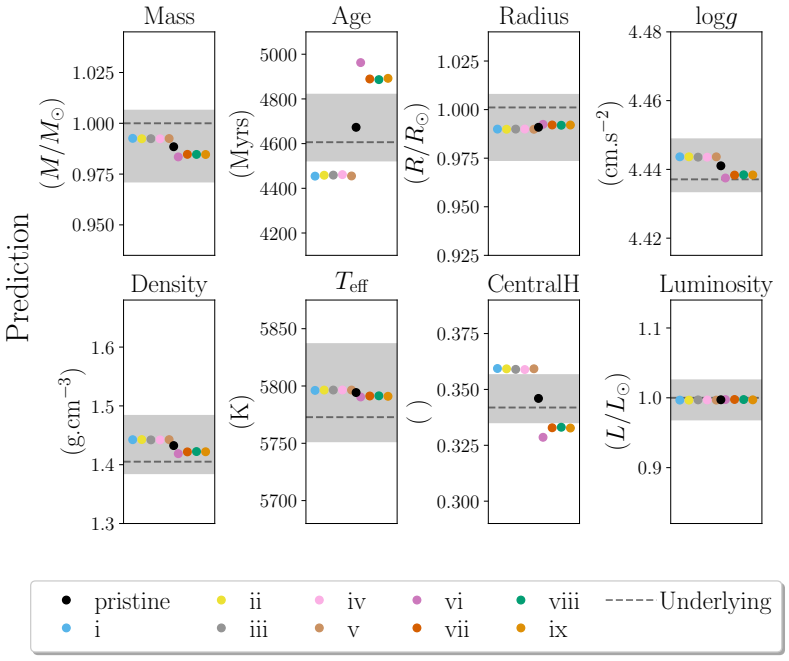


(a) AIMS fitting.

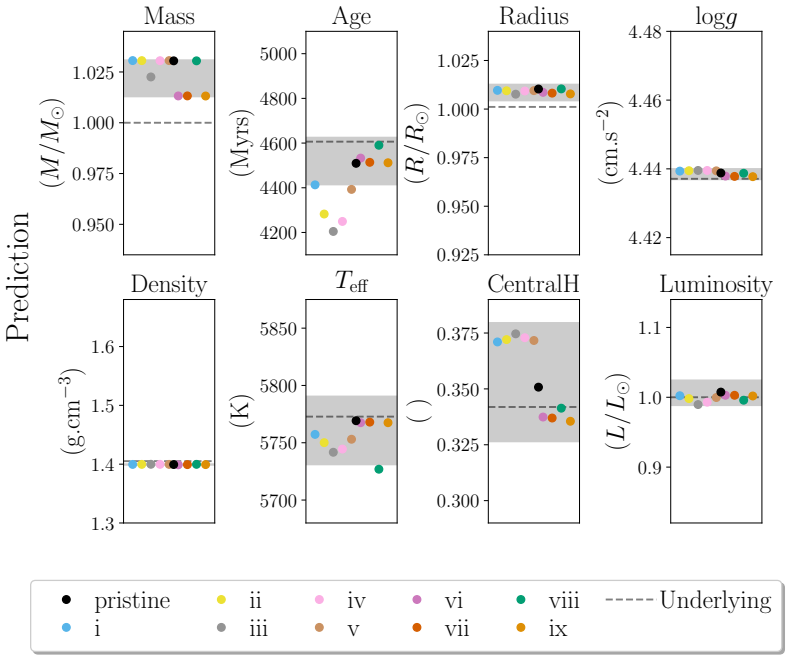


(b) YGM fitting.

Figure 5.5: The same as in Figure 5.4 but with the pristine data set based on the SSM.

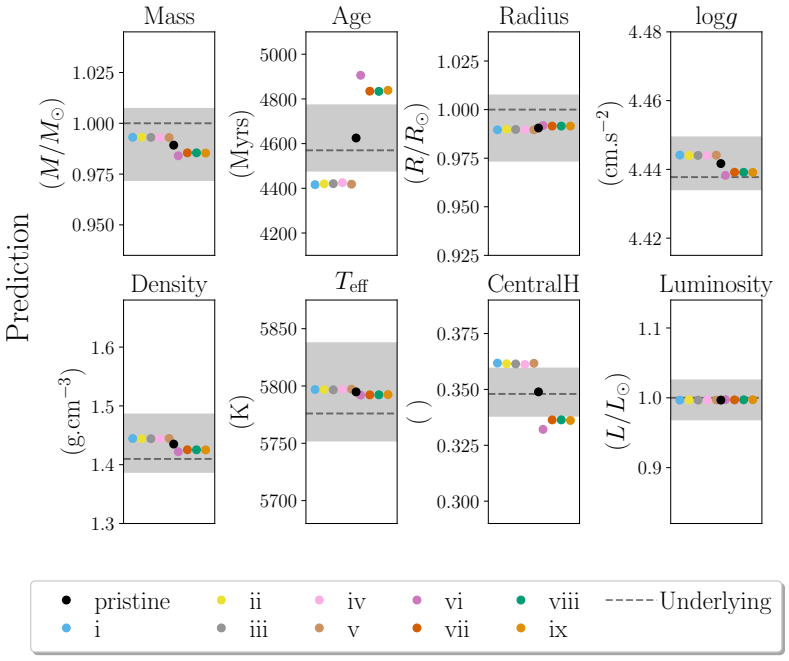


(a) AIMS fitting.

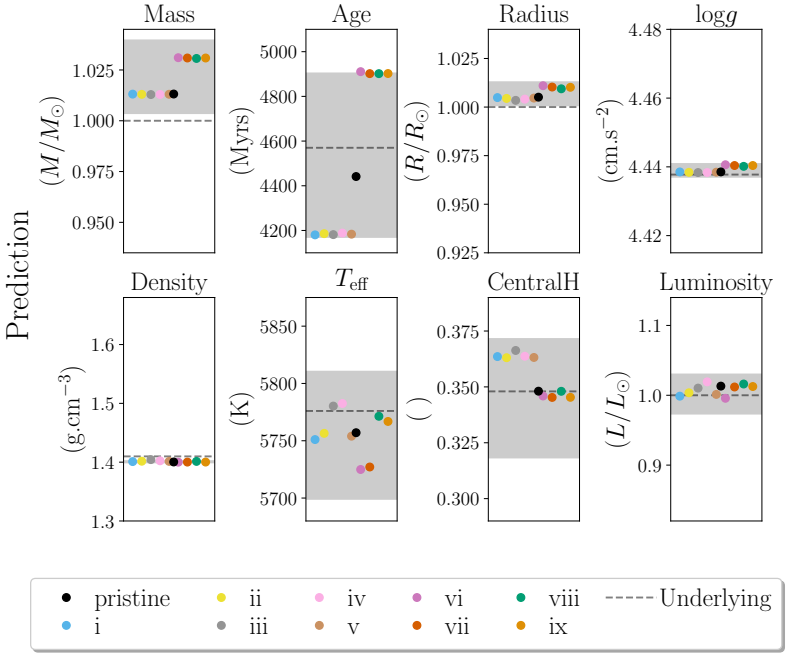


(b) YGM fitting.

Figure 5.6: The same as in Figure 5.4 with the pristine data set based on the most solar-like model from the AIMS grid but applying a luminosity constraint.



(a) AIMS fitting.



(b) YGM fitting.

Figure 5.7: The same as Figure 5.4) but with the pristine data set based on the SSM and applying a luminosity constraint.

constraint.

In the plots where the absolute spread of stellar property estimates is discernible from the pristine values we can see that there are two distinct groups of points above and below the pristine results. These correspond to where the underlying frequency shift at $l=0$ is larger than the $l=2$ shift and hence whether d_{02} is greater or smaller than for the pristine set. For example, if the shifted d_{02} is larger than the pristine d_{02}, i.e. $\delta\nu_0 > \delta\nu_2$, then the fitting will find a smaller age.

To verify that models were equally well fitted to shifted data sets as to the pristine frequency ratios we calculated the ratios of log-likelihoods between fits and found them to be approximately unity. We also tested the analysis methods described above by replacing the r_{02} frequency constraint with r_{01} data, and separately using both r_{02} and r_{01} data simultaneously. The resulting posterior estimates of stellar properties showed very similar patterns to the r_{02} results presented here.

5.6.1 The Sun

In addition to the various combinations tested above, we also studied the bias we would expect for frequency shifts from a Sun-like star compared to the pristine frequencies. As described earlier, shifts were calculated using $i = 90°$, $B_{\rm rel}= 1$ and $\lambda_{\rm min,max} = [3.3, 40.6]°$. Another set of frequencies was computed for the same parameter values except using $B_{\rm rel}=4.0$ to imitate a Sun with stronger magnetic field strength.

Figure 5.8 shows the results. The estimated stellar properties were not notably different from those one would obtain from pristine frequencies for the solar-like case ($B_{\rm rel}=1$). However, with four-times the field strength we should expect a bias for all properties, being at the 1σ level for the estimated age and central hydrogen content.

5.7 Discussion and conclusions

We have demonstrated that for some distributions and strengths of surface stellar activity the oscillation frequencies would experience a shift that impacts the proper-

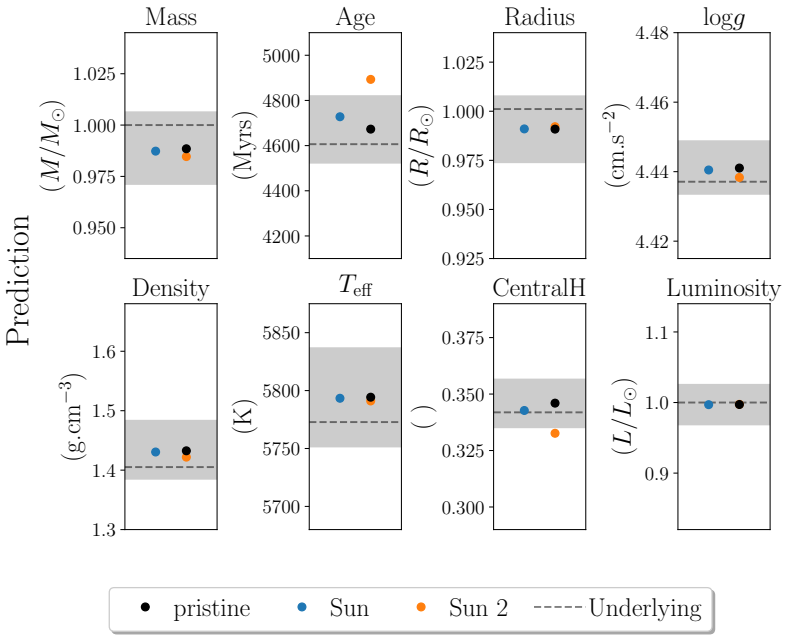


Figure 5.8: Results from fitting to two additional artificial sets: one using frequency shifts that we would expect to see in the Sun, and the other with same inclination angle and activity distribution as we would observe but with four times the magnetic field strength, i.e. $B = 4$. Fitting was conducted by AIMS including a luminosity constraint and the pristine data set was based on the most solar-like AIMS model.

ties obtained from stellar modelling pipelines when applied to separation ratios. The shifts we measure depend on a star's magnetic field strength, the activity distribution on the stellar surface, and it's inclination angle (the angle affecting which azimuthal mode components are detectable). Measured shifts therefore show a complex relationship between these variables. We generated several artificial sets of 'measured' frequencies using shifts arising from various combinations of the above. By fitting to separation ratios (r_{02}) constructed from the frequencies, global properties for these fake stars were estimated by two pipelines and compared to results from a field-free star.

Our results showed that estimates on stellar properties split into two groupings either side of the pristine result based on whether the shifted d_{02} is greater or less than the pristine d_{02}. The most noticeable divide is in the age parameter which is lower for an increased d_{02}. By extension, given the small range of metallicities here, a lower age will automatically result in a higher central hydrogen abundance. In

general we see a greater mass for those shifted data sets with larger d_{02} which is as expected given we are taking a cut in $T_{\rm eff}$.

The division into two groups implies that by measuring the shift in small frequency separation, it is possible to determine the direction of the biases, i.e. whether the property is an under- or over- estimate. Since the size of the bias depends on the frequency shifts experienced by the modes, and is therefore a complex function of the inclination angle, activity strength and distribution, it is more difficult to estimate the size of the bias. By pairing this with the methods of Thomas et al. (2019) it is possible to constrain the active latitudes present on the star using observations of frequency shifts over time along with the stellar inclination angle. If there is some way to estimate the star's magnetic field strength relative to the Sun then it could be possible to understand the expected size of the bias on properties, albeit with fairly large uncertainty.

We found that, in general, to experience a bias in property estimation larger magnetic field strengths are necessary. The deviations from the underlying properties would be larger for stars with a stronger magnetic field since this simply increases the magnitude of frequency shifts experienced by the modes. For the case of the Sun ($B_{\rm rel}$=1) the frequency shifts due to activity would not produce a significantly biased estimation of solar properties at the levels of precision tested here. In order for these to be affected the field strength would need to be approximately four times stronger.

The measured frequencies from stars with intermediate inclination angles are least susceptible to magnetic activity effects. As discussed in Section 5.3, this is due to the relative visibilities of the individual azimuthal modes and how their contributions to a central mode frequency are balanced. At these inclinations the field strength would need to be high for the observed separation ratios to be shifted far enough from the pristine r_{02} to have an impact on the estimated stellar properties. This can be seen in Figure 5.3 where the $B_{\rm min}$ value peaks at $i \sim 55°$, where $B_{\rm min}$

is the field strength required to produce shifted ratios discernibly different to the pristine ratios. However, these inclinations only account for $\sim$20% of stars (between $45-60°$). Assuming an underlying isotropic distribution of inclination angles, the relative number of stars observed as a function of i is proportional to $\sin(i)$, therefore observing a star with high inclination is more likely. For the lowest ($<40°$) or highest ($>70°$) inclinations $B_{\rm min}$ is lower. This means that for stars at these inclination angles more care must be taken to consider the bias on stellar properties due to activity.

For this analysis we took frequency uncertainties commensurate with *Kepler* data of a year or more. For shorter duration observations the frequency resolution will be reduced thereby minimising these effects since estimates of stellar properties will have larger uncertainties. The significance of the bias in properties will depend on the quality of data provided to the modelling pipeline. This includes additional spectroscopic measurements and whether low-frequency modes are constrained separately to separation ratios.

We find that for asteroseismic observations of Sun-like targets we can expect magnetic activity to affect mode frequencies which will bias the results from stellar modelling analysis. For most stellar properties we studied this offset should not be an issue since it is smaller than the uncertainties, including those on mass. However, for age and central hydrogen content the effect could be significant. Our results are consistent with those of Pérez Hernández et al. (2019) who investigated the biases arising on stellar property determination due to the influence of magnetic activity on the small separation of low-degree modes. They reported that biases are typically lower than systematic uncertainties but are largest for age estimates which is in agreement with our findings. Particular care must be taken when analysing long duration observations of stars with stronger magnetic field strengths than the Sun for which we expect higher magnitude frequency shifts. The same is true for stars with very high or low inclination angles where, for the same field strength, the

shift in measured frequency separation ratios is easier to observe and therefore will produce a more significant bias.

An obvious next step is to assess the fraction of asteroseismic targets in the *Kepler* and TESS samples that might be susceptible to these effects based on results from asteroseismic signatures of stellar activity cycles (e.g. Salabert et al., 2011; Régulo et al., 2016; Salabert et al., 2016b; Kiefer et al., 2017; Santos et al., 2018; Kiefer et al., 2019) and proxies of magnetic activity (e.g. see Mathur et al. 2019 and references therein).

6 Mixture model for background fitting

6.1 Introduction

Asteroseismology is a particularly powerful tool for deriving stellar properties (Chaplin & Miglio, 2013) such as mass, radius and age, and, as we demonstrate in this chapter, effective temperature and bulk metallicity. These quantities are a focus for many areas of study, from testing our theories of stellar evolution (e.g. Silva Aguirre et al., 2011; Chaplin & Miglio, 2013, and references therein) and magnetic activity (e.g. Metcalfe et al., 2007; Karoff et al., 2019), to parameterising exoplanet hosts (e.g. Christensen-Dalsgaard et al., 2010b; Huber et al., 2013; Batalha, 2014; Van Eylen et al., 2015), ensemble asteroseismology (Chaplin et al., 2011a, e.g.), and studies of the galaxy (e.g. Miglio et al., 2013; Stello et al., 2015). Acquiring these fundamental properties for large ensembles of stars is key to understand how stellar observations relate to the physics of stars.

However, commonly used methods to obtain stellar properties are laborious, requiring observed quantities to be compared to theoretical predictions at multiple stages in a statistically-motivated way. Typically, as a first step a power spectrum model is fit to asteroseismic observations to acquire global oscillation parameters or, in some cases, peak-bagging to individual mode frequencies. This often requires human input to provide initial guesses or place sensible priors on model parameters. Following this is often another fitting step where asteroseismic measurements obtained from the previous stage are compared to large grids of stellar models. Here approximations are necessary in order for this stage to be computationally viable.

In addition to using asteroseismic measurements, this analysis also requires more classical observations (from external photometry or spectroscopy sources). On top of this, the asteroseismic analysis requires post-fitting corrections using calibration stars for which we need high quality data and detailed modelling. Although these methods are very useful for building large catalogues of stellar properties, the analysis process is very involved. They must be thorough so as to consider all sources of uncertainty and adequately marginalise over all model parameters.

