Stellar Variance for Asteroseismic Parameter Estimation and Inferences on the Evolutionary State and Binary Population of Red Giant Stars.

by

Caitlin Dawn Jones

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

School of Physics and Astronomy
College of Engineering and Physical Sciences
University of Birmingham
September 2017

Abstract

The latest generation of space missions have performed large scale observations of stars and this has been revolutionary in the field of asteroseismology. The ability to characterise thousands of stars has been instrumental in understanding the interiors of stars and the evolution of the Galaxy. This thesis focuses on studying red giant stars, both on an individual basis and as a population, using a robust asteroseismic metric we define based on the bandpass filtered estimate of the stellar variance. Here we present results of testing asteroseismic scaling relations, and the assumptions needed to create realistic simulated power spectra. The resulting synthetic datasets then inform three other investigations. We present the results of an investigation into determining the binary population of *Kepler* red giant branch stars using our variance metric. The inferred fraction of $57.4 \pm 2.5\%$ is consistent with previous work on main sequence stars. Results of using our variance metric as part of an analysis pipeline, designed to automate the detection of solar-like oscillations and determine global asteroseismic parameters in K2 and CoRoT data are presented. Finally, we present a discussion of using our variance metric to highlight structural differences between red giant branch and red clump stars.

Acknowledgements

There are many people deserving of huge thanks for the part they have played in the completion of this thesis. Firstly to Mum and Dad, thank you for your love, support, indulgence of my childhood fascination with all things space and total belief in me. My sisters, Bethan and Rhiannon, thanks for always being there and making me laugh. I promise to now find time to come to the rugby with you. Greg, my wonderful partner and biggest supporter, I couldn't have taken this on without you, thank you for everything, always.

I am lucky to have worked in a group as welcoming and full of talented people as HiROS. Everyone that has been part of my time in the group has played a part in my work, however I owe particular thanks to Professor Bill Chaplin for his ever supportive, patient and honest supervision, and Dr Andrea Miglio for his assistance as a supervisor and TRILEGAL tech support. Thanks also to Dr Guy Davies and Professor Yvonne Elsworth for all their help and collaboration. To all my office-mates through the years, James, Thomas, Mathew, Ben, Oliver, Diego and Hugo, your friendship and all our enlightning(?) conversations have been a true highlight of my time in Birmingham.

I would also to like to express my thanks to my former physics teacher, Mr Anthony Clowser, the staff of the Haematology department at QEHB, James Saxon, and Cadbury's, for the chocolate buttons that fuelled most of the last four years.

Financial support was provided by the Science and Technology Facilities Council, The University of Birmingham School of Physics and Astronomy, and the Stellar Astrophysics Centre.

Contents

List of Figures

List of Tables

Chapter 1

Introduction I: Stellar Evolution

1.1 Motivation

A new generation of space missions performing large scale surveys has been revolutionary in the study of stars. In particular, the data we now have available to us have completely opened up the field of asteroseismology. The study of the oscillations of many thousands of stars is allowing us to better understand the previously hidden inner workings of stars and has been instrumental in helping to characterize newly discovered exoplanet systems, and in determining the structure and evolution of the Milky Way.

Red giant stars are particularly good targets for asteroseismic observation. Their high luminosities means we can observe them up to large distances in the Galaxy, and their high-amplitude, low frequency oscillations allow us to make many thousands of high signal-to-noise observations which are used in a multitude of applications within asteroseismology, from precise characterisation of the stellar interior, to determining stellar age and the evolution of entire stellar populations.

In this thesis, using data from the NASA *Kepler* and K2 missions, the CNES CoRoT mission and sophisticated Galactic simulations, red giant stars are studied both on an individual basis and as an ensemble. This thesis is structured as follows; two introductory chapters covering the basics of stellar interior physics, stellar evolution and asteroseismology, followed by four chapters detailing the investigations undertaken. Chapter 3 describes the assumptions and methods used in creating realistic simulated power spectra and in defining a simple metric based on the bandpass filtered estimate of the stellar variance. The resulting synthetic datasets underpin the rest of the work. Chapter 4 presents the results of an investigation seeking to infer the population of binary stars present among *Kepler* red giants, while Chapter 5 details the development and results obtained by a software pipeline designed to determine the global asteroseismic properties of individual red giants in the CoRoT and K2 datasets. Finally, Chapter 6 again uses the stellar variance to investigate the behaviour of of red giants of different stages of evolution.

Before we can appreciate the application of asteroseismology to evolved stars, we must first understand the life cycle of these stars. The behaviour of the oscillations we study are intimately connected to the physical properties of the stars, their mass, temperature and luminosity, along with their age and evolutionary stage. While there are many different types of oscillations which have been observed, the red giant stars we are concerned with exhibit oscillations driven by the turbulent convection in their outer layers. These oscillations were first observed in the Sun, and we therefore refer to them as solar-like oscillations.

In this chapter, we will follow the evolution of a solar-type star from its birth in a collapsing cloud, through its main sequence lifetime to its red giant phase. In detail, we will look at the internal structure changes, the fuel source, and the evolution of the temperature and luminosity of the star. Only then can we understand the driving of the solar-like oscillations and the inferences we can make from them.

1.2 The H-R diagram

The Hertzsprung-Russell (H-R) diagram is a scatter plot of luminosity or absolute magnitude against temperature or colour. Stars tend to fall into specific regions of the diagram which correspond to stages of their evolution. Figure 1.1 is such a diagram showing these parameters from the Hipparcos catalogue (van Leeuwen, 2007). The structures labelled on this diagram are the stages which we will focus on; the main sequence, the red giant branch and the red clump.