Signatures of fundamental stellar properties are present in the oscillation spectra of solar-like pulsators. It is well known that the large frequency separation, $\Delta\nu$, scales to good approximation with the square root of the mean density (Ulrich, 1986) and that the frequency of maximum power, $\nu_{\rm max}$, is dependent on the effective temperature and surface gravity of a star (Brown et al., 1991). Combining these two global seismic quantities enables measurements of stellar mass and radius via the famous scaling relations (see Chapter 2 for more detail).

Beyond the signal from acoustic modes themselves, the background convective granulation signal, most clearly visible at low frequencies, can reveal information about the global properties of stars. Correlations have been identified between stellar brightness variations and surface gravity, enabling estimates of $\log g$ to be made using granulation background signals (e.g. Bastien et al., 2013; Pande et al., 2018; Bugnet et al., 2018). Links between granulation amplitude and metallicity have also been suggested (Corsaro et al., 2017). Empirical scaling relations have been derived between granulation parameters and fundamental stellar properties (e.g. Mathur et al., 2011b; Kallinger et al., 2014) including effective temperature, mass, radius and surface gravity. The possibility of obtaining effective temperature from asteroseismic observations alone would be incredibly useful. However, this quantity is usually determined from external photometry or spectroscopy measurements which have limited cross-over with asteroseismic observations.

The vast quantity of observations we have available thanks to *Kepler* and TESS

give us wide-ranging opportunities for astrophysical study but at the same time present new challenges for data analysis. Our current techniques must evolve to meet this demand. With such large volumes of data we can begin to build models guided by our current theoretical understanding along with the observations we already have for populations of stars. Encoding this *prior* belief into our analysis is a crucial concept to guide our fitting to realistic solutions, more rapidly and with less human involvement. In this work we detail our data-driven approach to building a *generative model* from which we can derive bulk stellar properties. Such a model means we have a full description of the data and how they were generated whilst incorporating our prior beliefs in probabilistic way.

Using data from extensive spectroscopic surveys, Ness et al. (2015) created a generative model, named *The Cannon*, capable of determining "stellar labels" including surface gravity, metallicity and effective temperature from spectroscopic data. Their data-driven method relies on training their model on a set of reference objects before applying to other survey objects, thereby achieving a root mean square (rms) difference comparable to measurement uncertainties elsewhere in the literature. Ness et al. (2018) proceeded to use *The Cannon* to predict stellar labels but replaced the previous spectroscopic data with the auto-correlation function of the power spectrum to achieve a precision similar to spectroscopy (< 100K in effective temperature and < 0.1dex in the logarithm of the surface gravity). However, neither study considered measurement uncertainties on the training data labels in their analysis. Other data-driven techniques have included using random forests to predict bulk stellar properties from classical observations (e.g. spectroscopy) and global asteroseismic measurements (Bellinger et al., 2016; Bugnet et al., 2018, 2019). Convolutional Neural Networks (CNNs) have been implemented to successfully derive stellar properties from 1D photometric time series data (Blancato et al., 2020) and separately from asteroseismic measurements in the form of an image (Hon et al., 2020). However, using a CNN requires the input observations to be re-scaled/standardised to have

the same dimensions since the model treats the data as an image which may result in information loss. Full propagation and marginalisation over uncertainties is also not considered.

Several studies have shown that a connection exists between the properties of a star and its oscillatory behaviour. Using this concept we implemented a data-driven approach to build a generative model relating the fundamental properties of a star to the shape of its frequency-power spectrum. A multivariate normal (MVN) distribution model was trained to learn the covariance between stellar properties and parameters of a model power spectrum for two populations of evolved stars: red giant branch (RGB) and red clump (RC). Once trained these MVNs were combined to form a mixture prior which was implemented when background fitting to the power spectra of new stars. Our model is therefore capable of distinguishing between RGB and RC stars.

By fitting directly to asteroseismic observations, posterior information over both stellar properties and power spectrum model parameters could be derived. Many previous approaches lacked marginalisation over the uncertainties of their model parameters thus risking improper treatment of errors. Our method considered the observational measurement uncertainties when training the population models and marginalised over all model parameters during application to new, unseen stars.

Our method was split into two steps: the 'training' step to generate the model prior, and the 'testing' step where we assumed that the model holds for all stars. Like other data-driven methods, our approach relied on having a set of *reference* objects for which labels are known and which sufficiently cover the parameter space of interest. The quality of label transfer is dependent on how well covered the parameter space is by the training data.

The layout of this Chapter is as follows. We begin by describing the data used for this work and any necessary preparation. We then proceed to detail the methodology of our approach including how the prior was trained and formed into a Bayesian

mixture model before being applied to unseen stars. By first applying our method to simulated observations we could ensure the derived model was a good representation of the toy data. Following this, we present the results from applying our technique to stars from the APOKASC-2 catalogue showing that we obtain consistent estimates of stellar properties using just the observed power spectrum data. Finally we conclude with the capabilities of our method along with the limitations and ideas for future development.

6.2 Data preparation

Global stellar properties were taken from the second APOKASC catalogue (Pinsonneault et al., 2018), hereafter denoted as APOKASC-2, which combined asteroseismic data produced by the *Kepler* Asteroseismology Science Consortium (KASC) with spectroscopic parameters from the Apache Point Observatory Galactic Evolution Experiment (APOGEE), hence APOKASC. Within the catalogue the effective temperature, $T_{\rm eff}$, and metallicity, [Fe/H], along with their uncertainties were taken from the fourteenth data release of the Sloan Digital Sky Survey (DR14) (Abolfathi et al., 2018). The $\nu_{\rm max}$ and $\Delta\nu$ values were a weighted average of results from five different pipelines (see Hekker et al. 2011b for details of each approach) and masses were calculated using corrected scaling relations with $\nu_{\rm max}$, $\Delta\nu$ and $T_{\rm eff}$. As discussed later, the data sets were divided into RGB and RC stars using evolutionary states derived using the method in Elsworth et al. (2017) where possible or inferred from DR13 otherwise. *Kepler* magnitudes, K, were obtained by querying the VizieR web service (Ochsenbein et al., 2000) using the KIC numbers and were assumed to have a 1% uncertainty (Huber et al., 2016).

To construct a training data set that contained labels describing the power spectra as well as stellar properties we performed a preliminary background fit to a collection of 2000 evolved stars. We refer to this as the 'pre-fitting' stage. *Kepler* lightcurves were obtained using the open source Python package `Lightkurve`

(Lightkurve Collaboration et al., 2018) which searches the data archive at the Mikulski Archive for Space Telescopes (MAST) website [1]. We used the PDCSAP flux for all available quarters of long cadence data. The time series were stitched together and outliers outside of 4 standard deviations were removed. A power spectrum was computed using `Lightkurve`'s inbuilt functionality and re-binned (taking the mean) to a bin width of 1μHz after having truncated the spectrum to remove frequencies below 3μHz.

Combining the resulting parameters from background pre-fitting with stellar properties taken from the APOKASC-2 catalogue, we built a reference data set with which to train our data-driven model. We removed any fits where the power spectra contained strange artefacts (e.g. sudden spikes of noise) or where the resulting fit parameters were deemed by eye to lie outside of the typical trends between parameters (these were obvious outliers). This was done so that our model was trained on high-quality sets of parameters but it should not affect how well our model works for difficult cases when applied to new data. As such, any stars considered outliers were added to a test data set on which our trained model was to be tested. We split the training data into two populations (1000 each for RGB and RC) to train separate MVN models. The stars making up our training set are shown in the $T_{\rm eff}$-$\nu_{\rm max}$ plane in Figure 6.1 covering $4250 \leq T_{\rm eff} \leq 5196$K and $9 \leq \nu_{\rm max} \leq 237\mu$Hz. Here the marker size indicates the $\Delta\nu$ of the stars with values ranging from $1.4 - 17.6\mu$Hz. Metallicity is shown using the colour scale and covers the range -1.55 – 0.53 dex for our sample.

6.3 Method

6.3.1 Generative model for stellar property determination

We now lay out the generative model for determining fundamental stellar properties directly from fitting to power spectra. This assumes that the relationship learnt

[1] https://archive.stsci.edu/

Parameter	Bounds	
$a_{1/2}$	$10 - 2000$	(ppm)
$b_{1/2}$	$3 - 300$	(μHz)
Γ	$1 - 100$	(μHz)
H	$1 - 5.10^4$	($\frac{\text{ppm}^2}{\mu\text{Hz}}$)
ν_{max}	$3 - 300$	(μHz)
W	$1 - 10^3$	($\frac{\text{ppm}^2}{\mu\text{Hz}}$)
T_{eff}	$4000 - 5500$	(K)
[Fe/H]	$-1.5 - 1.0$	(dex)
M	$0.4 - 3.0$	($M_\odot$)
$\Delta\nu$	$1.0 - 30.0$	(μHz)
K	$5 - 13$	(mag.)

Table 6.1: Uniform priors used for background fitting to stellar power spectra. These were the same for both RGB and RC stars. Those in the first section were used during the pre-fitting stage. All were used in the new model.

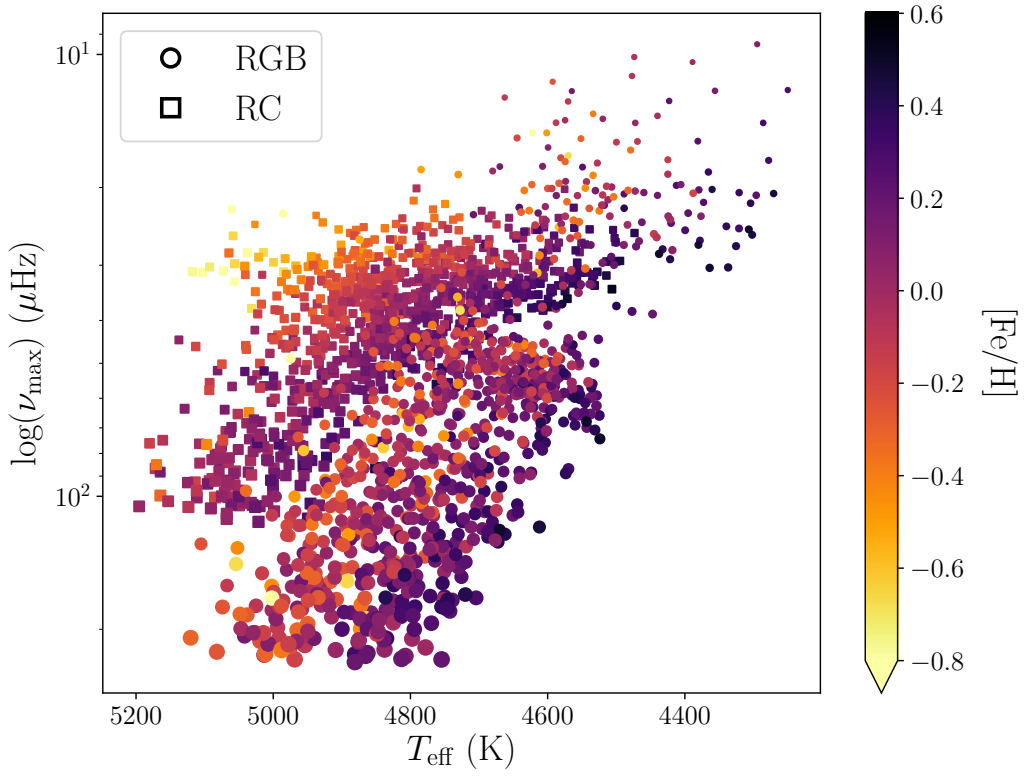


Figure 6.1: Log(ν_{max}) values against effective temperature for the 1000 RGB and 1000 RC stars making up our training set. The data are coloured by [Fe/H] and the marker size indicates the $\Delta\nu$.

from the training phase can be applied to all stars in the survey we are studying and therefore holds true for the testing phase. We now detail the components of the model and the fitting procedure.

Modelling the power spectrum

In order to extract information about a star we must fit a model to the data. This must have a complexity that is relevant to the data quality. In the optimum case we have a good enough signal-to-noise level that individual modes of oscillation can be identified and by peak-bagging we can obtain much information. However, there is plenty of information to gain from the global shape of a power spectrum and this is available for many more stars that are either less bright or have not been observed for as long. This will be the focus of this study: to constrain stellar properties using the background signal.

When modelling the global power spectrum of a solar-like oscillator there are several signals which must be accounted for. We must include terms for any convective background signal from the star itself, P_B, the modes of oscillation, P_O, as well as sources of noise, P_S. These features make up components of our model and each will now be defined.

Turbulent motion in the convective regions of solar-type stars produces not only a signal of the acoustic oscillations but also background signals. Typically, a Harvey-like profile (Harvey, 1985) is used to model the contributions from various processes to the background signal, including signal from activity or granulation (Aigrain et al., 2004; Michel et al., 2009). This profile as a function of the frequency, ν, has the form

$$\mathcal{H}(\nu; a, b, c) = \frac{\xi\, a^2/b}{1 + (\nu/b)^c}, \tag{6.1}$$

where a is an amplitude parameter, b is a characteristic frequency associated with the decay of stellar granules, and c is an exponent for which we fixed $c = 4$ as suggested by Michel et al. (2009) and Kallinger et al. (2010). The normalisation

constant, ξ, is used such that for an exponent of $c = 4$, $\xi = 2\sqrt{2}/\pi$. This relation can be understood by considering a symmetric signal pulse with an exponential rise and decay timescale, $\tau = (2\pi b)^{-1}$, for example from the rise and fall of stellar granules. The Fourier transform of this is a *super-Lorentzian* function as is described by the Harvey-like profile. Kallinger et al. (2014) compared several models for red giant spectra, with various background components, to determine how well they were able to reproduce the observations and showed that we should not represent the background signal with just a single term. They found that an appropriate model requires two super-Lorentzian components for the granulation signal, one of which encompasses the long-period signature of stellar activity. For our model the convective background contribution to the power spectrum therefore contains two such Harvey-like functions:

$$P_B(\nu) = \mathcal{H}(\nu; a_1, b_1) + \mathcal{H}(\nu; a_2, b_2), \tag{6.2}$$

where a_1/a_2 and b_1/b_2 are the rms amplitude and characteristic frequency associated with the $1^{\text{st}}/2^{\text{nd}}$ background components, respectively.

Next, the power excess due to solar-like oscillations is modulated by an envelope which in most cases is approximated as Gaussian in shape (Basu & Chaplin, 2017), although we acknowledge that this may not be true for the hottest and coolest solar-type stars. For this reason we use a Gaussian term to encompass the power excess due to modes of oscillation, with parameters to control the width, Γ, and height, H, of the envelope, and the central frequency being $\nu_{\max}$, the frequency at maximum power. The contribution from oscillations is therefore

$$P_O(\nu) = H \exp\Big(\frac{-(\nu - \nu_{\max})^2}{2\Gamma^2}\Big). \tag{6.3}$$

The final contributions we consider are from instrumental noise and shot noise. These are assumed to be Gaussian noise in the time domain and therefore we model

as a uniform background offset in the frequency domain, or white noise, with parameter W. Hence the term

$$P_S = W, \tag{6.4}$$

sets a floor for detectable oscillations and is independent of frequency. Other sources of noise, such as low-frequency drifts, need not be accounted for since the fitting range for power spectra are truncated.

Each data point in a lightcurve is collected over a discrete time interval Δt. Consequently, any oscillations or granulation within this time are averaged to produce a single value, so by taking a discrete measurement the continuous signal is convolved with a function describing the integration in time. A convolution in the time domain is simply a multiplication in Fourier space. By treating this function as a top hat with width Δt, the result is a frequency dependent attenuation of the true signal amplitude. If we assume that the cadence of observations is similar to this integration time then the attenuation factor is

$$\eta(\nu) = \mathrm{sinc}\Big(\frac{\pi}{2}\frac{\nu}{\nu_{\mathrm{nyq}}}\Big), \tag{6.5}$$

where ν_{nyq} is the Nyquist frequency ($\sim 283\mu$Hz for *Kepler* and $\sim 278\mu$Hz for TESS). This *apodisation* effect is well known and must be included in the power spectrum model (Chaplin et al., 2011b).

Combining all components and factors we define our model power spectrum as

$$\mathcal{M}(\nu) = P_S \; + \; \eta(\nu)^2 \; \Big[P_B(\nu) + P_O(\nu)\Big], \tag{6.6}$$

where $P_B(\nu)$, $P_O(\nu)$ and P_S are as defined above. Hence we have a set of parameters for our power spectrum model $\theta = \{a_1, b_1, a_2, b_2, \Gamma, H, \nu_{\mathrm{max}}, W\}$ which we seek to obtain for each star.

For pre-fitting, the background model (described in Section 6.3.1) was fitted to

individual power spectra using `CPNest`: a package to conduct sampling for Bayesian inference within Python (Veitch et al., 2017). The program requires upper and lower bounds to be specified for all parameters in the model which act as flat priors (see top section of Table 6.1). Further priors were added to guide the fitting procedure:

$$\begin{aligned} b_1 &< b_2, \\ (a_2 - a_1) &\sim \mathcal{N}(0, 50), \\ (b_2 - b_1) &\sim \mathcal{N}(30, 50), \\ \Gamma &\sim \mathcal{N}(\exp(0.85 \log \nu_{\max}) - 1.4, 20), \end{aligned} \tag{6.7}$$

The first relation attempts to limit degeneracies due to two components of the power spectrum model swapping in relative frequency position. The second and third terms are based on the scaling relations derived by Kallinger et al. (2014) and are very typical priors necessary for fitting to a power spectrum. The final prior is a typical power law relation for the mode width with $\nu_{\max}$ (Mosser et al., 2012a) along with some additional spread. Once trained our new model does not use the priors in Equation 6.7 when background fitting to unseen stars. The likelihood function used both during pre-fitting and normal fitting is also detailed in Section 6.3.1. For the pre-fitting procedure we used 200 live points for the nested sampling algorithm and each MCMC sampling chain had a maximum of 200 points.

Likelihood function

As discussed in Chapter 2 the noise properties of a power spectrum are such that the power is distributed about an underlying limit spectrum with χ^2 k degrees-of-freedom statistics which can be modelled using a Gamma distribution where $k = 2s$. Recall that the $2s$ degrees-of-freedom arises from the effects of rebinning the power spectrum where s is the number of bins that are combined into one. Referring back to Bayes theorem and Equations 2.4 and 2.5 this will set the likelihood function to

be

$$p(\hat{d}|\theta) = \prod_{i=1}^{N} \left[\frac{s^{s-1}}{(s-1)!} \frac{P_i^{s-1}}{\mathcal{M}(\nu_i)^s} e^{-\frac{sP_i}{\mathcal{M}(\nu_i)}} \right]. \tag{6.8}$$

$\mathcal{M}(\nu_i)$ and P_i are the modelled and observed power, respectively, per individual bin i. The product is taken over the entire data of length N.

It follows that the logarithmic likelihood, which is computationally easier to work with, is

$$\log p(\hat{d}|\theta) \propto -\sum_{i=1}^{N} \left[s \log \mathcal{M}(\nu_i) + \frac{s P_i}{\mathcal{M}(\nu_i)} - (s-1) \log P_i \right]. \tag{6.9}$$

The MVN model

Our model assumes that a star can be described using a set of global stellar properties and a power spectrum defined by the parameters detailed above. Combined, these make up the set of free parameters we seek to find for each star: $\mathbf{\Theta} = \{T_{\rm eff}, [\rm Fe/H], M, \Delta\nu, K, a_1, b_1, a_2, b_2, \Gamma, H, \nu_{\rm max}, W\}$ where $T_{\rm eff}$ is the stellar effective temperature, [Fe/H] the metallicity, M the mass, $\Delta\nu$ the large frequency separation, and K the *Kepler* magnitude. Since we already have these quantities for thousands of stars from *Kepler* and APOGEE (or can easily obtain them in the case of the background model parameters) we can construct a prior relating the parameters within $\mathbf{\Theta}$ in a data driven way, where the prior is taught on reference data, also known as the *training* data set.

Statistical analysis of these data reveals that there is a correlation between these quantities. We can therefore describe the population statistics using a MVN distribution comprising a mean and covariance. This is similar to an approach used by Lieu et al. (2017) to derive galaxy group and cluster mass from weak gravitational lensing measurements. The covariance function will contain the intrinsic scatter on parameters as well as the covariances between them. The MVN will act as a map to relate stellar properties to changes in the shape of a power spectrum.

However, we note that between some of the parameters there is much more extreme correlation than others, primarily involving $\nu_{\max}$ or K. For reasons discussed later, we choose to first remove some of this strong correlation by calculating the residual between logarithms of the parameters and an estimated version, i.e.

$$R_j = \begin{cases} \Theta_j - \mathcal{F}(\nu_{\max}, K) & \text{if } j \text{ refers to [Fe/H]}, \\ \log \Theta_j - \mathcal{F}(\nu_{\max}, K) & \text{otherwise.} \end{cases} \tag{6.10}$$

where j denotes one of the parameters in $\boldsymbol{\Theta}$ excluding $\nu_{\max}$ and K, and $\mathcal{F}$ is the function used to approximate the parameters. Note that [Fe/H] is already on a logarithmic scale so we need not take the logarithm to calculate its residual. The resulting residuals no longer have the wide range of covariances we see between the parameters themselves since extreme correlations have been removed. It is the residuals which are then modelled using MVN distributions rather than the parameters themselves. We found that this technique was required for convergence when fitting a MVN to the training data, and works since the covariance function no longer has both strongly positive and negative values to fit at the same time.

Similar to Kallinger et al. (2014) we use a power law relation to estimate the parameters such that $\mathcal{F}$ in logarithmic form is given by

$$\mathcal{F}(\nu_{\max}, K) = c_j\, K + d_j \;+\; e_j\, \log \nu_{\max} + f_j. \tag{6.11}$$

$\mathcal{F}$ was found by a simple least squares fit of this straight line relation to the training set in order to extract c_j, d_j, e_j and f_j for each parameter. These were then fixed for the rest of the study.

Our model assumes that the distribution of parameters, or at least their residuals, $\boldsymbol{R}$, of length r, for a population of stars can be described by a MVN distribution with a mean r-vector, $\boldsymbol{\mu}$, and $r \times r$ covariance matrix, $\boldsymbol{\Sigma}$. The probability density

of this distribution is

$$p(\boldsymbol{R}|\boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{r/2}\ |\boldsymbol{\Sigma}|^{1/2}} \exp\Big[-\frac{1}{2}\ (\boldsymbol{R} - \boldsymbol{\mu})^{\mathrm{T}}\ \boldsymbol{\Sigma}^{-1}\ (\boldsymbol{R} - \boldsymbol{\mu})\Big]. \tag{6.12}$$

Since $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ describe the population they are called *hyperparameters*.

Bayesian mixture model

Our model is applicable to evolved stars similar to those in our reference data set. These consist of a mix from two different populations: RGB and RC stars. RGB stars are those which have evolved off the main sequence, beginning their first ascent up the red giant branch, or up the asymptotic giant branch, and have an inert helium core surrounded by a hydrogen burning shell. Conversely, RC stars are those beyond the tip of the red giant branch which have started to burn helium in their cores. Although they have very different internal conditions the surface characteristics of these two types of star can be very similar, particularly the $T_{\rm eff}$, luminosity, and $\log g$. This leads to them occupying very similar regions in an HR diagram and therefore being notoriously difficult to distinguish using classical observations. Differentiating between these populations has been the focus of many studies (e.g. Bedding et al. 2011; Kallinger et al. 2012; Stello et al. 2013; Elsworth et al. 2017; Hon et al. 2018; Kuszlewicz et al. 2020).

Our method uses a mixture of two population models to build a prior. These are two separately trained MVN distributions described by $\boldsymbol{\mu}_{\rm RGB}$ and $\boldsymbol{\Sigma}_{\rm RGB}$ for the RGB stars, and $\boldsymbol{\mu}_{\rm RC}$ and $\boldsymbol{\Sigma}_{\rm RC}$ for the RC stars. We draw simultaneously from both models and use a weighting factor, λ, to balance the contribution from each. As explained below, the prior is split into three pieces. One part of the prior is the probability of obtaining the stellar properties and parameters, or more accurately their residuals, given the MVN distributions for RGB and RC stars as described by

the mixture, i.e.:

$$p(\boldsymbol{R}, \lambda|\boldsymbol{\mu}_{\rm RGB}, \boldsymbol{\Sigma}_{\rm RGB}, \boldsymbol{\mu}_{\rm RC}, \boldsymbol{\Sigma}_{\rm RC}) = \lambda\, p(\boldsymbol{R}|\boldsymbol{\mu}_{\rm RGB}, \boldsymbol{\Sigma}_{\rm RGB}) + (1-\lambda)\, p(\boldsymbol{R}|\boldsymbol{\mu}_{\rm RC}, \boldsymbol{\Sigma}_{\rm RC}). \tag{6.13}$$

As $\lambda \to 1$ draws from the prior are dominated by the RGB model, and as $\lambda \to 0$ the RC model dominates. A λ value somewhere between 0 and 1 is uncertain about the evolutionary state of the star.

The aim of this work is to be able to extract stellar properties directly from asteroseismic observations using a model built in a data driven way. To do so we use Bayes theorem, introduced in Chapter 1 but adapted here, where the posterior probability density on the parameters of interest ($\boldsymbol{\Theta}$ and λ) given the observations is

$$\begin{aligned} p(\boldsymbol{\Theta}, \lambda|\hat{d}, \mu_{\rm RGB}, \boldsymbol{\Sigma}_{\rm RGB}, \mu_{\rm RC}, \boldsymbol{\Sigma}_{\rm RC}) &\propto p(\hat{d}|\boldsymbol{\Theta})\; p(\boldsymbol{\Theta}, \lambda|\mu_{\rm RGB}, \boldsymbol{\Sigma}_{\rm RGB}, \mu_{\rm RC}, \boldsymbol{\Sigma}_{\rm RC}), \\ &\propto p(\hat{d}|\boldsymbol{\Theta})\; p(\boldsymbol{\Theta})\; p(\lambda)\; p(\boldsymbol{R}\lambda|\boldsymbol{\mu}_{\rm RGB}, \boldsymbol{\Sigma}_{\rm RGB}, \boldsymbol{\mu}_{\rm RC}, \boldsymbol{\Sigma}_{\rm RC}). \end{aligned} \tag{6.14}$$

As before, $p(\hat{d}|\boldsymbol{\Theta})$ is the likelihood probability of the data given the parameters $\boldsymbol{\Theta}$ and is defined below. We have split the prior probability into a flat prior over all $\boldsymbol{\Theta}$, $p(\boldsymbol{\Theta})$ (as detailed in Table 6.1), a flat prior between 0 and 1 for our weighting parameter λ, $p(\lambda)$, and the MVN distributions as a mixture prior over the residual parameters, $p(\boldsymbol{R}, \lambda|\mu_{\rm RGB}, \boldsymbol{\Sigma}_{\rm RGB}, \mu_{\rm RC}, \boldsymbol{\Sigma}_{\rm RC})$. For numerical stability it is better to work in terms of logarithmic probabilities since the product taken over a large number of small probabilities can quickly go beyond the computational numerical precision. Therefore the priors and likelihood will be defined in terms of their logarithms.

The multivariate probability density function, as defined in Equation 6.12, means that we must take the sum of exponentials which is computationally unstable. In-

stead, we define the logarithm of the mixture prior as

$$\log p(\boldsymbol{R}, \lambda|\boldsymbol{\mu}_{\rm RGB}, \boldsymbol{\Sigma}_{\rm RGB}, \boldsymbol{\mu}_{\rm RC}, \boldsymbol{\Sigma}_{\rm RC}) = \log\left(\sum_k \exp(a_k)\right), \tag{6.15}$$

where

$$a_k = \log(\lambda_k \times\ p(\boldsymbol{R}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)), \tag{6.16}$$

$$= \log\lambda_k\ -\ \frac{1}{2}\log\left((2\pi)^n|\boldsymbol{\Sigma}_k|\right)\ -\frac{1}{2}(\boldsymbol{R}-\boldsymbol{\mu}_k)^T\,\boldsymbol{\Sigma}_k^{-1}\,(\boldsymbol{R}-\boldsymbol{\mu}_k), \tag{6.17}$$

with k =RGB/RC, and $(\lambda_{\rm RGB} + \lambda_{\rm RC}) = 1$.

A probabilistic graphical model of our mixture model is shown in Figure 6.2.

6.3.2 Training step

In order to build a generative model we had to first train the MVN distribution to obtain the suitable values of $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$. This was done separately for RGB and RC stars. We assume the training data have been drawn from the same population distributions, with the same statistics, as unseen stars from which we would like to infer stellar properties. For training, the probability density for each model can be written as

$$p(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k|\boldsymbol{R}_k^{\rm T}) \propto p(\boldsymbol{\mu}_k)\,p(\boldsymbol{\Sigma}_k)\,p(\boldsymbol{R}_k^{\rm T}|\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k). \tag{6.18}$$

$p(\boldsymbol{\mu}_k)$ and $p(\boldsymbol{\Sigma}_k)$ are the hyperpriors for the mean and covariance matrix of the k population, and $\boldsymbol{R}_k^{\rm T}$ are the residuals from the corresponding training data. The superscript T simply denotes that we are using the training data.

Hyperpriors on $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$

Since $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$ are hyperparameters they require hyperpriors. Below we describe the hyperprior for a generic $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ but it is relevant for either the RGB or RC models.

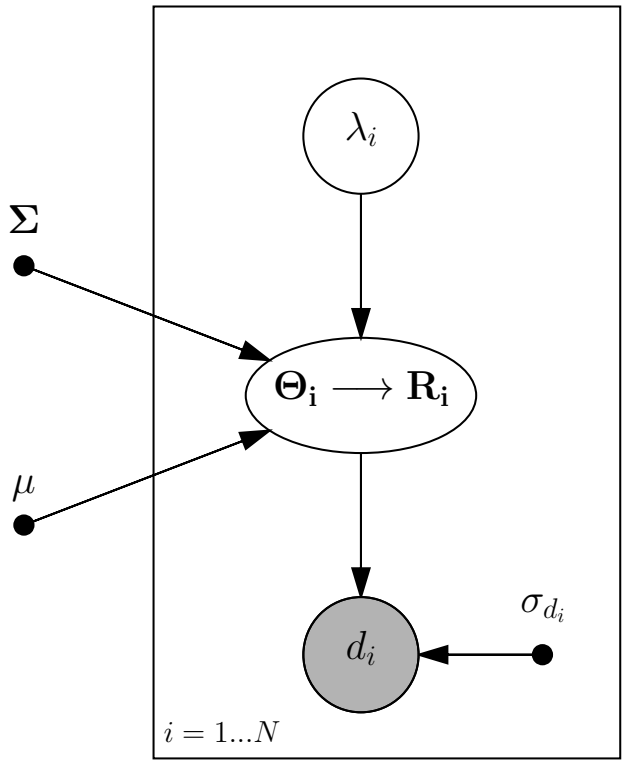


Figure 6.2: Probabilistic graphical model showing the model defined in Equation 6.14. Shaded circles indicate observed quantities, while dots are fixed parameters that are not probabilistically modelled, and open circles indicate free parameters to be determined by the fitting procedure. The parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ define the shape of the model prior which controls the distributions of the quantities in $\boldsymbol{\Theta}$. The parameters of each star describe a model power spectrum which combine with the observational uncertainty to control the distribution of bulk stellar properties. d_i represents the observations of a datum i for a total of N data. (Made using `daft` available at `docs.daft-pgm.org`.)

For a prior on the mean vector we use a normal distribution with a large variance to act as a weakly informative prior:

$$p(\boldsymbol{\mu}) = \frac{1}{\sqrt{2\pi\boldsymbol{\sigma}^2}} \exp\Big[-\frac{\boldsymbol{\mu}^2}{2\boldsymbol{\sigma}^2}\Big], \tag{6.19}$$

where we chose the variance $\boldsymbol{\sigma}^2 = 10$. This is centred on zero since the residuals should all be approximately zero.

Forming a prior for the covariance matrix is more complicated since it must

be positive-definite and symmetric. This makes it difficult to sample so we split building the prior into stages. Following the recommended approach of The Stan Development Team (2018) we can decompose $\mathbf{\Sigma}$ into an r-dimensional scale vector, $\boldsymbol{\sigma}$, and an $r \times r$ correlation matrix, $\mathbf{\Omega}$, such that:

$$\mathbf{\Sigma} = \text{diag}(\boldsymbol{\sigma}) \; . \; \mathbf{\Omega} \; . \; \text{diag}(\boldsymbol{\sigma}), \tag{6.20}$$

where diag() refers to forming a diagonal matrix for the calculation. A Lewandowski-Kurowicka-Joe (LKJ) distribution prior (Lewandowski et al., 2009) is typically placed on the correlation matrix which has the density

$$p(\mathbf{\Omega}|\eta) \propto |\mathbf{\Omega}|^{\eta-1}, \tag{6.21}$$

where η is a hyperparameter that controls the size of the correlations between the components of $\boldsymbol{R}$. For $\eta = 1$ the above relation collapses to a uniform distribution. If $\eta > 1$ the magnitude of correlations decreases, whereas for $\eta < 1$ then stronger correlation between $\boldsymbol{R}$ is favoured. Our model was implemented in `PyMC3` which supports a LKJ prior. Prior distributions were also provided for $\boldsymbol{\sigma}$, which must be positive, and η. For both we chose weakly informative log-normal priors:

$$\boldsymbol{\sigma} \sim \log \mathcal{N}(\log(0.03), 2), \qquad \eta \sim \log \mathcal{N}(\log(1.0), 0.1), \tag{6.22}$$

i.e. weakly preferring an identity matrix as implemented by Lieu et al. (2017). Since inverting $\mathbf{\Sigma}$ is an unstable and inefficient process, the `PyMC3` method in practice places a prior on the Cholesky decomposition of the covariance matrix

$$\mathbf{\Sigma} = LL^{\top}, \tag{6.23}$$

where L is a lower-triangle matrix. It is therefore possible to place priors over $\boldsymbol{\sigma}$

and η which combine to put a prior on L so that the underlying correlation matrix $\mathbf{\Omega}$, follows a LKJ distribution. From this, a prior on $\mathbf{\Sigma}$ can be reconstructed.