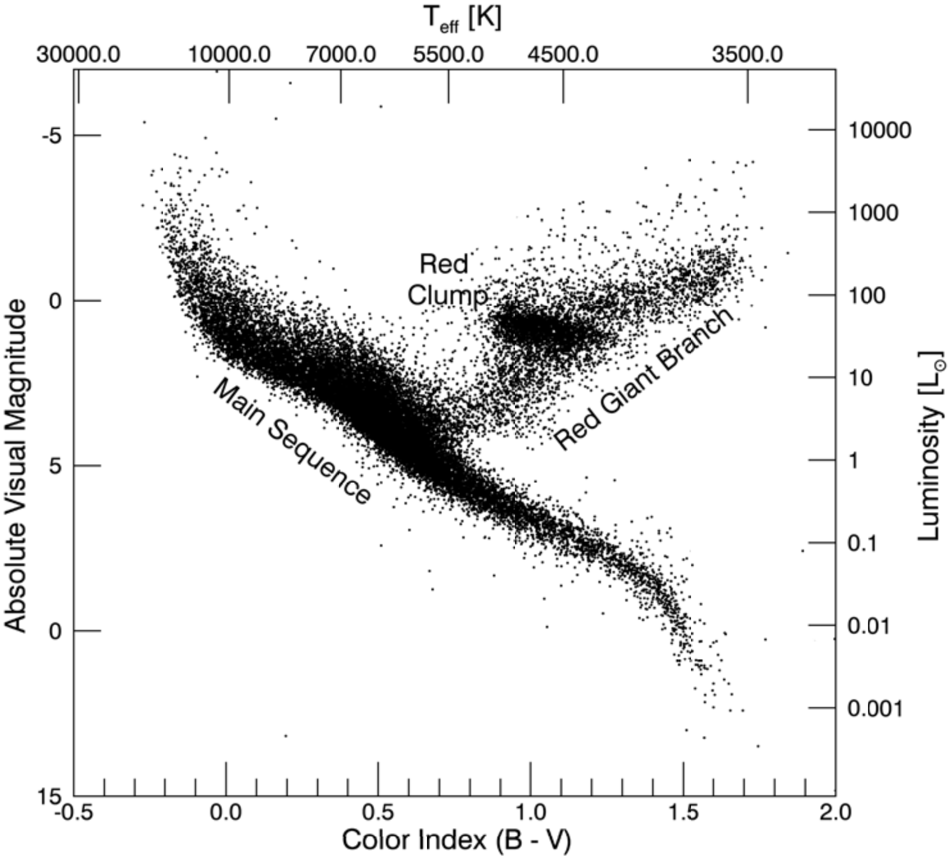


Figure 1.1: An H-R diagram created from the Hipparcos data catalog (van Leeuwen, 2007), with the main sequence, red giant branch and red clump phases labelled.

Stars move across this diagram as they evolve, and the course of the paths they take and how quickly they evolve is governed by their mass. The following discussion largely follows that of Kutner (2003), Chapters 6, 9 and 10, and my own undergraduate notes on the course delivered by Taroyan (2010), entitled *The Solar Interior*.

1.2.1 Formation and Pre-Main-Sequence Evolution

The Milky Way is made up of stars and diffuse gas and dust, primarily molecular hydrogen and helium, along with some heavier elements released in the death of stars. This is referred to as the interstellar medium. These clouds of interstellar medium vary in size, temperature and density. These molecular clouds can be thousands of times the mass of the Sun, and span tens of parsecs. They are very cool (a few tens of Kelvin) and though they are dense compared to the rest of interstellar space, they still have densities of around 10^{-20} times that of a star. They are held up against their own gravity by internal pressure. If the cloud begins to collapse due to shock waves from supernovae, or any other perturbation from a body passing close to the cloud, it begins to fragment, and these fragments are what forms stars.

Gas from the cloud begins to fall onto the area of overdensity, increasing its gravity and beginning the collapse. This forms a protostar and a disk of matter orbiting it, called the protoplanetary disc. At this point, the protostar itself is still too cool to fuse hydrogen. The energy source of the protostar is its gravitational potential energy, rather than fusion. The end of the protostar phase comes when it stops accruing material, and the cloud becomes opaque in the infrared. The new pre-main-sequence star, which now has reached its final mass, continues to contract under it's own gravity, until the temperature at its core is high enough for hydrogen fusion to begin and the collapse of the core has stopped due to the increased radiation pressure.

1.3 The Main Sequence

The band labelled in Figure 1.1 which runs diagonally across the the H-R diagram from the cool and least luminous red stars to the hottest, most luminous blue stars is called the main sequence. This is the phase in the star's life when the main energy source for the star is the fusion of hydrogen into helium in the core. The core is no longer contracting and the star is stable.

1.4 Equations of Stellar Structure

With a few simple assumptions, it is possible to describe the interior structure of our stable star using four equations. These equations relate the density, temperature and energy transport conditions of the interior, quantities which all evolve as the star ages, expands and cools.

Two of these equations relate the mass and pressure to radius, and two govern the relationship between temperature, energy transport and radius. All these equations describe the star in a steady state; they are constant in time, and assume that the star is a perfect sphere. This is a very simple model of stellar structure, referred to as the spherically symmetric quasi-static model.

1.4.1 Mass Conservation

Our first equation (Kutner, 2003) describes how the mass contained increases with radius.

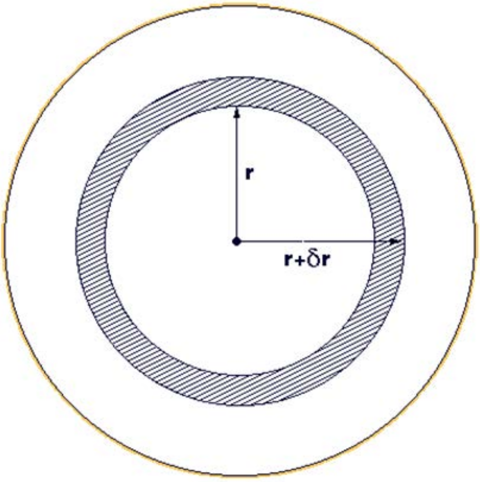


Figure 1.2: A schematic of a shell at radius r and of thickness δr within a spherically symmetric star. Adapted fromTaroyan (2010).

In the spherically symmetric star the increase in volume as we move outwards in the star by some small amount δr can be described as

$$\delta V = 4\pi r^2 \delta r, \tag{1.1}$$

then the change in mass contained over this region of the star is simply

$$\delta M = 4\pi r^2 \delta r \rho(r) \tag{1.2}$$

where $\rho(r)$ is the radius dependent density profile. This leads to the equation of mass conservation

$$\frac{dM(r)}{dr} = 4\pi r^2 \rho(r). \tag{1.3}$$

1.4.2 Hydrostatic Equilibrium

This equation (Kutner, 2003) states that the force due to the gradient of pressure within the star is balanced by the inward force of gravity. If we consider a small element of matter within the star, as a cube at radius r, as in Figure 1.3

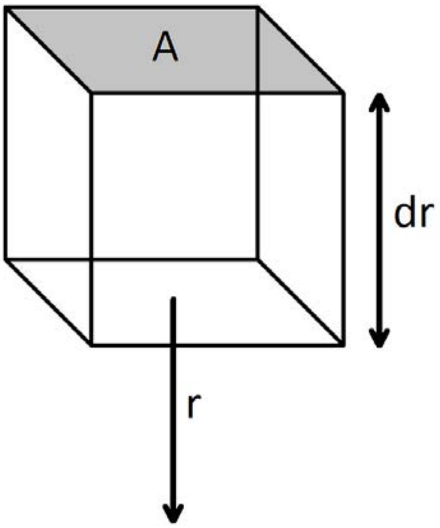


Figure 1.3: A schematic of a small element within the star. Situated at r, with height δr and area of each face A.