Fitting procedure

The posterior given in Equation 6.18 was estimated using `PyMC3`'s Hamilton Monte Carlo (HMC) algorithm (Duane et al., 1987) implementation via the No-U-Turn Sampler (NUTS) (Hoffman & Gelman, 2011). HMC is one of several sampling techniques used to estimate a posterior distribution. As with other sampling methods for posterior estimation, this involves a random walk through parameter space to sample the target distribution. However, contrary to standard Monte Carlo methods, HMC includes a momentum term to determine the trajectory of each of the walkers. In this way steps in parameter space are less random since each step depends on the gradient of the logarithm of the posterior and so it is more sensitive to its shape. A common analogy is to think of a frictionless ball rolling around a trough and mapping its profile of minima and maxima; the ball follows curved lines due to it's momentum. HMC conducts a more rapid and efficient walk than, for example, a Metropolis Hastings algorithm, leading to a faster convergence. After a certain number of steps the walker is 'kicked' in a random direction to follow a new trajectory ensuring there is still some randomness when exploring the parameter space. To continue the ball analogy, every so often the ball is flicked in a new random direction.

Although this method ensures efficient randomness, since HMC includes momentum, walkers move along curved paths and will eventually loop back on themselves if allowed to travel for long enough (without energy loss the ball will circle back to its initial position). NUTS is a method of proposing steps around parameter space by controlling how long to follow the curved trajectory before receiving a new kick. NUTS adaptively finds the optimum number of steps to take along the trajectory by determining when the path will turn back on itself and stopping the route before it does so. This class of samplers is particularly useful for more complex target

distributions, e.g. those with multiple modes, and those with a larger number of dimensions.

We are interested in finding the underlying MVN model which best described our population of stars. However, when training our model any calculation of the likelihood of obtaining our observed data given the model parameters must be computed using the total covariance. In order to take into account the uncertainty on our training data we set up this total model covariance comprised of an intrinsic covariance (describing the underlying population distribution) in combination with the typical variance we expect for the parameters, i.e.

$$\boldsymbol{\Sigma_T} = \boldsymbol{\Sigma} + \boldsymbol{\Sigma_O}, \tag{6.24}$$

where $\boldsymbol{\Sigma}$ is the intrinsic covariance and $\boldsymbol{\Sigma}_O = \mathrm{diag}([\sigma \boldsymbol{R}^2])$ is the variance arising from measurement uncertainty. Assuming un-correlated measurements, the latter is a diagonal matrix made from the uncertainties associated to each residual, $\sigma \boldsymbol{R}$.

We ran 4 chains with 1000 tuning steps followed by 1000 iterations kept for posterior inference. Convergence was assured by checking the trace plots of each parameter were sensible and that the Gelman-Rubin statistic (Gelman & Rubin, 1992) was $\hat{R} < 1.01$. We also ensured that there were no numerical divergences which would arise if the sampler reaches regions of parameter space where it could not explore sufficiently. If many divergences occur this may result in an insufficient number of samples to estimate the posterior and implies that there could be biases in the chains. Ways to remove divergences are to increase the acceptance rate or to re-parameterise the model. For this work we used an acceptance probability of 0.99.

6.3.3 Testing step

During the training phase we used known labels of power spectrum parameters and stellar properties to separately train two MVN models to find the hyperparameters that best describe the RGB and RC populations. For the testing phase we instead

compute the inverse where the stellar properties are not known but can now be inferred using the trained model and observations of the test stars. Whereas the training step did not use the mixture model (while it is learning the two separate models), this is implemented in the testing phase.

Our method assumes that the same generative model is true for all other stars in the same population as those from the training set. We implement the mixture prior using mean values from the posteriors of the trained $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. As previously mentioned this is paired with uniform bounds on all parameters for $p(\boldsymbol{\Theta})$ (as required by `CPNest`) and a flat prior on λ, but otherwise no other functional prior. Combining these priors with the likelihood as described above into Equation 6.14 we obtain the posteriors on $\boldsymbol{\Theta}$ and λ.

Fitting Procedure

As with the background fit we used to obtain the training data sets, we implemented `CPNest` to estimate the joint posterior of $\boldsymbol{\Theta}$ and λ via nested sampling. Although the primary output of this algorithm is the model evidence, it also produces samples from the posterior over which we can marginalise to obtain estimates of the stellar properties. Nested sampling methods have been found to perform better than other sampling techniques, such as MCMC, when the posteriors are multimodal or have degeneracies or phase transitions present. We expect our situation to be like this since RGB and RC overlap in some regions of parameter space.

For the fitting procedure we used 1000 live points and maximum number of steps in the MCMC chain of 5000. We found that the minimum number of live points was 1000, below which there were insufficient samples for estimating the posterior. We also found that using less than 2000 MCMC steps resulted in the parameter space not being adequately explored when replacing deactivated points and using 5000 steps gave the best performance for distinguishing between RGB and RC stars. Numbers greater than the chosen values saw no improvement in the fitting. The process was terminated when the change in the logarithm of the evidence was less

than a tolerance level of 0.1, at which point the fitting was deemed to have converged.

6.4 Results

6.4.1 Application to artificial stars

We first tested our model on simulated data. Artificial stellar properties and periodogram parameters were generated for a collection of 550 stars by taking samples from a pre-defined MVN distribution. The mean was set equal to the typical values for APOKASC-2 RGB stars, and an underlying covariance matrix was constructed by defining a Cholesky decomposed matrix, L, where $\boldsymbol{\Sigma_U} = L.L^\top$. L was chosen such that the artificial data produced from the MVN distribution resembled typical APOKASC-2 data for RGB stars. To generate artificial data the total covariance was used. This was made by combining the underlying covariance matrix, $\boldsymbol{\Sigma_U}$, with observational uncertainty, i.e.

$$\boldsymbol{\Sigma_T} = \boldsymbol{\Sigma_U} + \boldsymbol{I\sigma}^2, \tag{6.25}$$

where $\boldsymbol{\sigma}$ are typical observational uncertainties for APOKASC-2 RGB stars. Note that this is similar to Equation 6.24 but this covariance matrix is for all parameters in $\boldsymbol{\Theta}$, not just residuals. For each toy star a periodogram was simulated from the artificial power spectrum parameters. Noise was added to each spectrum following the description in Anderson et al. (1990) before being rebinned identically to the real stellar spectra. The data were split into a training set of 500 'stars' and 50 for testing.

As described in Section 6.3.1, our model prior was first trained to capture the relation between stellar properties and periodogram parameters for our toy stellar population. Figure 6.3 shows the posterior estimates for the trained covariance matrix in solid blue along with the true underlying covariance used to generate the data indicated by the dashed orange line. The absolute values of the covariance elements

Figure 6.3: Posteriors of each element in the underlying covariance matrix $\mathbf{\Sigma}_{\boldsymbol{U}}$ after training on 500 toy stars is shown in blue. The dashed orange line indicates the true intrinsic covariance which is known for our artificial data.

are unimportant; the main point is that the posteriors are consistent with the input underlying covariance which we know for our artificial data. This indicates that the model was successfully trained to identify the underlying population statistics. After training we then had a MVN distribution for our artificial population which could act as a prior when fitting to 'new' toy stars.

Next, the model was fit to the 50 new stars in the testing data set. Figure 6.4 shows that estimates of the fitted stellar properties show good agreement with the true values, known since the data were simulated. The rms between our model

predictions and known values are of a similar size to the average uncertainty on predictions (the precision) and the bias values are significantly smaller.

6.4.2 Application to real stars

Our model was applied to real data, this time by training two separate MVN priors each on 1000 RGB or RC stars. Taking the mean values from the posteriors of the distributions' parameters we formed our mixture prior for the testing step. Our model was applied to 2322 unseen test stars to obtain posterior estimates of stellar properties.

Figures 6.5 & 6.6 show a typical fit to an individual test star (KIC7936407) and the marginalised posterior estimates. In Figure 6.5 the re-binned observations are in blue and the fitted model is overlaid in orange with the shaded region showing the 5σ uncertainty. Also included as dashed lines are four of the components making up the power spectrum model; the two Harvey-like profiles, $\mathcal{H}_1$ and $\mathcal{H}_2$, the Gaussian mode envelope, P_O, and the white noise component, W. Figure 6.6 shows the posterior estimates for each bulk stellar parameter and λ. The blue line and shaded region indicate the true value (the quoted APOKASC-2 value aside from $\nu_{\rm max}$ which is from our pre-fitting stage), along with its uncertainty. For most parameters this blue region is narrow compared the posterior apart from $T_{\rm eff}$ and $\nu_{\rm max}$.

Our method is able to infer stellar property estimates from asteroseismic observations of the global power spectrum. The left-side panels of Figure 6.7 show the estimated mean predicted values for newly fitted stars against truth values for our test data set. The latter were taken from the APOKASC-2 catalogue, excluding $\nu_{\rm max}$ for which we used values from our pre-fitting stage. Blue and orange points show the RGB and RC results respectively. The right-side panel shows the corresponding histogram of the z-score which is equal to the difference between our predictions and APOKASC-2 values divided by the total uncertainty, calculated as the quadrature sum of uncertainties. The solid black curve represents an $\mathcal{N}(0, 1)$ distribution for

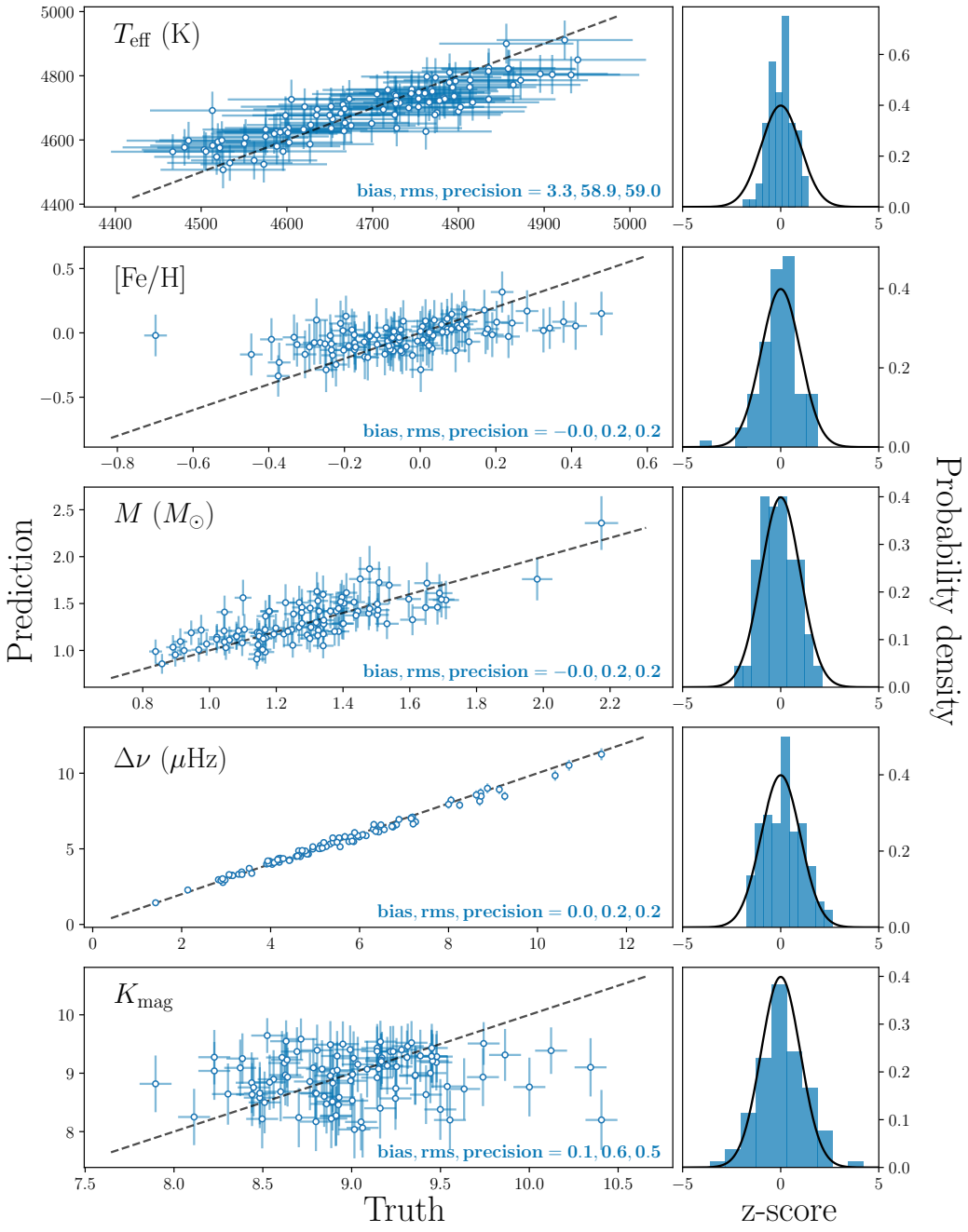


Figure 6.4: The predictions for bulk stellar properties of the 50 toy test stars (left) and corresponding histogram (right) showing the z-score equal to predicted-truth values divided by the joint uncertainty. For reference the solid black curve indicates a $\mathcal{N}(0, 1)$ distribution.

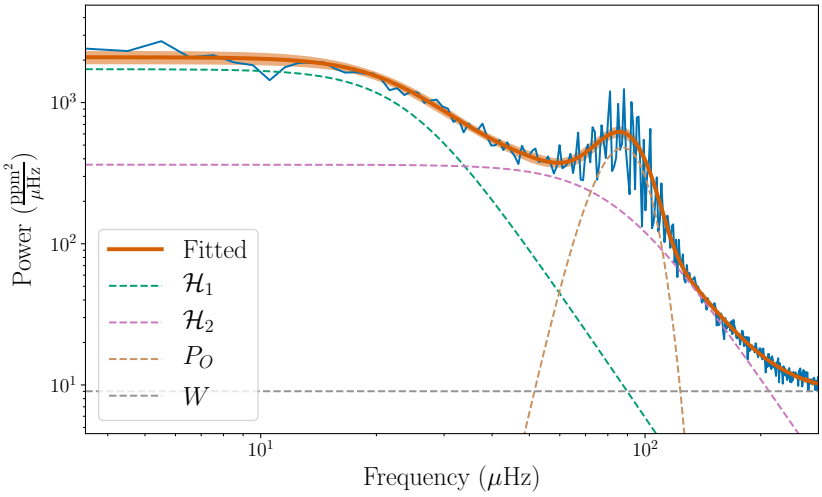


Figure 6.5: The fitted background power spectrum model for an example test star (KIC7936407) shown in solid orange with a shaded region to indicate the 5σ uncertainty as a function of frequency. The data, in blue, has been re-binned to a bin width of 1μHz. Included as dashed lines are the two Harvey-like components, $\mathcal{H}_1$ and $\mathcal{H}_2$, the Gaussian mode envelope, P_O, and the white noise component, W, to the model.

reference. Any bias in the results can be seen as an offset between the centre of this curve and the histogram peak and is a reported statistic in each plot. For predictions which have the same level of uncertainty and intrinsic scatter the histograms and $\mathcal{N}(0, 1)$ would match. If the predictions have uncertainties smaller than those of the truth values then the results distribution would be narrower than the black curve since the residuals would be closer to zero.