We simply need to balance the forces. Acting radially upwards is the pressure exerted on the lower face, and acting radially downwards is the pressure exerted on the upper face and the gravitational force due to the mass within radius r, or more simply, the weight of the element. In equilibrium, we equate these forces

$$P(r)A = P(r+dr)A + \frac{GM(r)}{r^2}\rho(r)Adr \tag{1.4}$$

and in rearranging

$$P(r+dr)A - P(r)A = -\frac{GM(r)}{r^2}\rho(r)Adr \tag{1.5}$$

and considering dr tending to zero, we can write

$$\frac{P(r+dr)-P(r)}{dr} = -\frac{dP(r)}{dr} \tag{1.6}$$

leading us to the equation of hydrostatic balance:

$$\frac{dP(r)}{dr} = -\frac{GM(r)\rho(r)}{r^2}. \tag{1.7}$$

1.4.3 Energy Conservation

This equation (e.g. Kutner, 2003) relates the rate of energy release and the rate of energy transport. Figure 1.4 shows the same diagram of a thin shell within the star, and the notation for the rate of energy flow across the two spheres of radius r and $r + \delta r$.

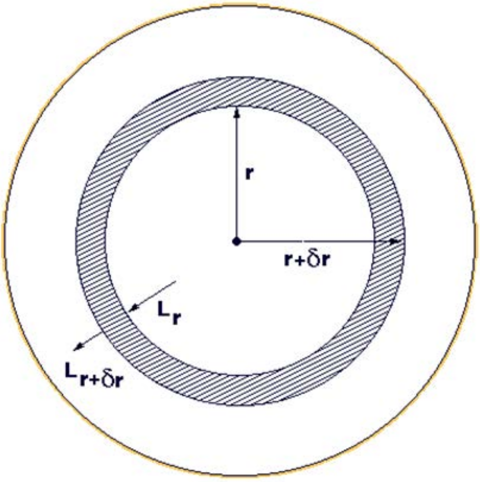


Figure 1.4: A schematic of the same shell as in Figure 1.2, with the rate of energy flow across the two spheres bounding the shell. Adapted from Taroyan (2010).

Following the same descriptions of the change in volume and mass across this shell

$$\delta V = 4\pi r^2 \delta r \tag{1.8}$$

$$\delta M = 4\pi r^2 \delta r \rho(r) \tag{1.9}$$

We can define some value ϵ as the energy released per unit time per unit mass, and hence the energy release in this shell is

$$\delta M \epsilon = 4\pi r^2 \rho(r) \delta r \epsilon \tag{1.10}$$

Again, conservation of energy requires that the energy released and transported across the shell must balance, and following the same process as the hydrostatic balance equation, we arrive at the equation of conservation of energy:

$$\frac{dL(r)}{dr} = 4\pi r^2 \rho(r) \epsilon. \tag{1.11}$$

1.4.4 Radiative Transport

We must also consider how energy is transported radially outwards in the star. The temperature gradient in the interior of the star depends on the method of energy transport. Energy is mainly transported by radiation and convection. Conduction is far less efficient than radiation. In stars which support solar like oscillations, the outer regions of the star are convective, while the deeper interior is radiative.

In radiation, the energy flux is directly proportional to the gradient of temperature

$$f(r) = -\gamma \frac{dT(r)}{dr} \tag{1.12}$$

where γ is a coefficient of conductivity. Across our sphere of radius r, the energy flow is simply described by

$$L_r = 4\pi r^2 f(r). \tag{1.13}$$

In combining these two equations, we arrive at

$$\frac{dT(r)}{dr} = -\frac{1}{4\pi r^2 \gamma} L_r \tag{1.14}$$

the equation of radiative transport (eg. Mao (2003), valid in the non-convective region of the interior. Note that the negative sign means that heat moves from regions of high temperature to low temperature.

1.4.5 Convection

In the outer layers of our solar-like oscillating star, the interior is convective, and this is the region where the oscillations are driven. Whether a region is convectively unstable depends on the temperature gradient within the region. If we consider some element in the convective zone at radius r, where it is in equilibrium with its surroundings, at pressure $P(r)$ and density $\rho(r)$.

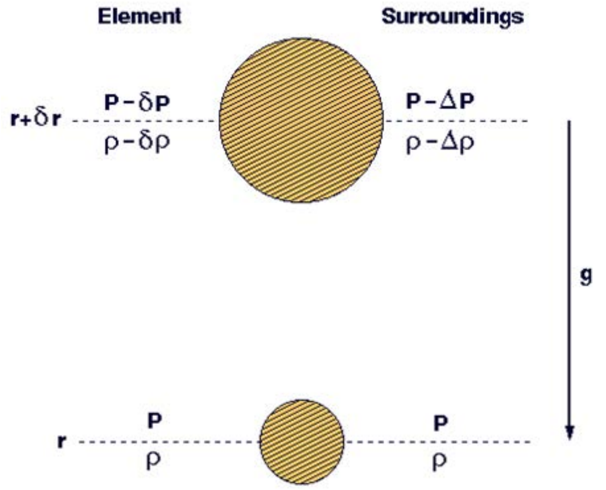


Figure 1.5: Cartoon showing the change in pressure and density of some convectively unstable element as it rises a distance of δr, as well as these properties of its surroundings. Adapted from Taroyan (2010).

It then rises to $r + \delta r$, where its pressure and density become $P + \delta P$ and $\rho + \delta\rho$, but depending on the temperature gradient, these may not be in equilibrium with the new surroundings at $r + \delta r$, which we can define as $P + \Delta P$ and $\rho + \Delta\rho$. If $\rho + \delta\rho > \rho + \Delta\rho$ then the element will sink, if the opposite is true it will continue to rise. The condition for convective instability is therefore

$$\rho + \delta\rho < \rho + \Delta\rho \tag{1.15}$$

and whether this is satisfied depends on the rate of the element's expansion due to the decreasing pressure, and the rate at which the surrounding density decreases with radius. If we assume that the element rises at a speed much less than the sound speed, then the pressure changes with radius of the element and it's surroundings are the same, or

$$\delta P = \Delta P \tag{1.16}$$

and we will also assume that the element rises adiabatically, requiring

$$PV^{\gamma} = \text{constant} \tag{1.17}$$

where γ is the adaibatic constant. As the material acts as an ideal gas, the volume of the element is inversely proportional to it's density and we can say

$$\frac{P}{\rho^{\gamma}} = \text{constant} \tag{1.18}$$

and hence

$$\frac{P + \delta P}{(\rho + \delta\rho)^{\gamma}} = \frac{P}{\rho^{\gamma}} \tag{1.19}$$

assuming that $\delta\rho$ is small, the change in density of the element can be written as

$$\delta\rho = \frac{\rho}{\gamma P}\delta P. \tag{1.20}$$

Now considering the change in density of the surroundings, $\Delta\rho$, over a small rise in radius, δr

$$\Delta\rho = \frac{d\rho}{dr}\delta r \tag{1.21}$$

the condition for convective instability can be written a

$$\frac{\rho}{\gamma P}\delta P < \frac{d\rho}{dr}\delta r \tag{1.22}$$

and recalling equation 1.16, that the pressure of the element will be the same as it's surroundings, so over a small δr

$$\frac{\delta P}{\delta r} = \frac{dP}{dr} \tag{1.23}$$

or

$$\frac{\rho}{\gamma P}\frac{dP}{dr} < \frac{d\rho}{dr} \tag{1.24}$$

and dividing by $\frac{dP}{dr}$ and rearranging:

$$\frac{P}{\rho}\frac{d\rho}{dP} < \frac{1}{\gamma} \tag{1.25}$$

We can combine this with our assumption of an ideal gas:

$$\frac{dP}{P} = \frac{d\rho}{\rho} + \frac{dT}{T} \tag{1.26}$$

and arrive at the condition for convection, in terms of the temperature gradient:

$$\frac{P}{T}\frac{dT}{dP} > \frac{\gamma - 1}{\gamma}. \tag{1.27}$$

If this condition is satisfied, then large scale motions of plasma transport energy radially outwards within the star. This is essential for the driving of solar like oscillations, and is present in the outer convective regions of all the solar like oscillators.

1.4.6 Interior Structure

The interior of the star, as governed by the above equations has both radiative and convective zones. The radiative zone of a sun-like star is in the deep interior, while the outer regions of the star are convective. Figure 1.6 is a simple cartoon of the interior structure of a main sequence star. Typical radii of these regions are given by modelling, in a $1M_{\odot}$ star, the core occupies the inner $\approx 25\%$ by radius, this is overlaid by the radiative zone out to 70% of the radius, and the outer 30% is the convective zone (Christensen-Dalsgaard et al., 1996).

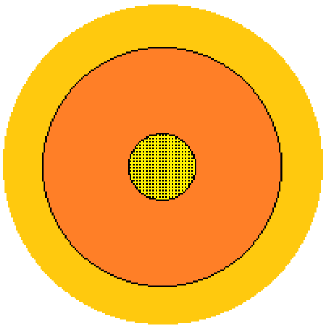


Figure 1.6: A cartoon diagram of a main sequence star. The outer region of the star is convective, the deeper interior is radiative, and the shaded region is the core, where nuclear fusion takes place.

The definition of the main sequence is that the star produces it's energy by hydrogen fusion in the core of the star.

1.4.7 Nuclear Fusion

The hydrogen core of the star is the source of fuel for the entire main-sequence lifetime of the star. Through nuclear fusion, hydrogen is converted into helium in a reaction called the proton-proton chain. The below combination (e.g. Kutner, 2003) of two protons, or hydrogen nuclei (${}^{1}_{1}\mathrm{H}$);

$$
{}^{1}_{1}\mathrm{H} +^{1}_{1}\mathrm{H} \rightarrow^{2}_{1}\mathrm{H} + (\mathrm{e}^{+}) + \nu \tag{1.28}
$$

forms a deuterium nucleus, a positron and a neutrino. The deuterium nucleus then reacts with another proton

$$^{2}_{1}\mathrm{H} +^{1}_{1}\mathrm{H} \rightarrow^{3}_{2}\mathrm{He} + \gamma \tag{1.29}$$

to form Helium-3 and a gamma ray. Two of these Helium-3 nuclei then combine to result in

$$^{3}_{2}\mathrm{He} +^{3}_{2}\mathrm{He} \rightarrow^{4}_{2}\mathrm{He} +^{1}_{1}\mathrm{H} +^{1}_{1}\mathrm{H}. \tag{1.30}$$

The energy carried by the gamma ray and positron heat the surrounding gas which then releases photons.

1.5 After the Main Sequence

The details of post-main-sequence evolution depends on the mass of the star. In stars with a mass less than around $2.25\mathrm{M}_{\odot}$ (though this is dependent on metallicity, e.g. Girardi (1999)), the temperature in the helium core of the star is too cool to undergo fusion itself and it contracts under gravity until it's density is so high that it is supported by electron degeneracy pressure. The hydrogen above the core is compressed by the gravitational pull of the core and therefore heated to the point that hydrogen fusion restarts in a "shell" surrounding the core. This in turn heats the outer layers of the star, causing them to expand, then cool. This leads to a cooler, larger star with a slightly increased luminosity, referred to as a subgiant.

1.6 The Red Giant Branch

Due to the larger temperature gradient between the hydrogen burning shell and the surface after this initial expansion, the rate of energy transport outwards increases (see above discussion). The temperature of the surface stays approximately constant as it continues to expand and consequently, the luminosity increases. At this stage, the star lies in the red giant branch in the Hertzsprung-Russell diagram. This is labelled in Figure 1.1, note that these stars occupy a much narrower range of temperature than the main sequence.

1.6.1 Interior Structure

The structure of these more evolved stars is slightly different than during the main sequence. Having exhausted the supply of hydrogen in the core, hydrogen fusion now takes place in a shell surrounding an inert, helium core. There is still a radiative zone and a convective zone, however now the outer envelope of the star has expanded dramatically, causing the star to be much more luminous and have a much deeper convective zone than during its main sequence lifetime. Figure 1.7 is a schematic drawing of the structure of a red giant branch star.