Our results are shown to be consistent with truth values. We report rms statistics similar to the uncertainty in our predictions and have biases which are 3 to 9 times smaller for RGB stars and 3 to 16 times smaller for RC. The uncertainties from our estimates of stellar properties are, in some cases, larger than those of the APOKASC-2 catalogue with ~ 0.2 dex in [Fe/H], $\sim 15\%$ in mass and $\sim 4\%$ in $\Delta\nu$. In the case of $\Delta\nu$, the reported APOKASC-2 values are a weighted mean over results from five different pipelines. Their uncertainties are an average taken over the dispersion in each pipeline's measurements around their normalised mean. Pinsonneault et al. (2018) found this to usually be smaller than the predicted uncertainties of each

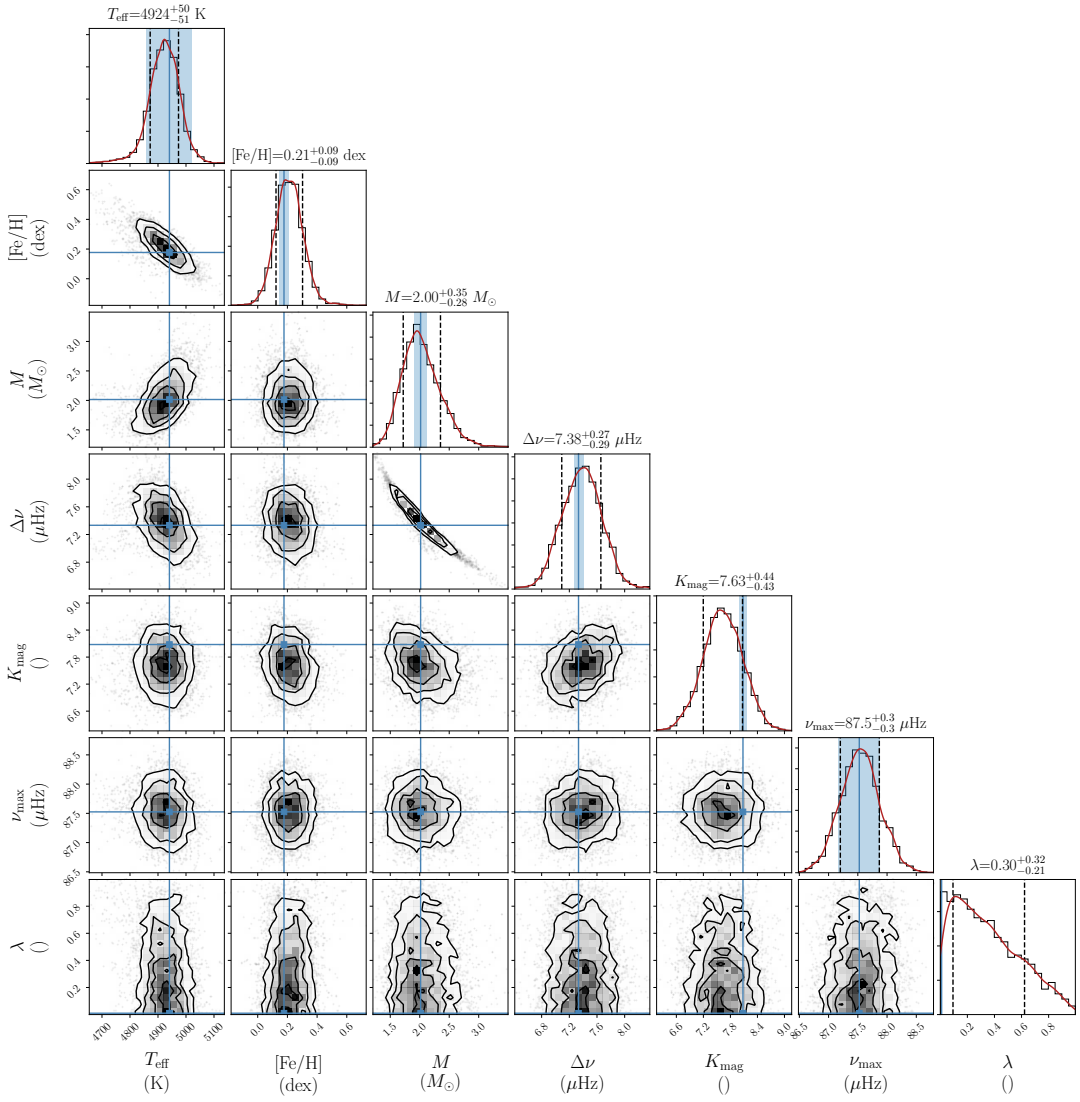


Figure 6.6: The posterior distributions for bulk stellar properties and λ for an example test star (KIC7936407). The red curve is a Gaussian kernel density estimate to guide the eye to an approximate probability density function for each posterior. The solid blue line and shaded blue region indicate the truth values and associated uncertainties as taken from the APOKASC-2 catalogue, except $\nu_{\rm max}$ which is taken from our pre-fitting analysis. The 68% credible intervals are indicated by vertical dashed lines and are used as the upper and lower uncertainty estimates labelled above each histogram along with the median.

pipeline which may explain the smaller uncertainty than those produced by our model. For our method there is also likely to be a lack of constraint when fitting to a global power spectrum rather than to individual oscillation modes or to a different representation of the periodogram data (e.g. an auto-correlation function). The larger uncertainty on $\Delta\nu$ is commensurate with a larger uncertainty on mass which is calculated via a scaling relation with $\Delta\nu$. The [Fe/H] values in the APOKASC-2 catalogue are derived using spectroscopic data and hence will have smaller associated uncertainties than we are able to obtain using asteroseismic observations alone. Our estimates of effective temperature, however, have a similar level of uncertainty to APOKASC-2 values for RGB stars ($\sim$ 70K) and are generally narrower for RC stars ($>$ 56K). Our uncertainty predictions on $\nu_{\rm max}$ are similar to the quoted APOKASC-2 values if not slightly narrower with typical fractional uncertainties of $\sim$ 3% for RGB stars and $\sim$ 6% for RC. These smaller uncertainties are more consistent with those predicted by the only Bayesian pipeline from the collection of five used in compiling the APOKASC-2 catalogue. In order to more tightly constrain these stellar properties using our model we require additional observations, such as those from spectroscopy, or we could include more detail in the model, such as structure within the mode envelope.

We found that our model can largely distinguish between RGB and RC stars. For those that are more difficult to identify this is reflected in the posterior distribution for λ being more uniform in shape since the model cannot decide between $\lambda = 1$ for an RGB star or $\lambda = 0$ for an RC. The estimated uncertainty on λ is hence wider for these stars. Figure 6.8 shows the spread in inferred λ values for the test set of stars which we took as the mode of each posterior. The colour indicates the true evolutionary state, although it is worth noting when comparing with our results that these were taken from the APOKASC-2 catalogue and have no reported uncertainties. Most stars in the test data set were correctly identified (orange points for RC stars lying near 0.0 and few around 1.0). For stars that were difficult to

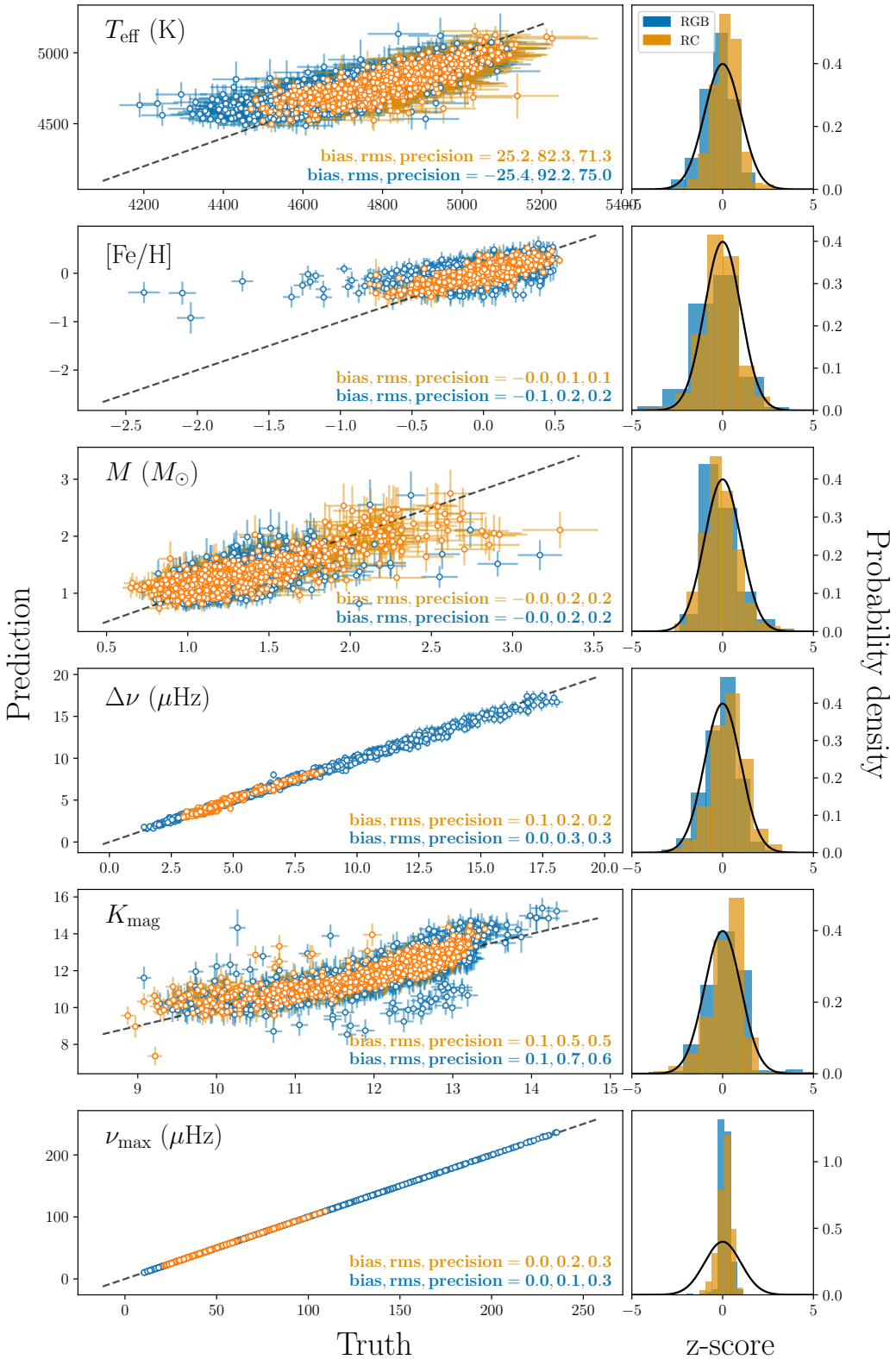


Figure 6.7: The predictions for bulk stellar properties of the 1469 RGB and 853 RC test stars (left) and corresponding histogram (right) showing the z-score equal to predicted-truth values divided by the joint uncertainty. For reference the solid black curve indicates a $\mathcal{N}(0, 1)$ distribution.

distinguish (i.e. $\lambda \sim 0.5$) the uncertainty was larger.

As an approximate metric for our model's ability to predict evolutionary state we computed the percentage accuracies in the intervals $\lambda > 0.9$ (i.e. RGB) and $\lambda < 0.1$ (i.e. RC) to be 94% and 82%, respectively. It is worth noting that the calculation for this metric uses only the median values of the λ distributions. Therefore it does not incorporate that the RGB stars for which the model incorrectly predicts $\lambda \sim 0.1$ generally have much larger uncertainties than those where the model correctly predicts RC stars as having $\lambda \sim 0.1$. This can be seen in the left hand region of Figure 6.8 where at low λ values the orange points cluster lower than the blue on the y-axis (i.e. they have smaller uncertainties). Since the accuracy metric is not computed using estimated λ uncertainties more uncertain predictions are not down-weighted compared to more accurate predictions in the percentage calculation, hence the low accuracy of 80-90%. For this reason any preliminary classification using our model should be considerate of the uncertainty on the predictions.

The most sensitive property to the choice of evolutionary stage is $T_{\rm eff}$ as expected. Our results also imply that there is a sensitivity to metallicity in the background power spectrum, although not as strong as suggested by Corsaro et al. (2017). We are able to infer metallicity from the power spectra of RC stars, however, for RGB stars there is insufficient information in the asteroseismic observations alone. It could simply be that the stars whose metallicity are most poorly derived are those from regions of parameter space that were sparsely populated by the training data. One way to resolve this is to fill the training set with more stars that populate this region in [Fe/H]. This result is, however, not altogether surprising since the evolution of RC stars is much more metallicity dependent than that of RGB stars (e.g. Hekker et al., 2017).

For evolved stars there is a well known relationship between mass and age. This means mass is an interesting quantity since if well constrained it can lead to good age estimates. It can be seen from the example joint posterior in Figure 6.6 that

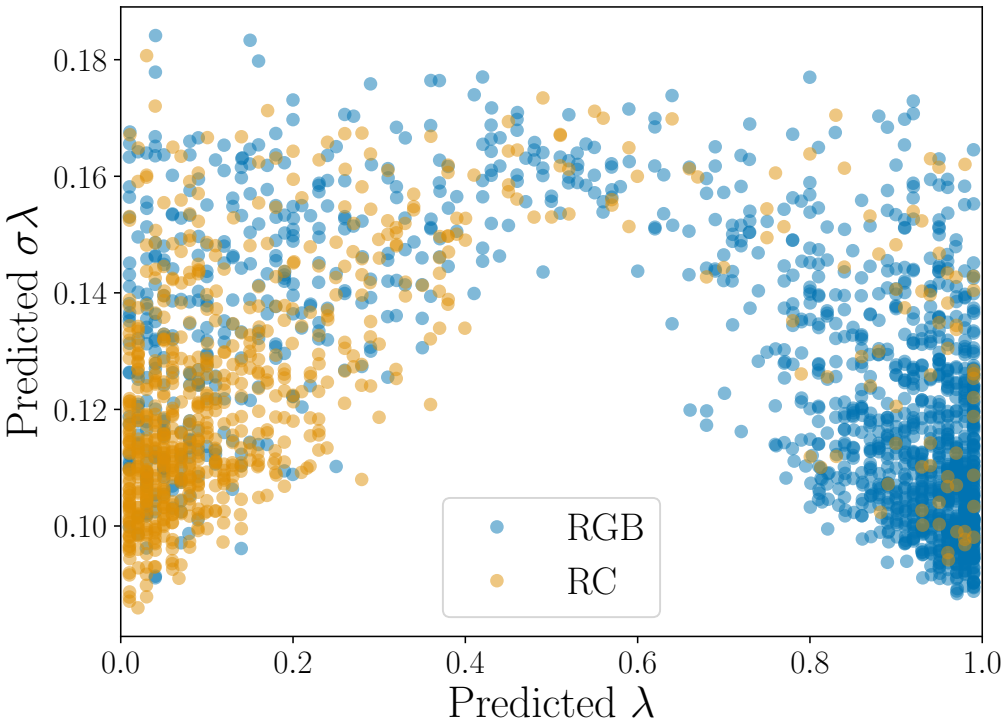


Figure 6.8: The inferred evolutionary stage of each test star as determined by the λ parameter where $\lambda = 1.0$ indicates a RGB star and $\lambda = 0.0$ a RC star. The coloured points specify the true type of star which is known for the test data.

there is a correlation between $\Delta\nu$ and mass. Since our estimated uncertainty is much larger than typical uncertainties (e.g. from APOKASC-2), and provided we can get a more tightly constrained estimate on $\Delta\nu$, we can use it to condition our posterior allowing us to constrain the estimate for mass further. If we were to incorporate additional information, such as a better estimate on $\Delta\nu$ from `PBJam` (Nielsen et al., 2021) with an uncertainty of 0.5%, then our marginalised posterior can be tightened correspondingly. By applying this condition on the posteriors obtained using our method, this enables mass estimates to reach a precision of $\sim 6\%$ for RGB stars and $\sim 5\%$ for RC. Along with good [Fe/H] estimates from spectroscopy (perhaps via *The Cannon*) this implies one can in principle obtain age estimates to $\sim 19\%$ and $\sim 16\%$ for RGB and RC stars, respectively, since $\sigma_{\rm Age} \sim 3.2\sigma_{\rm mass}$ for evolved stars (Davies & Miglio, 2016). This aligns with the expectations from literature for what should be possible using asteroseismology (Lebreton & Montalbán, 2009; Silva Aguirre et al., 2015; Miglio et al., 2020, e.g.). Obtaining precise ages is of interest to many research fields beyond asteroseismology including the study of exoplanets, stellar evolution and Galactic archaeology.

6.5 Conclusion

Using data driven techniques we were able to derive a generative model that describes the RGB and RC populations well enough to extract bulk stellar properties from their global power spectrum. We trained a multivariate normal distribution to learn the covariance relation between stellar properties and the shape of a power spectrum for populations of evolved stars observed by *Kepler*. As such, aside from constructing the training set with properties obtained from stellar modelling, we are not explicitly reliant on physical stellar models. Our model was trained on data constructed using bulk properties from the APOKASC-2 catalogue and periodogram parameters derived from pre-fitting to asteroseismic observations of the training stars. Unlike *The Cannon*, our method takes into account observational

uncertainties on the training data. Once separate distributions were obtained for RGB and RC populations we then fixed the two models and combined them to form a mixture prior. This was used for background fitting to unseen stars without requiring foreknowledge of the evolutionary stage. Fitting was therefore guided by the population hyperparameters.

We have demonstrated that with our method it is possible to derive bulk stellar properties directly from asteroseismic observations alone by implementing transfer learning. Our method relies on the training data being a good representation of the population of evolved stars and that it covers parameter space well. We assume that our generative model holds true for any further objects of the same type, i.e. they are samples from same stellar populations. We also rely on the notion that stars with same stellar properties (at least for those considered in our study) will have power spectra that are identical except for observational uncertainties or intrinsic scatter. Hence with these assumptions we can fix the model after training and use it as prior for the proceeding stage.

We found that there is enough information in a global background power spectrum to constrain some fundamental properties and to distinguish between observations of an RGB or RC star. Our estimates were consistent with those of APOKASC-2 and in the case of effective temperature our predictions have smaller uncertainties. In particular, being able to obtain estimates for $T_{\rm eff}$, and in some cases [Fe/H], from asteroseismic observations alone would be useful to the stellar community. These are quantities typically required as inputs to stellar modelling pipelines and usually must be obtained using external sources of observations. With our method this is not necessary, although we should note that our estimates are still stellar model dependent since we train on APOKASC-2 data.

Our method requires fewer stages of analysis and is overall faster than those which involve fitting stellar models to observations. This is because the model has already learnt the map between variables. It takes longer than *The Cannon* because

we make posterior estimates on parameters rather than purely a maximum likelihood estimate. Hence we obtain errors on estimates at the cost of more computationally expensive inference. Another advantage of our method is that once the model has been trained, applying a background fit to observations requires no user input for the prior since it is based on population statistics.

The basic idea of this method is to extract as much information as possible from the granulation signal and mode envelope to derive stellar characteristics. For some properties our estimates have larger than typical uncertainty which implies that there is insufficient information to provide better constraints from the asteroseismic observations alone. One possible solution to this is to expand the model to more fully describe the shape of a power spectrum, for example to include individual oscillatory modes. However, this increase in dimensionality will make computation more expensive. One would also have to be careful when constructing the training data set to ensure that it sufficiently populates the much larger parameter space. An alternative could also be to add other sources of data. There is further information for each star available by using spectroscopic data so there is scope to connect an asteroseismic model such as this with a spectroscopic approach like *The Cannon*. As discussed above, supplying a more tightly constrained, external measure of $\Delta\nu$ would allow the possibility of conditioning our posterior estimates on $\Delta\nu$ in order to obtain more precise masses and hence ages.