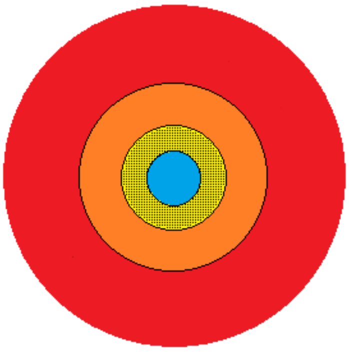


Figure 1.7: A cartoon diagram of a red giant branch star. The outer region of the star is convective and the star now has an inflated outer envelope, and fusion takes place in a hydrogen shell surrounding an inert helium core.

Hydrogen fusion continues in a shell around the core. This process forms helium, which accretes onto the inert core, causing it to become more massive, and to begin contracting under its own gravity, and as the density increases, the core gets hotter. The density eventually becomes so great that the electrons become degenerate, and therefore incompressible, and the core stops contracting. The hydrogen burning shell continues to deposit helium onto the degenerate core, which continues to increase in temperature without increasing in pressure. When it reaches around 10^8 K, helium fusion begins.

1.7 The Red Clump

These now core-helium-burning stars occupy a narrow range of luminosities due to them having a common core mass (Bedding et al., 2011). These stars occupy the Red Clump in the H-R diagram, as shown in Figure 1.1.

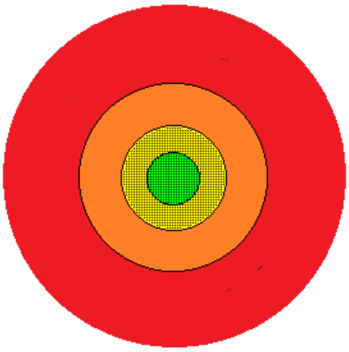


Figure 1.8: A cartoon diagram of a red clump star. The outer region of the star is convective and the star now has an inflated outer envelope, and fusion takes place in a hydrogen shell and in the helium core.

At 10^8 K, helium fuses into carbon and energy through the triple alpha process (e.g Kutner (2003)).

$$ {}^4_2\mathrm{He} + {}^4_2\mathrm{He} \rightarrow {}^8_4\mathrm{Be} + \gamma \tag{1.31}$$

$$ {}^8_4\mathrm{Be} + {}^4_2\mathrm{He} \rightarrow {}^{12}_6\mathrm{C} + \gamma \tag{1.32}$$

This process continues to heat the core, and in its degenerate state, the core does not obey the ideal gas laws and does not expand with the increase in pressure and temperature. The extreme high density and temperature of the core causes a runaway nuclear reaction, releasing a huge amount of energy in an event known as the helium flash. This lasts minutes (Dearborn et al., 2006). Due to this release of energy, the temperature in the core is raised sufficiently that thermal pressure lifts the degeneracy, the core can expand and continues fusing helium at a stable rate.

In stars with a mass of $> 2.25\mathrm{M}_\odot$, the temperature needed for helium fusion is reached before the core becomes degenerate and so the star does not go through the flash. Instead, helium fusion occurs at a stable rate and the stars span a range of core masses, and hence, luminosities. These stars appear as a much more disperse grouping in the H-R diagram and are referred to as the secondary clump. Red Clump stars and Red Giant Branch stars can have very similar surface properties while having very different interior properties. Asteroseismology can diagnose these different evolutionary states, even in these outwardly very similar stars, and we will touch on this in Chapter 6.

1.8 Multiple Star Systems

While the above discusses the formation and evolution of a single star, this is certainly not the case for many stars in our Galaxy. It is common for two stars to form from the same collapsing gas cloud, and for these stars to remain gravitationally bound throughout their lifetimes. We leave discussion of binary systems to Chapter 4. This contains discussion of how it is possible to detect these binary pairs, and also the investigation based on asteroseismology aiming to determine the overall fraction of red giants which exist in these binary systems.

1.9 Up next...

Following is a discussion of the background physics and techniques of asteroseismology necessary to understand the work undertaken, the differences in the oscillations between red giant branch and red clump stars and how we can use asteroseismology to distinguish between the two, and an introduction to the filtered variance metric used to diagnose binarity and in automating the oscillation detection process.

Chapter 2

Introduction II: Asteroseismology

The interior structures of stars discussed in Chapter 1 are not observable directly by any telescope. Arthur Eddington opens his book *The Internal Constitution of the Stars* (Eddington, 1926) with the now classic comment

> At first sight it would seem that the deep interior of the Sun and stars is less accessible to scientific investigation than any other region of the universe. Our telescopes may probe farther and farther into the depths of space; but how can we ever obtain certain knowledge of that which is hidden behind substantial barriers? What appliance can pierce through the outer layers of a star and test the conditions within?

This “appliance” turned out to be asteroseismology. Asteroseismology uses the natural oscillations present in stars to study their otherwise invisible interiors. The term itself comes from the Greek *aster* meaning star and *seismos* meaning tremor. It is closely related to the older field of helioseismology, which provides important information about the conditions in the solar interior which inform stellar models today.

It was not until 1960 that the first observations hinting at a way to observe the solar interior were made. A scanning spectroheliographic technique detailed in Leighton et al. (1961) measured the Doppler shift in a spectral line. The field of view was moved across the disc of the Sun, capturing an image in a single wavelength of light. Slight blue shifts from the spectral line were captured on one image and slight red shifts on the other, and these were then superimposed. Two separate observations of this type were made, and subtracting these two images from each other shows the change in velocity on the surface. Figure 2.1 below shows an example of a Doppler plate, taken from this seminal work.

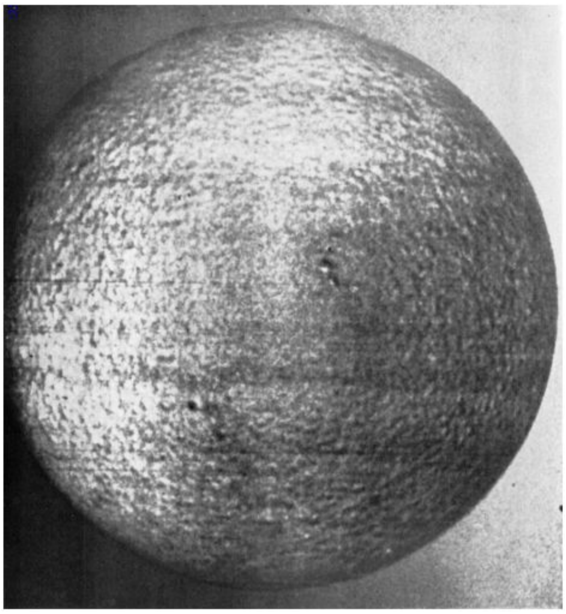


Figure 2.1: A Doppler difference plate, where dark areas are receding and light areas approaching. (Leighton et al., 1961)

The surprising result was that a structured velocity field with 5-minute periodicity was evident, not only random convection patterns as was expected. This was the beginning of what we now call helioseismology. However it was not until a paper by Ulrich (1970) that these periodic velocity changes were attributed to trapped acoustic standing wave modes oscillating within a spherical cavity within the Sun.

There are many different classes of pulsating star, as shown in the pulsation H-R diagram in Figure 2.2. Their oscillations have various different driving mechanisms and types of oscillation, and asteroseismology techniques can be used to study all of these stars, however, the red giant stars we are studying reside in the solar-like regime. Solar-like refers to the driving and damping of the oscillations in the same way as happens in the Sun, in the outer convective region of the star.

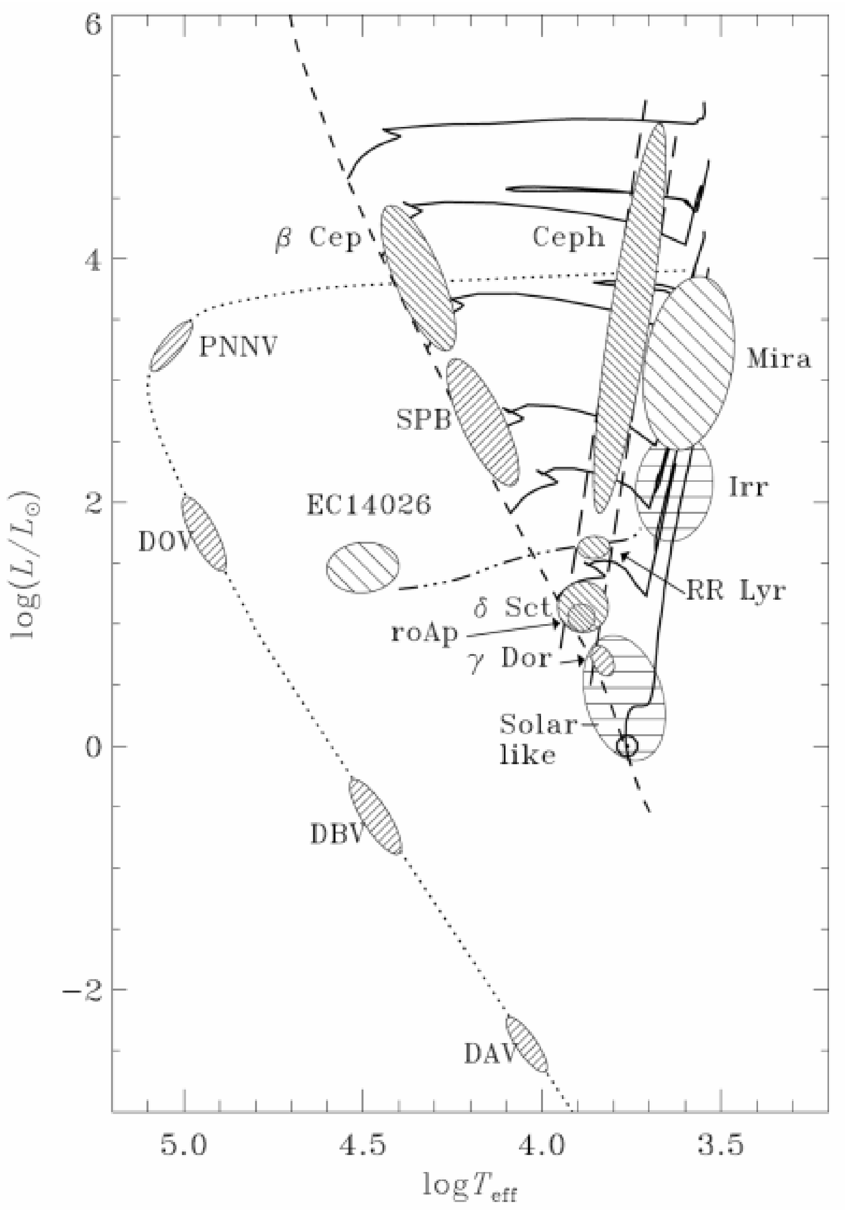


Figure 2.2: An H-R diagram showing the various classes of pulsating star which can be studied using asteroseismology. (Christensen-Dalsgaard, 1998)

In this chapter, the physical processes behind these oscillations will be introduced, as well as some of the notation used to describe the oscillations. We will go on to find out what we can observe of these modes; how they appear at the stellar surface, and the techniques we use to observe them. These observations are typically studied in the form of power spectra. We will learn how power spectra are created, and the interpretation of the parameters we can extract from them. We will also discuss the instruments and simulation algorithms which provide the data we use.

2.1 Basic Asteroseismology

2.1.1 Spherical Harmonics

Modes are described by spherical harmonics as solutions to the spherical wave equation and can be described by three wavenumbers: the radial order n, describing the number of nodes the mode has along the stellar radius; the spherical degree l, which describes the number of nodes on the stellar surface; and the azimuthal order m, which describes how many of the node planes bisect the equator. (Christensen-Dalsgaard, 2003). The frequency of a mode is largely dependent on l and n, and is described to good approximation by:

$$\nu_{\mathrm{nl}} = \Delta\nu\left(n + \frac{l}{2} + \epsilon\right) + \delta\nu_{ll+2}(n). \tag{2.1}$$

The terms above are defined as follows: $\Delta\nu$ is the inverse of the sound travel time across the star, and is the dominant frequency spacing in the spectrum. This is referred to as the large frequency separation, and is the spacing between modes of consecutive order n and the same l. ϵ is a small phase shift due to reflection at the upper turning point (Basu and Chaplin, in the press), and it is dependent on the conditions near the surface of the star. Following the first term in Equation 2.1, there is a degeneracy in the frequencies such that $\nu_{nl} \approx \nu_{n+1l-2}$. The second term is referred to as the small frequency separation and lifts this degeneracy.