Despite producing error estimates on properties we note that our method is not fully Bayesian since we do not marginalise over the population hyperparameters when fitting to new stars. Our model could be adapted to become fully hierarchical by having latent parameters on stellar properties which must be probabilistically consistent with the population at the same time as fitting the observations. The population hyperparameters would ensure consistency between stars but the model would be able to respond as new information entered the system. With this model configuration the only difference between training and testing data would be how

tightly constrained the uncertainties were: narrow for stars with known properties which would in turn guide the population hyperparameters, and wide for new stars which could be tighten by the process of fitting hierarchically, i.e. simultaneous training and testing stages. This model would then also include uncertainty on the map between bulk properties and periodogram parameters.

In its current form our model can be applied to *Kepler* and *K2* observations. However, there are a few important differences that must be considered when applying a model trained on *Kepler* data to TESS observations. Among other difficulties, the two telescopes have different band-passes where *Kepler* observed in $\sim 400 - 900$nm to target solar-type oscillators whilst TESS observes in the region $\sim 600 - 1100$nm targeting cooler stars. This means that some parts of the signal, such as surface granulation, will have a smaller amplitude in the redder band-pass. The systematic noise of TESS is also higher than that of *Kepler* which may limit the model's ability to fit to observations (see Yuxi et al. (2020) for further discussion). However, for application to TESS stars we could simply re-calibrate the model by either re-training using TESS data or by using a correction scaling factor based on TESS magnitude, e.g. by including the factor of 0.85 suggested by Campante et al. (2016), or a similar, and temperature dependent, bolometric correction as suggested by Lund (2019). The situation is similar if one attempts to make predictions on stars whose properties lie outside of the parameter space covered by the training data: the training set would have to be expanded to cover these regions.

We would like to note that although this method works well for evolved stars, it is likely to be more difficult to apply to main sequence populations. Red giant stars are more self-similar and there would be additional complexities to consider for younger stars. However, with the large quantities of data from the PLATO mission, which targets solar-like stars, applying our method may be possible after training the model on new data sets.

An alternative method to that discussed here, and one which is potentially more

easily scaleable to larger and more complex models, is to use a Bayesian Neural Network (BNN). This also follows machine learning principles by learning an arbitrary map between stellar properties and parameters describing a power spectrum model, and would be able to rapidly make inference on new observations. Neural networks produce highly flexible models which can capture abstract features in the data. Increasing the complexity from a simple global background fit to one which fits to individual modes and perhaps also takes in non-asteroseismic observations would be relatively trivial. After all, neural networks with several nodes and hidden layers already have a huger number of parameters to optimise so adding ~ 20 more is similarly manageable. However, a method such as this would have its own unique problems. The main difference between a BNN and a standard neural network is that instead of each weight and bias being a point value they are distributions. As such, computation is much more intensive during training. Hyperparameter tuning is challenging and BNNs notoriously suffer from weak identifiability between model parameters and have multimodality problems in the posterior. Overcoming these obstacles is a highly active area of research since the potential usages would be vast (see e.g. Depeweg et al. 2017; Pourzanjani et al. 2017; Mitros & Mac Namee 2019; Sen et al. 2020).

As shown in this work, we are able to obtain reasonable estimates for $\nu_{\rm max}$ and $\Delta\nu$ without bespoke priors for each object. This has usefulness beyond estimating physical properties via scaling relations. By combining this model as a pre-analysis step to a more rigorous fitting program such as `PBjam`, we could provide good, 'hands-free' priors over these quantities to replace the human input priors that are currently required. Having an estimate of the background level, $\nu_{\rm max}$ and $T_{\rm eff}$ is also a requirement for the FAMED pipeline (Corsaro et al., 2020) which could be provided by our methods.

Bibliography

Abolfathi B., et al., 2018, The Astrophysical Journal Supplement, 235, 42

Adelberger E. G., et al., 1998, Reviews of Modern Physics, 70, 1265

Adelberger E. G., et al., 2011, Reviews of Modern Physics, 83, 195

Aerts C., Christensen-Dalsgaard J., Kurtz D. W., 2010, Asteroseismology. Springer Science

Aigrain S., Favata F., Gilmore G., 2004, Astronomy and Astrophysics, 414, 1139

Aigrain S., et al., 2015, Monthly Notices of the Royal Astronomical Society, 450, 3211

Aigrain S., et al., 2016, The Effects of Stellar Activity on Detecting and Characterising Planets, doi:10.5281/zenodo.154565, `https://doi.org/10.5281/zenodo.154565`

Alvarado-Gómez J. D., et al., 2015, Astronomy and Astrophysics, 582, A38

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

Anguera Gubau M., Palle P. L., Perez Hernandez F., Regulo C., Roca Cortes T., 1992, Astronomy and Astrophysics, 255, 363

Antia H. M., Basu S., 1994, Astronomy and Astrophysics Supplement, 107, 421

Appourchaux T., 2003, Astronomy and Astrophysics, 412, 903

Appourchaux T., 2008, Astronomische Nachrichten, 329, 485

Appourchaux T., 2014, A crash course on data analysis in asteroseismology. Cambridge University Press, p. 123

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

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

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

Appourchaux T., et al., 2012, Astronomy and Astrophysics, 543, A54

Badnell N. R., Bautista M. A., Butler K., Delahaye F., Mendoza C., Palmeri P., Zeippen C. J., Seaton M. J., 2005, Monthly Notices of the Royal Astronomical Society, 360, 458

Baglin A., Auvergne M., Barge P., Deleuil M., Catala C., Michel E., Weiss W., COROT Team 2006, in The CoRoT Mission Pre-Launch Status - Stellar Seismology and Planet Finding. p. 33

Baliunas S. L., et al., 1995, The Astrophysical Journal, 438, 269

Ball W. H., Gizon L., 2014, Astronomy and Astrophysics, 568, A123

Barclay T., 2020, The TESS extended mission, `https://heasarc.gsfc.nasa.gov/docs/tess/the-tess-extended-mission.html`

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

Basri G., et al., 2010, The Astrophysical Journal Letters, 713, L155

Basri G., Walkowicz L. M., Reiners A., 2013, The Astrophysical Journal, 769, 37

Bastien F. A., Stassun K. G., Basri G., Pepper J., 2013, Nature, 500, 427

Basu S., 2016, Living Reviews in Solar Physics, 13, 2

Basu S., Chaplin W. J., 2017, Asteroseismic Data Analysis: Foundations and Techniques. Princeton University Press

Basu S., Kinnane A., 2018, The Astrophysical Journal, 869, 8

Basu S., Chaplin W. J., Elsworth Y., 2010, The Astrophysical Journal, 710, 1596

Batalha N. M., 2014, Proceedings of the National Academy of Science, 111, 12647

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

Bazot M., et al., 2018, Astronomy and Astrophysics, 619, L9

Bazot M., Benomar O., Christensen-Dalsgaard J., Gizon L., Hanasoge S., Nielsen M., Petit P., Sreenivasan K. R., 2019, Astronomy and Astrophysics, 623, A125

Beck P. G., et al., 2011, Science, 332, 205

Beck P. G., et al., 2020, in Neiner C., Weiss W. W., Baade D., Griffin R. E., Lovekin C. C., Moffat A. F. J., eds, Stars and their Variability Observed from Space. pp 75–79 (`arXiv:2001.04912`)

Bedding T. R., et al., 2011, Nature, 471, 608

Belkacem K., Goupil M. J., Dupret M. A., Samadi R., Baudin F., Noels A., Mosser B., 2011, Astronomy and Astrophysics, 530, A142

Bell K. J., 2020, Research Notes of the American Astronomical Society, 4, 19

Bellinger E. P., Angelou G. C., Hekker S., Basu S., Ball W. H., Guggenberger E., 2016, The Astrophysical Journal, 830, 31

Benomar O., et al., 2013, The Astrophysical Journal, 767, 158

Benomar O., et al., 2018, Science, 361, 1231

Berdyugina S. V., 2005, Living Reviews in Solar Physics, 2, 8

Blancato K., Ness M., Huber D., Lu Y., Angus R., 2020, arXiv e-prints, p. arXiv:2005.09682

Böhm-Vitense E., 2007, The Astrophysical Journal, 657, 486

Boro Saikia S., et al., 2016, Astronomy and Astrophysics, 594, A29

Boro Saikia S., et al., 2018a, Astronomy and Astrophysics, 616, A108

Boro Saikia S., et al., 2018b, Astronomy and Astrophysics, 620, L11

Borucki W. J., et al., 2010, Science, 327, 977

Bradshaw S. J., Hartigan P., 2014, The Astrophysical Journal, 795, 79

Brandenburg A., Saar S. H., Turpin C. R., 1998, The Astrophysical Journal Letters, 498, L51

Braun D. C., Duvall T. L. J., Labonte B. J., 1988, The Astrophysical Journal, 335, 1015

Brogaard K., et al., 2018, Monthly Notices of the Royal Astronomical Society, 476, 3729

Broomhall A. M., 2017, Solar Physics, 292, 67

Broomhall A. M., Nakariakov V. M., 2015, Solar Physics, 290, 3095

Broomhall A. M., Chaplin W. J., Elsworth Y., Fletcher S. T., New R., 2009, The Astrophysical Journal Letters, 700, L162

Broomhall A. M., Chaplin W. J., Elsworth Y., New R., 2011, Monthly Notices of the Royal Astronomical Society, 413, 2978

Broomhall A. M., Chaplin W. J., Elsworth Y., Simoniello R., 2012, Monthly Notices of the Royal Astronomical Society, 420, 1405

Broomhall A. M., Chatterjee P., Howe R., Norton A. A., Thompson M. J., 2014, Space Science Reviews, 186, 191

Broomhall A. M., Pugh C. E., Nakariakov V. M., 2015, Advances in Space Research, 56, 2706

Brown T. M., Gilliland R. L., Noyes R. W., Ramsey L. W., 1991, The Astrophysical Journal, 368, 599

Brun A. S., Browning M. K., 2017, Living Reviews in Solar Physics, 14, 4

Buchhave L. A., Latham D. W., 2015, The Astrophysical Journal, 808, 187

Bugnet L., García R. A., Davies G. R., Mathur S., Corsaro E., Hall O. J., Rendle B. M., 2018, Astronomy and Astrophysics, 620, A38

Bugnet L., García R. A., Mathur S., Davies G. R., Hall O. J., Lund M. N., Rendle B. M., 2019, doi:10.1051/0004-6361/201834780, 624, A79

Burtseva O., Hill F., Kholikov S., Chou D. Y., 2009, Solar Physics, 258, 1

Bushby P., Mason J., 2004, Astronomy and Geophysics, 45, 4.07

Campante T. L., et al., 2016, The Astrophysical Journal, 830, 138

Carrington R. C., 1858, Monthly Notices of the Royal Astronomical Society, 19, 1

Chaplin W. J., 2014, Sounding the solar cycle with helioseismology: Implications for asteroseismology. Cambridge University Press, p. 1

Chaplin W. J., Basu S., 2015, in Balogh A., Hudson H., Petrovay K., von Steiger R., eds, , The Solar Activity Cycle: Physical Causes and Consequences. Springer New York, pp 437–456, doi:10.1007/978-1-4939-2584-1_15, https://doi.org/10.1007/978-1-4939-2584-1_15

Chaplin W. J., Miglio A., 2013, Annual Review of Astronomy and Astrophysics, 51, 353

Chaplin W. J., et al., 1996, Solar Physics, 168, 1

Chaplin W. J., Elsworth Y., Isaak G. R., Lines R., McLeod C. P., Miller B. A., New R., 1998, Monthly Notices of the Royal Astronomical Society, 300, 1077

Chaplin W. J., Elsworth Y., Isaak G. R., Miller B. A., New R., 2000, Monthly Notices of the Royal Astronomical Society, 313, 32

Chaplin W. J., Appourchaux T., Elsworth Y., Isaak G. R., New R., 2001, Monthly Notices of the Royal Astronomical Society, 324, 910

Chaplin W. J., Elsworth Y., Isaak G. R., Miller B. A., New R., 2004, Monthly Notices of the Royal Astronomical Society, 352, 1102

Chaplin W. J., Elsworth Y., Miller B. A., New R., Verner G. A., 2005, The Astrophysical Journal Letters, 635, L105

Chaplin W. J., Elsworth Y., Houdek G., New R., 2007a, Monthly Notices of the Royal Astronomical Society, 377, 17

Chaplin W. J., Elsworth Y., Miller B. A., Verner G. A., New R., 2007b, The Astrophysical Journal, 659, 1749

Chaplin W. J., Elsworth Y., New R., Toutain T., 2008, Monthly Notices of the Royal Astronomical Society, 384, 1668

Chaplin W. J., et al., 2011a, Science, 332, 213

Chaplin W. J., et al., 2011b, The Astrophysical Journal, 732, 54

Chaplin W. J., et al., 2011c, The Astrophysical Journal Letters, 732, L5

Chaplin W. J., Elsworth Y., Davies G. R., Campante T. L., Handberg R., Miglio A., Basu S., 2014, Monthly Notices of the Royal Astronomical Society, 445, 946

Chaplin W. J., et al., 2019, Monthly Notices of the Royal Astronomical Society, 489, L86

Christensen-Dalsgaard J., 1988, in Christensen-Dalsgaard J., Frandsen S., eds, Advances in Helio- and Asteroseismology Vol. 123, Advances in Helio- and Asteroseismology. p. 295

Christensen-Dalsgaard J., 2002, Reviews of Modern Physics, 74, 1073

Christensen-Dalsgaard J., 2014, Asteroseismology lecture notes, 5 edn. Institut for Fysik og Astronomi, Aarhus Universitet, Teoretisk Astrofysik Center, Danmarks Grundforskningsfond

Christensen-Dalsgaard J., et al., 2010a, The Astrophysical Journal Letters, 713, L164

Christensen-Dalsgaard J., et al., 2010b, The Astrophysical Journal, 713, L164

Clette F., Svalgaard L., Vaquero J. M., Cliver E. W., 2014, Space Science Reviews, 186, 35

Corsaro E., De Ridder J., 2014, Astronomy and Astrophysics, 571, A71

Corsaro E., et al., 2017, Astronomy and Astrophysics, 605, A3

Corsaro E., McKeever J. M., Kuszlewicz J. S., 2020, Astronomy and Astrophysics, 640, A130

Cowling T. G., 1941, Monthly Notices of the Royal Astronomical Society, 101, 367

Cranmer S. R., Bastien F. A., Stassun K. G., Saar S. H., 2014, The Astrophysical Journal, 781, 124

Creevey O. L., et al., 2017, Astronomy and Astrophysics, 601, A67

Davenport J., 2015, PhD thesis, University of Washington

Davies G. R., Miglio A., 2016, Astronomische Nachrichten, 337, 774

Davies G. R., et al., 2015, Monthly Notices of the Royal Astronomical Society, 446, 2959

Davies G. R., et al., 2016, Monthly Notices of the Royal Astronomical Society, 456, 2183

Demarque P., Guenther D. B., Li L. H., Mazumdar A., Straka C. W., 2008, Astrophysics and Space Science, 316, 31

Deng L.-C., et al., 2012, Research in Astronomy and Astrophysics, 12, 735

Depeweg S., Hernández-Lobato J. M., Doshi-Velez F., Udluft S., 2017, arXiv e-prints, p. arXiv:1706.08495

Dikpati M., Gilman P. A., 2009, Space Science Reviews, 144, 67

Donati J. F., Brown S. F., 1997, Astronomy and Astrophysics, 326, 1135

Duane S., Kennedy A. D., Pendleton B. J., Roweth D., 1987, Physics Letters B, 195, 216

Duncan D. K., et al., 1991, The Astrophysical Journal Supplement, 76, 383

ESA 2017, Technical report, PLATO Definition Study Report. ESA

Eaton J. A., Hall D. S., 1979, Astrophysical Journal, 227, 907

Elsworth Y., Howe R., Isaak G. R., McLeod C. P., New R., 1990, Nature, 345, 322

Elsworth Y., Hekker S., Basu S., Davies G. R., 2017, Monthly Notices of the Royal Astronomical Society, 466, 3344

Fabbian D., et al., 2017, Astronomische Nachrichten, 338, 753

Fares R., 2014, in Petit P., Jardine M., Spruit H. C., eds, IAU Symposium Vol. 302, Magnetic Fields throughout Stellar Evolution. pp 180–189 (arXiv:1310.5562), doi:10.1017/S1743921314002026

Fares R., Moutou C., Donati J. F., Catala C., Shkolnik E. L., Jardine M. M., Cameron A. C., Deleuil M., 2013, Monthly Notices of the Royal Astronomical Society, 435, 1451

Ferguson J. W., Alexander D. R., Allard F., Barman T., Bodnarik J. G., Hauschildt P. H., Heffner-Wong A., Tamanai A., 2005, The Astrophysical Journal, 623, 585

Field D. A., 1991, Advances in Engineering Software and Workstations, 13, 263

Fletcher S. T., Chaplin W. J., Elsworth Y., New R., 2009, The Astrophysical Journal, 694, 144