2.1.2 Mode Excitation and Damping

The outer layers of the solar-type and red giant stars are convective, and it is turbulent motions of the plasma in these near surface layers which stochastically drive and damp the oscillations. This leads to essentially random driving. A typical analogy to use is to imagine a bell being continuously struck with sand. The random strikes of the sand grains will eventually cause the bell to ring at it's resonant frequency. The same is true for our oscillating star, however it will have thousands of resonant frequencies.

2.1.3 Modes at the Stellar Surface

In stars other than the Sun, only low-l modes can be observed due to the fact that the stellar surfaces cannot be resolved, and the effect which the perturbations of the stellar surface caused by high l-degree modes have on the light intensity average out across the stellar surface for modes of greater than $l = 2$ in most stars, though $l = 3$ modes are detectable in some stars. A representation of the appearance of these modes is shown by Figure 2.3, though of course the true perturbations are a superposition of these and all the other modes excited in the stars.

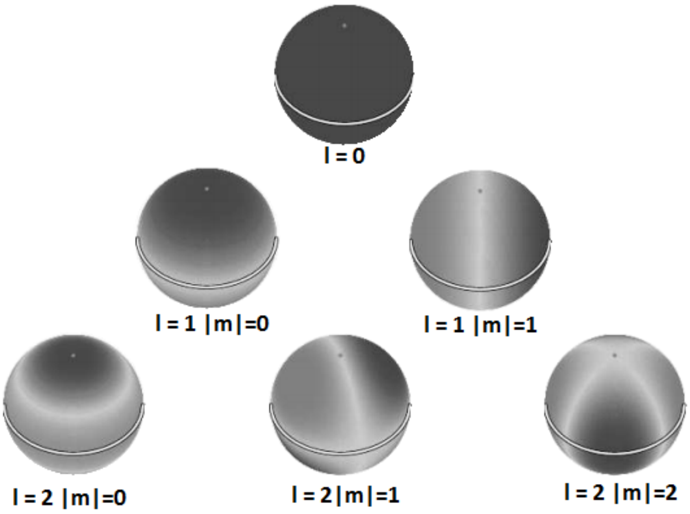


Figure 2.3: The displacement of the stellar surface at the extremes of pulsation cycles for the individual modes which are detectable on most stars which display solar like oscillations. Dark areas are expanding and light areas are contracting. Figure adapted from Telting and Schrivers (1997).

It is the superposition of these and thousands more modes which were observed by Leighton et al. (1961). The five-minute periodicity corresponds to the frequency at which the Sun's oscillations are strongest. The propagation of the oscillation modes cause motions of the stellar surface, and it is these motions which we are able to measure by the Doppler shifts as in the spectroheliograph technique and the Birmingham Solar Oscillations Network (BiSON) (Chaplin et al., 1996), but also as fluctuations in the luminosity of the star. As part of the stellar surface expands, it cools and becomes dimmer and when it contracts, it heats and becomes brighter. This causes tiny but detectable variations in the light levels measured from that star, and it is these which we use to study the oscillations of other stars with data from the *Kepler* and *CoRoT* space missions. This technique involves measuring the brightness of the star to great precision for a long period of time, in *Kepler's* case for 4 years. This time series data can be studied in terms of the frequencies by creating a power density spectrum.

2.1.4 The Fourier Transform

In basic terms, any time series signal (here, our lightcurve) can be described as a superposition of a number of sinusoids. The Fourier transform is a tool used to recover these sinusoids such that they can be studied in terms of their amplitude and frequency. It is common to express these sinusoids as the sum of a series of complex exponentials, and so the Fourier transform $\mathcal{F}(\mu)$ of a time series signal $f(t)$ may be defined as

$$\mathcal{F}(\nu) = \int_{-\infty}^{\infty} f(t) e^{-i2\pi t\nu} dt. \tag{2.2}$$

Since our data are discretely sampled observations rather than a continuous functions, it is correct to use the discrete Fourier transform, which assumes that the signal is composed of N samples which have a regular sample length of Δt, we would then instead of Equation 2.2 use

$$\mathcal{F}_{\text{discrete}}(\nu) = \Delta t \sum_{n=1}^{n=N} e^{-i2\pi\nu t_n}. \tag{2.3}$$

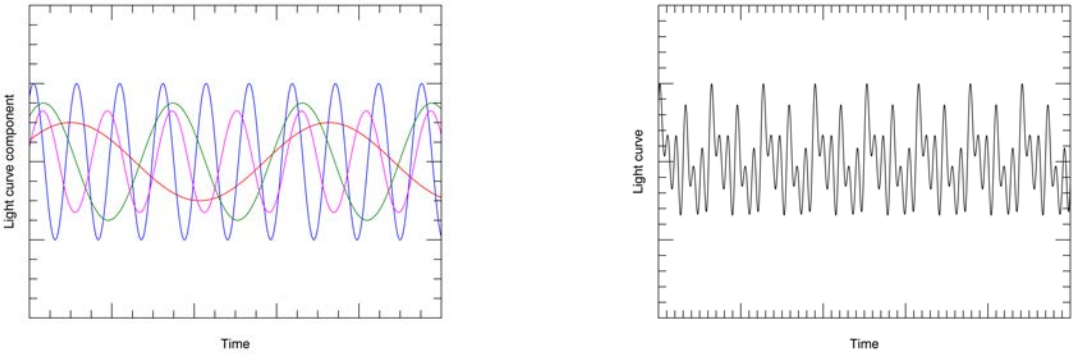
Since our oscillations are excited stochastically, meaning that they are driven by random fluctuations in the turbulent convection zone, the phase information contained in the complex part of the transform is not useful, we typically deal with the frequency spectrum in terms of power spectral density (PSD), rather than amplitude, and this can be defined as

$$\mathrm{PSD}(\nu) = |\mathcal{F}_{\mathrm{discrete}}(\nu)|^2. \tag{2.4}$$

There is an issue with using a true Fourier transform for this work. The discrete Fourier transform assumes a regular sampling time Δt. For *Kepler* data, this isn't strictly true. While the sampling time is regular at spacecraft time, since *Kepler* is moving, the sample length is not truly regular, and this along with any gaps due to instrumental issues will induce noise features into the transform.