Fletcher S. T., Broomhall A.-M., Salabert D., Basu S., Chaplin W. J., Elsworth Y., Garcia R. A., New R., 2010, The Astrophysical Journal Letters, 718, L19

Foreman-Mackey D., 2016, Journal of Open Source Software, 1, 24

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, Publications of the ASP, 125, 306

Formicola A., et al., 2004, Physics Letters B, 591, 61

Fossat E., et al., 2017, Astronomy and Astrophysics, 604, A40

Fröhlich C., et al., 1995, Solar Physics, 162, 101

Frohlich C., et al., 1997, Solar Physics, 170, 1

Fuller J., Cantiello M., Stello D., Garcia R. A., Bildsten L., 2015, Science, 350, 423

Gabriel A. H., et al., 1995, Solar Physics, 162, 61

Gai N., Basu S., Chaplin W. J., Elsworth Y., 2011, The Astrophysical Journal, 730, 63

Gaia Collaboration 2016, Astronomy and Astrophysics, 595, A2

Gaia Collaboration 2018, Astronomy and Astrophysics, 616, A1

García R. A., 2015, in EAS Publications Series. pp 193–259 (arXiv:1510.02651), doi:10.1051/eas/1573004

García R. A., Ballot J., 2019, Living Reviews in Solar Physics, 16, 4

García Saravia Ortiz de Montellano A., Hekker S., Themeßl N., 2018, Monthly Notices of the Royal Astronomical Society, 476, 1470

García R. A., Turck-Chièze S., Jiménez-Reyes S. J., Ballot J., Pallé P. L., Eff-Darwich A., Mathur S., Provost J., 2007, Science, 316, 1591

García R. A., et al., 2008, Astronomische Nachrichten, 329, 476

García R. A., Mathur S., Salabert D., Ballot J., Régulo C., Metcalfe T. S., Baglin A., 2010, Science, 329, 1032

García R. A., et al., 2014, Astronomy and Astrophysics, 572, A34

Garraffo C., Drake J. J., Cohen O., 2016, Astronomy and Astrophysics, 595, A110

Gaulme P., et al., 2016, The Astrophysical Journal, 832, 121

Gelman A., Rubin D. B., 1992, Statistical Science, 7, 457

Gelman A., Carlin J. B., Stern H. S., Dunson D. B., Vehtari A., Rubin D. B., 2013, Bayesian Data Analysis, 3 edn

Giles H. A. C., Collier Cameron A., Haywood R. D., 2017, Monthly Notices of the Royal Astronomical Society, 472, 1618

Gilliland R. L., et al., 2011, The Astrophysical Journal Supplement, 197, 6

Gizon L., Solanki S. K., 2003, The Astrophysical Journal, 589, 1009

Gizon L., Solanki S. K., 2004, Solar Physics, 220, 169

Goldreich P., Keeley D. A., 1977, The Astrophysical Journal, 211, 934

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, Vol.˜5, No.˜1, p.˜65-80, 2010, 5, 65

Gough D. O., Thompson M. J., 1990, Monthly Notices of the Royal Astronomical Society, 242, 25

Grenander U., 1959, in Probability and Statistics: The Harald CramГқr Volume.

Grevesse N., Noels A., 1993, in Prantzos N., Vangioni-Flam E., Casse M., eds, Origin and Evolution of the Elements. pp 15–25

Grevesse N., Sauval A. J., 1998, Space Science Reviews, 85, 161

Gruberbauer M., Guenther D. B., MacLeod K., Kallinger T., 2013, Monthly Notices of the Royal Astronomical Society, 435, 242

Grundahl F., ChristensenвЙПЬDalsgaard J., Arentoft T., Frandsen S., Kjeldsen H., Jørgensen U. G., Kjaergaard P., 2008, Communications in Asteroseismology, 157, 273

Hale S. J., Howe R., Chaplin W. J., Davies G. R., Elsworth Y. P., 2016, Solar Physics, 291, 1

Hall O. J., et al., 2021, Nature Astronomy,

Handberg R., Campante T. L., 2011, Astronomy and Astrophysics, 527, A56

Handberg R., Lund M. N., 2014, Monthly Notices of the Royal Astronomical Society, 445, 2698

Handler G., 2009, Monthly Notices of the Royal Astronomical Society, 398, 1339

Harvey J., 1985, in Rolfe E., Battrick B., eds, ESA Special Publication Vol. 235, Future Missions in Solar, Heliospheric & Space Plasma Physics. p. 199

Harvey K. L., White O. R., 1999, The Astrophysical Journal, 515, 812

Harvey J. W., et al., 1996, Science, 272, 1284

Hathaway D. H., 2015, Living Reviews in Solar Physics, 12, 4

Haywood R. D., et al., 2014, Monthly Notices of the Royal Astronomical Society, 443, 2517

Hekker S., 2020, Frontiers in Astronomy and Space Sciences, 7, 3

Hekker S., et al., 2011a, Monthly Notices of the Royal Astronomical Society, 414, 2594

Hekker S., et al., 2011b, Astronomy and Astrophysics, 525, A131

Hekker S., Elsworth Y., Basu S., Bellinger E., 2017, in European Physical Journal Web of Conferences. p. 04006 (arXiv:1705.05879), doi:10.1051/epjconf/201716004006

Henry T. J., Soderblom D. R., Donahue R. A., Baliunas S. L., 1996, Astronomical Journal, 111, 439

Hoffman M. D., Gelman A., 2011, arXiv e-prints, p. arXiv:1111.4246

Hon M., Stello D., Yu J., 2018, Monthly Notices of the Royal Astronomical Society, 476, 3233

Hon M., Bellinger E. P., Hekker S., Stello D., Kuszlewicz J. S., 2020, Monthly Notices of the Royal Astronomical Society, 499, 2445

Houdek G., Balmforth N. J., Christensen-Dalsgaard J., Gough D. O., 1999, Astronomy and Astrophysics, 351, 582

Howard R., Labonte B. J., 1981, Solar Physics, 74, 131

Howe R., Komm R. W., Hill F., 2002, Astrophysical Journal, 580, 1172

Howe R., Komm R. W., Hill F., Haber D. A., Hindman B. W., 2004, The Astrophysical Journal, 608, 562

Howe R., Haber D. A., Hindman B. W., Komm R., Hill F., Gonzalez Hernandez I., 2008, in Howe R., Komm R. W., Balasubramaniam K. S., Petrie G. J. D., eds, Astronomical Society of the Pacific Conference Series Vol. 383, Subsurface and Atmospheric Influences on Solar Activity. p. 305

Howe R., Davies G. R., Chaplin W. J., Elsworth Y. P., Hale S. J., 2015, Monthly Notices of the Royal Astronomical Society, 454, 4120

Howe R., Basu S., Davies G. R., Ball W. H., Chaplin W. J., Elsworth Y., Komm R., 2017, Monthly Notices of the Royal Astronomical Society, 464, 4777

Howe R., Chaplin W. J., Davies G. R., Elsworth Y., Basu S., Broomhall A. M., 2018, Monthly Notices of the Royal Astronomical Society, 480, L79

Howe R., et al., 2020, Monthly Notices of the Royal Astronomical Society, 493, L49

Howell S. B., et al., 2014, Publications of the Astronomical Society of the Pacific, 126, 398

Huber D., 2018, in Campante T. L., Santos N. C., Monteiro M. J. P. F. G., eds, Asteroseismology and Exoplanets: Listening to the Stars and Searching for New Worlds Vol. 49, Asteroseismology and Exoplanets: Listening to the Stars and Searching for New Worlds. p. 119 (arXiv:1711.01281), doi:10.1007/978-3-319-59315-9_6

Huber D., et al., 2011a, The Astrophysical Journal, 731, 94

Huber D., et al., 2011b, The Astrophysical Journal, 743, 143

Huber D., et al., 2013, The Astrophysical Journal, 767, 127

Huber D., et al., 2016, Astrophysical Journal Supplement, 224, 2

Huber D., et al., 2017, The Astrophysical Journal, 844, 102

Iglesias C. A., Rogers F. J., 1996, The Astrophysical Journal, 464, 943

Ioannidis P., Huber K. F., Schmitt J. H. M. M., 2016, Astronomy and Astrophysics, 585, A72

Irwin A. W., 2012, FreeEOS: Equation of State for stellar interiors calculations (ascl:1211.002)

Isaacson H., Fischer D., 2010, The Astrophysical Journal, 725, 875

Jain K., Tripathy S. C., Bhatnagar A., Kumar B., 2000, Solar Physics, 192, 487

Jain K., Tripathy S., Hill F., Salabert D., García R. A., Broomhall A.-M., 2018, IAU Symposium, 340, 27

Jeffers S. V., et al., 2018, Monthly Notices of the Royal Astronomical Society, 479, 5266

Jenkins J. S., et al., 2006, Monthly Notices of the Royal Astronomical Society, 372, 163

Jimenez-Reyes S. J., Regulo C., Palle P. L., Roca Cortes T., 1998, Astronomy and Astrophysics, 329, 1119

Jiménez-Reyes S. J., García R. A., Jiménez A., Chaplin W. J., 2003, The Astrophysical Journal, 595, 446

Jiménez-Reyes S. J., García R. A., Chaplin W. J., Korzennik S. G., 2004, Astrophysical Journal, 610, L65

Jiménez-Reyes S. J., et al., 2008, Monthly Notices of the Royal Astronomical Society, 389, 1780

Jiménez A., Roca Cortés T., Jiménez-Reyes S. J., 2002, Solar Physics, 209, 247

Johnson J. A., et al., 2017, Astronomical Journal, 154, 108

Kallinger T., 2019, arXiv e-prints, p. arXiv:1906.09428

Kallinger T., et al., 2010, Astronomy and Astrophysics, 522, A1

Kallinger T., et al., 2012, Astronomy and Astrophysics, 541, A51

Kallinger T., et al., 2014, Astronomy and Astrophysics, 570, A41

Kallinger T., Beck P. G., Stello D., Garcia R. A., 2018, Astronomy and Astrophysics, 616, A104

Karoff C., et al., 2013, Monthly Notices of the Royal Astronomical Society, 433, 3227

Karoff C., et al., 2016, Nature Communications, 7, 11058

Karoff C., et al., 2018, The Astrophysical Journal, 852, 46

Karoff C., Metcalfe T. S., Montet B. T., Jannsen N. E., Santos A. R. G., Nielsen M. B., Chaplin W. J., 2019, Monthly Notices of the Royal Astronomical Society, 485, 5096

Kiefer R., Broomhall A.-M., 2020, Monthly Notices of the Royal Astronomical Society, 496, 4593

Kiefer R., Schad A., Davies G., Roth M., 2017, Astronomy and Astrophysics, 598, A77

Kiefer R., Broomhall A.-M., Ball W. H., 2019, Frontiers in Astronomy and Space Sciences, 6, 52

Kjeldsen H., Bedding T. R., 1995, Astronomy and Astrophysics, 293, 87

Kjeldsen H., Bedding T. R., Christensen-Dalsgaard J., 2008, The Astrophysical Journal Letters, 683, L175

Kochukhov O., Lundin A., Romanyuk I., Kudryavtsev D., 2011, Astrophysical Journal, 726, 24

Kochukhov O., Lüftinger T., Neiner C., Alecian E., MiMeS Collaboration 2014, Astronomy and Astrophysics, 565, A83

Kolmogorov A., 1933, Giornale dellвПЉНInstituto Italiano degli Attuari, 4, 83

Komm R. W., Howe R., Hill F., 2000a, Astrophysical Journal, 531, 1094

Komm R. W., Howe R., Hill F., 2000b, The Astrophysical Journal, 543, 472

Krishna Swamy K. S., 1966, The Astrophysical Journal, 145, 174

Kuszlewicz J. S., Hekker S., Bell K. J., 2020, Monthly Notices of the Royal Astronomical Society, 497, 4843

Lammer H., et al., 2012, Earth, Planets, and Space, 64, 179

Lanza A. F., Rodonò M., Pagano I., Barge P., Llebaria A., 2003, Astronomy and Astrophysics, 403, 1135

Lanza A. F., Bonomo A. S., Rodonò M., 2007, Astronomy and Astrophysics, 464, 741

Lebreton Y., Goupil M. J., 2014, Astronomy and Astrophysics, 569, A21

Lebreton Y., Montalbán J., 2009, in Mamajek E. E., Soderblom D. R., Wyse R. F. G., eds, The Ages of Stars Vol. 258, The Ages of Stars. pp 419–430 (arXiv:0811.2908), doi:10.1017/S1743921309032074

Lebreton Y., Goupil M. J., Montalbán J., 2014, in EAS Publications Series. pp 177–223 (arXiv:1410.5337), doi:10.1051/eas/1465005

Lehmann L. T., Hussain G. A. J., Jardine M. M., Mackay D. H., Vidotto A. A., 2019, Monthly Notices of the Royal Astronomical Society, 483, 5246

Lehmann L. T., Hussain G. A. J., Vidotto A. A., Jardine M. M., Mackay D. H., 2020, Monthly Notices of the Royal Astronomical Society, 500, 1243

Lewandowski D., Kurowicka D., Joe H., 2009, Journal of Multivariate Analysis, 100, 1989

Li T., Bedding T. R., Huber D., Ball W. H., Stello D., Murphy S. J., Bland-Hawthorn J., 2018, Monthly Notices of the Royal Astronomical Society, 475, 981

Libbrecht K. G., Woodard M. F., 1990, Nature, 345, 779

Lieu M., Farr W. M., Betancourt M., Smith G. P., Sereno M., McCarthy I. G., 2017, Monthly Notices of the Royal Astronomical Society, 468, 4872

Lightkurve Collaboration et al., 2018, Lightkurve: Kepler and TESS time series analysis in Python, Astrophysics Source Code Library (ascl:1812.013)

Linsky J. L., 1999, in Butler C. J., Doyle J. G., eds, Astronomical Society of the Pacific Conference Series Vol. 158, Solar and Stellar Activity: Similarities and Differences. p. 3

Lopes I., Turck-Chieze S., 1994, Astronomy and Astrophysics, 290, 845

Lund M. N., 2019, Monthly Notices of the Royal Astronomical Society, 489, 1072

Lund M. N., et al., 2017, Astrophysical Journal, 835, 172

Mackay D. J. C., 2003, Information Theory, Inference and Learning Algorithms

Maehara H., Shibayama T., Notsu Y., Notsu S., Honda S., Nogami D., Shibata K., 2015, Earth, Planets, and Space, 67, 59

Majewski S. R., et al., 2017, Astronomical Journal, 154, 94

Makarov V. V., Beichman C. A., Catanzarite J. H., Fischer D. A., Lebreton J., Malbet F., Shao M., 2009, The Astrophysical Journal Letters, 707, L73

Mandal S., Karak B. B., Banerjee D., 2017, The Astrophysical Journal, 851, 70

Mathur S., 2014, in Ballet J., Martins F., Bournaud F., Monier R., Reylé C., eds, SF2A-2014: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics. pp 257–262

Mathur S., et al., 2010, Astronomy and Astrophysics, 511, A46

Mathur S., García R. A., Salabert D., Ballot J., Régulo C., Metcalfe T. S., Baglin A., 2011a, in GONG-SoHO 24: A New Era of Seismology of the Sun and Solar-Like Stars. p. 012045 (arXiv:1011.4102), doi:10.1088/1742-6596/271/1/012045

Mathur S., et al., 2011b, The Astrophysical Journal, 741, 119

Mathur S., et al., 2012, The Astrophysical Journal, 749, 152

Mathur S., García R. A., Morgenthaler A., Salabert D., Petit P., Ballot J., Régulo C., Catala C., 2013, Astronomy and Astrophysics, 550, A32

Mathur S., Salabert D., García R. A., Ceillier T., 2014a, Journal of Space Weather and Space Climate, 4, A15

Mathur S., et al., 2014b, Astronomy and Astrophysics, 562, A124

Mathur S., García R. A., Huber D., Regulo C., Stello D., Beck P. G., Houmani K., Salabert D., 2016, The Astrophysical Journal, 827, 50

Mathur S., García R. A., Bugnet L., Santos Â. R. G., Santiago N., Beck P. G., 2019, Frontiers in Astronomy and Space Sciences, 6, 46

Maunder E. W., 1904, Monthly Notices of the Royal Astronomical Society, 64, 747

McQuillan A., Mazeh T., Aigrain S., 2013, The Astrophysical Journal Letters, 775, L11

McQuillan A., Mazeh T., Aigrain S., 2014, The Astrophysical Journal Supplement, 211, 24

Meibom S., Mathieu R. D., Stassun K. G., 2009, The Astrophysical Journal, 695, 679

Meibom S., Barnes S. A., Platais I., Gilliland R. L., Latham D. W., Mathieu R. D., 2015, Nature, 517, 589

Metcalfe T. S., van Saders J., 2017, Solar Physics, 292, 126

Metcalfe T. S., Dziembowski W. A., Judge P. G., Snow M., 2007, Monthly Notices of the Royal Astronomical Society, 379, L16

Metcalfe T. S., Basu S., Henry T. J., Soderblom D. R., Judge P. G., Knölker M., Mathur S., Rempel M., 2010, The Astrophysical Journal Letters, 723, L213

Metcalfe T. S., et al., 2013, The Astrophysical Journal Letters, 763, L26

Metcalfe T. S., et al., 2014, The Astrophysical Journal Supplement, 214, 27

Metcalfe T. S., Egeland R., van Saders J., 2016, The Astrophysical Journal Letters, 826, L2

Micela G., 2017, in Deeg H. J., Belmonte J. A., eds, , Handbook of Exoplanets. Springer International Publishing, pp 1–14, doi:10.1007/978-3-319-30648-3_19-1, https://doi.org/10.1007/978-3-319-30648-3_19-1

Michel E., Samadi R., Baudin F., Barban C., Appourchaux T., Auvergne M., 2009, Astronomy and Astrophysics, 495, 979

Miglio A., et al., 2013, Monthly Notices of the Royal Astronomical Society, 429, 423

Miglio A., et al., 2016, Monthly Notices of the Royal Astronomical Society, 461, 760

Miglio A., et al., 2017, Astronomische Nachrichten, 338, 644

Miglio A., et al., 2020, arXiv e-prints, p. arXiv:2004.14806

Mitros J., Mac Namee B., 2019, arXiv e-prints, p. arXiv:1912.01530

Montet B. T., Tovar G., Foreman-Mackey D., 2017, The Astrophysical Journal, 851, 116

Moreno-Insertis F., Solanki S. K., 2000, Monthly Notices of the Royal Astronomical Society, 313, 411

Morris B. M., Hebb L., Davenport J. R. A., Rohn G., Hawley S. L., 2017, The Astrophysical Journal, 846, 99

Morris B. M., Agol E., Hebb L., Hawley S. L., 2018, Astronomical Journal, 156, 91

Mosser B., Appourchaux T., 2009, Astronomy and Astrophysics, 508, 877

Mosser B., et al., 2011, Astronomy and Astrophysics, 532, A86

Mosser B., et al., 2012a, Astronomy and Astrophysics, 537, A30

Mosser B., et al., 2012b, Astronomy and Astrophysics, 540, A143

Mosser B., et al., 2013, Astronomy and Astrophysics, 550, A126

Mosser B., et al., 2014, Astronomy and Astrophysics, 572, L5

Mosser B., Vrard M., Belkacem K., Deheuvels S., Goupil M. J., 2015, Astronomy and Astrophysics, 584, A50

Murphy S. J., 2015, Monthly Notices of the Royal Astronomical Society, 453, 2569

Murphy S. J., Shibahashi H., Kurtz D. W., 2013, Monthly Notices of the Royal Astronomical Society, 430, 2986

Ness M., Hogg D. W., Rix H. W., Ho A. Y. Q., Zasowski G., 2015, The Astrophysical Journal, 808, 16

Ness M. K., Silva Aguirre V., Lund M. N., Cantiello M., Foreman-Mackey D., Hogg D. W., Angus R., 2018, The Astrophysical Journal, 866, 15

Netto Y., Valio A., 2020, Astronomy and Astrophysics, 635, A78

Newton E. R., Irwin J., Charbonneau D., Berta-Thompson Z. K., Dittmann J. A., West A. A., 2016, The Astrophysical Journal, 821, 93

Nielsen M. B., Gizon L., Schunker H., Karoff C., 2013, Astronomy and Astrophysics, 557, L10

Nielsen M. B., Gizon L., Cameron R. H., Miesch M., 2019, Astronomy and Astrophysics, 622, A85

Nielsen M. B., et al., 2021, Astronomical Journal, 161, 62

Noyes R. W., Hartmann L. W., Baliunas S. L., Duncan D. K., Vaughan A. H., 1984, The Astrophysical Journal, 279, 763

Ochsenbein F., Bauer P., Marcout J., 2000, Astronomy and Astrophysics Supplement, 143, 23

Oláh K., et al., 2009, Astronomy and Astrophysics, 501, 703

Oshagh M., Santos N. C., Boisse I., Boué G., Montalto M., Dumusque X., Haghighipour N., 2013, Astronomy and Astrophysics, 556, A19

Otí Floranes H., Christensen-Dalsgaard J., Thompson M. J., 2005, Monthly Notices of the Royal Astronomical Society, 356, 671

Pallé P. L., Regulo C., Roca Cortes T., 1989, Astronomy and Astrophysics, 224, 253

Pallé P. L., Régulo C., Roca Cortés T., 1990, The Spectrum of Solar p-Modes and the Solar Activity Cycle. Springer, p. 129, doi:10.1007/3-540-53091-6_73

Pande D., Bedding T. R., Huber D., Kjeldsen H., 2018, Monthly Notices of the Royal Astronomical Society, 480, 467

Pérez Hernández F., García R. A., Corsaro E., Triana S. A., De Ridder J., 2016, Astronomy and Astrophysics, 591, A99

Pérez Hernández F., García R. A., Mathur S., Santos A. R. G., Régulo C., 2019, Frontiers in Astronomy and Space Sciences, 6, 41

Pevtsov A. A., Virtanen I., Mursula K., Tlatov A., Bertello L., 2016, Astronomy and Astrophysics, 585, A40

Pinsonneault M. H., et al., 2014, The Astrophysical Journal Supplement, 215, 19

Pinsonneault M. H., et al., 2018, The Astrophysical Journal Supplement, 239, 32

Piskunov N., Kochukhov O., 2002, Astronomy and Astrophysics, 381, 736

Pont F., Sing D. K., Gibson N. P., Aigrain S., Henry G., Husnoo N., 2013, Monthly Notices of the Royal Astronomical Society, 432, 2917

Pourzanjani A., Jiang R., Petzold L., 2017, Improving the Identifiability of Neural Networks for Bayesian Inference

Quirion P.-O., Christensen-Dalsgaard J., Arentoft T., 2010, The Astrophysical Journal, 725, 2176

Radick R. R., Lockwood G. W., Skiff B. A., Baliunas S. L., 1998, The Astrophysical Journal Supplement, 118, 239

Rauer H., et al., 2014, Experimental Astronomy, 38, 249

Reese D. R., 2016, AIMS: Asteroseismic Inference on a Massive Scale (ascl:1611.014)

Reese D., Lignières F., Rieutord M., 2006, Astronomy and Astrophysics, 455, 621

Reese D. R., et al., 2016, Astronomy and Astrophysics, 592, A14

Régulo C., García R. A., Ballot J., 2016, Astronomy and Astrophysics, 589, A103

Reinhold T., Reiners A., Basri G., 2013, Astronomy and Astrophysics, 560, A4

Reinhold T., Cameron R. H., Gizon L., 2017, Astronomy and Astrophysics, 603, A52

Reinhold T., Shapiro A. I., Solanki S. K., Montet B. T., Krivova N. A., Cameron R. H., Amazo-Gómez E. M., 2020, Science, 368, 518

Rendle B. M., et al., 2019a, Monthly Notices of the Royal Astronomical Society, 484, 771

Rendle B. M., et al., 2019b, Monthly Notices of the Royal Astronomical Society, 490, 4465

Rhodes E. J., et al., 2011, in Prasad Choudhary D., Strassmeier K. G., eds, Proceedings of the International Astronomical Union Vol. 273, The Physics of Sun and Star Spots. pp 389–393, doi:10.1017/S1743921311015614

Ricker G., 2019, in AAS/Division for Extreme Solar Systems Abstracts. p. 100.01

Ricker G. R., et al., 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003

Rodrigues T. S., et al., 2017, Monthly Notices of the Royal Astronomical Society, 467, 1433

Roettenbacher R. M., et al., 2016, Nature, 533, 217

Rogers F. J., Nayfonov A., 2002, The Astrophysical Journal, 576, 1064

Rosén L., Kochukhov O., 2012, How reliable is Zeeman Doppler imaging without simultaneous temperature reconstruction? (`arXiv:1210.0789`), doi:10.1051/0004-6361/201219972

Roxburgh I. W., 2015, Astronomy and Astrophysics, 574, A45

Roxburgh I. W., Vorontsov S. V., 2003, Astronomy and Astrophysics, 411, 215

Rusomarov N., Kochukhov O., Ryabchikova T., Piskunov N., 2015, Astronomy and Astrophysics, 573, A123

Salabert D., Chaplin W. J., Elsworth Y., New R., Verner G. A., 2007, Astronomy and Astrophysics, 463, 1181

Salabert D., García R. A., Pallé P. L., Jiménez-Reyes S. J., 2009, Astronomy and Astrophysics, 504, L1

Salabert D., Régulo C., Ballot J., García R. A., Mathur S., 2011, Astronomy and Astrophysics, 530, A127

Salabert D., García R. A., Turck-Chièze S., 2015, Astronomy and Astrophysics, 578, A137

Salabert D., et al., 2016a, in 19th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun (CS19). Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun. p. 30 (arXiv:1610.00990), doi:10.5281/zenodo.159255

Salabert D., et al., 2016b, Astronomy and Astrophysics, 589, A118

Salabert D., García R. A., Jiménez A., Bertello L., Corsaro E., Pallé P. L., 2017, Astronomy and Astrophysics, 608, A87

Salabert D., Régulo C., Pérez Hernández F., García R. A., 2018, Astronomy and Astrophysics, 611, A84

Sanchez S., Fournier A., Aubert J., 2014, The Astrophysical Journal, 781, 8

Santos A. R. G., Cunha M. S., Avelino P. P., Chaplin W. J., Campante T. L., 2016, Monthly Notices of the Royal Astronomical Society, 461, 224

Santos A. R. G., Cunha M. S., Avelino P. P., García R. A., Mathur S., 2017, Astronomy and Astrophysics, 599, A1

Santos A. R. G., et al., 2018, The Astrophysical Journal Supplement, 237, 17

Santos A. R. G., García R. A., Mathur S., Bugnet L., van Saders J. L., Metcalfe T. S., Simonian G. V. A., Pinsonneault M. H., 2019a, Astrophysical Journal Supplement, 244, 21

Santos A. R. G., et al., 2019b, The Astrophysical Journal, 883, 65

Schaefer B. E., King J. R., Deliyannis C. P., 2000, The Astrophysical Journal, 529, 1026

Scherrer P. H., et al., 1995, Solar Physics, 162, 129

Scherrer P. H., et al., 2012, Solar Physics, 275, 207

Schunker H., Schou J., Gaulme P., Gizon L., 2018, Solar Physics, 293, 95

Scuflaire R., Théado S., Montalbán J., Miglio A., Bourge P. O., Godart M., Thoul A., Noels A., 2008a, Astrophysics and Space Science, 316, 83

Scuflaire R., Montalbán J., Théado S., Bourge P. O., Miglio A., Godart M., Thoul A., Noels A., 2008b, Astrophysics and Space Science, 316, 149

Semel M., 1989, Astronomy and Astrophysics, 225, 456

Sen D., Papamarkou T., Dunson D., 2020, arXiv e-prints, p. arXiv:2008.08044

Severino G., Magrì M., Oliviero M., Straus T., Jefferies S. M., 2001, The Astrophysical Journal, 561, 444

Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, The Astrophysical Journal, 822, 15

Sheeley N. R. J., 2010, in Cranmer S. R., Hoeksema J. T., Kohl J. L., eds, Astronomical Society of the Pacific Conference Series Vol. 428, SOHO-23: Understanding a Peculiar Solar Minimum. p. 3 (arXiv:1005.3834)

Shibayama T., et al., 2013, The Astrophysical Journal Supplement, 209, 5

Silva A. V. R., 2003, Astrophysical Journal, 585, L147

Silva Aguirre V., et al., 2011, The Astrophysical Journal Letters, 740, L2

Silva Aguirre V., et al., 2013, The Astrophysical Journal, 769, 141

Silva Aguirre V., et al., 2015, Monthly Notices of the Royal Astronomical Society, 452, 2127

Silva Aguirre V., et al., 2017, The Astrophysical Journal, 835, 173

Simoniello R., Finsterle W., Salabert D., García R. A., Turck-Chièze S., Jiménez A., Roth M., 2012, Astronomy and Astrophysics, 539, A135

Skrutskie M. F., et al., 2006, Astronomical Journal, 131, 1163

Smirnov N. V., 1939, Bull Moscow University, 2, 3616

Solanki S. K., 2003, Astronomy and Astrophysics Reviews, 11, 153

Solanki S. K., Wenzler T., Schmitt D., 2008, Astronomy and Astrophysics, 483, 623

Sonoi T., Samadi R., Belkacem K., Ludwig H. G., Caffau E., Mosser B., 2015, Astronomy and Astrophysics, 583, A112

Stahn T., Gizon L., 2008, Solar Physics, 251, 31

Steigman G., 2010, Journal of Cosmology and Astroparticle Physics, 4, 029

Stello D., Chaplin W. J., Basu S., Elsworth Y., Bedding T. R., 2009a, Monthly Notices of the Royal Astronomical Society, 400, L80

Stello D., et al., 2009b, The Astrophysical Journal, 700, 1589

Stello D., et al., 2013, The Astrophysical Journal Letters, 765, L41

Stello D., et al., 2015, The Astrophysical Journal Letters, 809, L3

Tassoul M., 1980, The Astrophysical Journal Supplement, 43, 469

Tassoul M., 1990, The Astrophysical Journal, 358, 313

The Stan Development Team 2018, Stan Modeling Language Users Guide and Reference Manual, http://mc-stan.org

Thomas A. E. L., et al., 2019, Monthly Notices of the Royal Astronomical Society, 485, 3857

Thomas A. E. L., Chaplin W. J., Basu S., Rendle B., Davies G., Miglio A., 2021, Monthly Notices of the Royal Astronomical Society, 502, 5808

Thoul A. A., Bahcall J. N., Loeb A., 1994, The Astrophysical Journal, 421, 828

Toutain T., Wehrli C., 1997, in Schmieder B., del Toro Iniesta J. C., Vazquez M., eds, Astronomical Society of the Pacific Conference Series Vol. 118, 1st Advances in Solar Physics Euroconference. Advances in Physics of Sunspots. p. 254

Tregloan-Reed J., Unda-Sanzana E., 2019, Astronomy and Astrophysics, 630, A114

Tripathy S. C., Hill F., Jain K., Leibacher J. W., 2006, in Fletcher K., Thompson M., eds, ESA Special Publication Vol. 624, Proceedings of SOHO 18/GONG 2006/HELAS I, Beyond the spherical Sun. p. 93

Tripathy S. C., Hill F., Jain K., Leibacher J. W., 2007, Solar Physics, 243, 105

Turck-Chièze S., et al., 2004, The Astrophysical Journal, 604, 455

Ulrich R. K., 1986, The Astrophysical Journal Letters, 306, L37

Unknown authors 800BC, I Ching

Usoskin I. G., 2013, Living Reviews in Solar Physics, 10, 1

Usoskin I. G., 2017, Living Reviews in Solar Physics, 14, 3

Van Eylen V., et al., 2014, The Astrophysical Journal, 782, 14

Van Eylen V., et al., 2015, in European Physical Journal Web of Conferences. p. 02005 (arXiv:1412.4848), doi:10.1051/epjconf/201510102005

Van Eylen V., Agentoft C., Lundkvist M. S., Kjeldsen H., Owen J. E., Fulton B. J., Petigura E., Snellen I., 2018, Monthly Notices of the Royal Astronomical Society, 479, 4786

Vandakurov Y. V., 1967, Astronomicheskii Zhurnal, 44, 786

Vanderburg A., Plavchan P., Johnson J. A., Ciardi D. R., Swift J., Kane S. R., 2016, Monthly Notices of the Royal Astronomical Society, 459, 3565

Veitch J., Del Pozzo W., Cody ed1d1a8d 2017, Johnveitch/Cpnest: Pypi Release, doi:10.5281/zenodo.825456

Viani L. S., Basu S., Chaplin W. J., Davies G. R., Elsworth Y., 2017, The Astrophysical Journal, 843, 11

Vorontsov S. V., 1991, Soviet Astronomy, 35, 400

Vrard M., Mosser B., Samadi R., 2016, Astronomy and Astrophysics, 588, A87

Walker G., et al., 2003, Publications of the Astronomical Society of the Pacific, 115, 1023

Walkowicz L. M., Basri G., Valenti J. A., 2013, The Astrophysical Journal Supplement, 205, 17

White T. R., Bedding T. R., Stello D., Christensen-Dalsgaard J., Huber D., Kjeldsen H., 2011, The Astrophysical Journal, 743, 161

Wohler B., Jenkins J. M., 2016, Technical report, The Kepler science processing pipeline source code road map. NASA Ames Research Center

Woodard M. F., Noyes R. W., 1985, Nature, 318, 449

Wright J. T., Marcy G. W., Butler R. P., Vogt S. S., 2004, The Astrophysical Journal Supplement, 152, 261

Wu C.-J., Ip W.-H., Huang L.-C., 2015, The Astrophysical Journal, 798, 92

Wu C. J., Krivova N. A., Solanki S. K., Usoskin I. G., 2018, Astronomy and Astrophysics, 620, A120

Yu J., Huber D., Bedding T. R., Stello D., Hon M., Murphy S. J., Khanna S., 2018, The Astrophysical Journal Supplement, 236, 42

Yuxi Lu Angus R., Agüeros M. A., Blancato K., Ness M., Curtis J. L., Grunblatt S., 2020, arXiv e-prints, p. arXiv:2008.03351

Zhang J., et al., 2020, The Astrophysical Journal Supplement, 247, 9

Zinn J. C., et al., 2020, The Astrophysical Journal Supplement, 251, 23

van Saders J. L., Ceillier T., Metcalfe T. S., Silva Aguirre V., Pinsonneault M. H., García R. A., Mathur S., Davies G. R., 2016, Nature, 529, 181