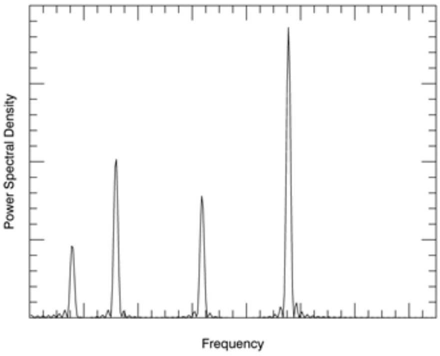
To mitigate this problem, we instead employ a technique called the Lomb-Scargle periodogram (Lomb, 1976), a technique similar to a Fourier transform which is a method of estimating the power spectrum based on fitting sinusoids to the signal directly, using a least squares fit.

The resulting transform is referred to as the power spectral density, or more commonly, a power spectrum. The power spectrum of a time series shows how energy is distributed among the frequency components which form the signal in time. Figure 2.4 shows a toy example of the use of a power spectrum. Figure 2.4(c) is the power spectrum of this toy light curve. The four components of different frequency and power are clearly appreciable.



(a) Four sinusoidal components of different amplitudes and frequencies.

(b) The superposition of the four components. This is our toy light curve.



(c) The Lomb-Scargle periodogram of the time series, recovering the frequency of the four components above.

Figure 2.4

The light curve in Figure 2.4(b) is the superposition of the four sinusoidal components shown in Figure 2.4(a). For the real time series data that this thesis is concerned with, the picture is of course more complicated than this noiseless toy example consisting of a few simple components.

2.1.5 Components of a Solar-like Power Spectrum

The form of the power spectrum is common to all solar-like oscillators in that it consists of a background signal, made up of signatures from stellar activity, granulation and instrumental effects such as shot noise, along with a clear excess of power above the background where the modes are found. Figure 2.5 shows an example power spectrum with the noiseless version (or limit spectrum) of the granulation and white noise components along with the smoothed mode region overplotted.

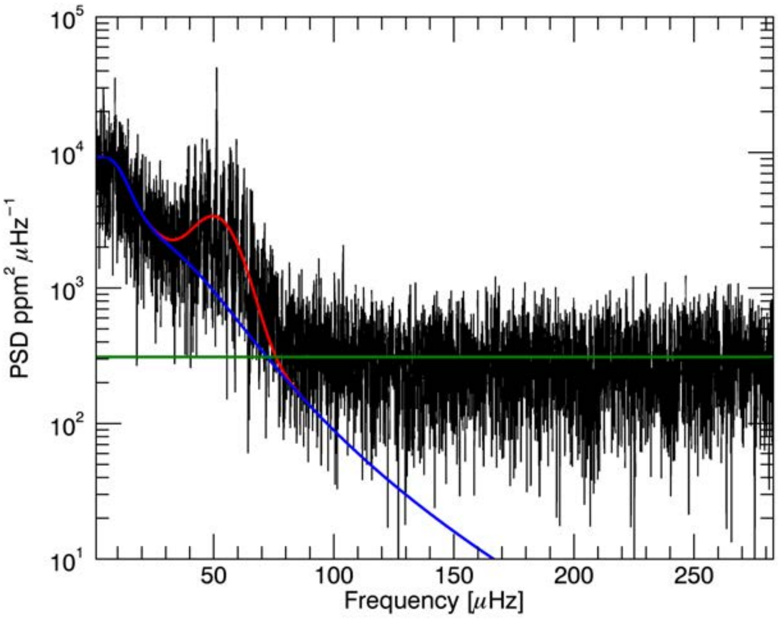


Figure 2.5: An example power spectrum of a *Kepler* solar-like oscillator. The three components shown are the granulation background in blue, the smoothed mode envelope in red, and the shot noise in green. Data courtesy García et al. (2011).

The exact forms of these components are discussed at length in Chapter 3. The white noise component comes from photon noise on the light curve observations, the granulation background is due to the presence of granules on the stellar surface. These granules are the appearance of convective cells at the surface of the star. These granules increase and decrease in brightness with a characteristic timescale and amplitude, appearing in the power spectrum in a form that closely resembles a zero-centred Lorentzian profile.

It is perhaps useful to think of the individual oscillations temporal behaviour as a damped harmonic oscillator, and therefore the individual modes take the form of a Lorentzian centred on the frequency of the acoustic mode as defined in equation 2.1, where their width reflects the lifetime of the mode. The Gaussian envelope shown in red in Figure 2.5 which modulates the heights of the mode profiles is a phenomenological description, however it is intuitive given that the granulation responsible for driving the modes has a characteristic frequency, that we might expect the power of the modes to be centred on this frequency and have a gradual fall-off around this.

There are several important asteroseismic parameters we can extract from the power spectrum which are related to the global properties of the star. Asteroseismic scaling relations associate these seismic properties of a star with its intrinsic properties such as mass, luminosity and temperature, and are used to determine quantities such as radius and surface gravity of stars.

Our first asteroseismic parameter is $\nu_{\rm max}$; the frequency where we observe the maximum oscillation power. This corresponds to the peak of the Gaussian envelope shown in Figure 2.5, and is the frequency which contributes the most power to the light curve, and hence the frequency of the mode of highest amplitude. The maximum frequency at which acoustic modes are trapped within the star is called the acoustic cutoff frequency of the atmosphere of the star (Brown et al., 1991). $\nu_{\rm max}$ is proportional to the acoustic cutoff frequency $\nu_{\rm ac}$, which scales as

$$\nu_{\rm max} \propto \nu_{\rm ac} \propto g T_{\rm eff}^{0.5} \tag{2.5}$$

where g is the stellar surface gravity and $T_{\rm eff}$ is the effective temperature. In writing g in terms of mass and radius $g \propto MR^{-2}$ and then radius in terms of luminosity and temperature from $L \propto R^2 T_{\rm eff}^4$, we come to the relation

$$\nu_{\rm max} = \nu_{\rm max,\odot} \left(\frac{M/M_\odot}{L/L_\odot} \right) T^{3.5} \tag{2.6}$$

from e.g. Kjeldsen and Bedding (1995). This is scaled against the solar value of $\nu_{\rm max,\odot} = 3090\ \mu$Hz. This can lead us to an estimate of stellar mass.

Another feature of the power spectrum is the large frequency separation, $\Delta\nu$. See Figure 2.6 for an example power spectrum with $\Delta\nu$ marked. $\Delta\nu$ denotes the frequency spacing between modes of the same degree ℓ and consecutive order n.

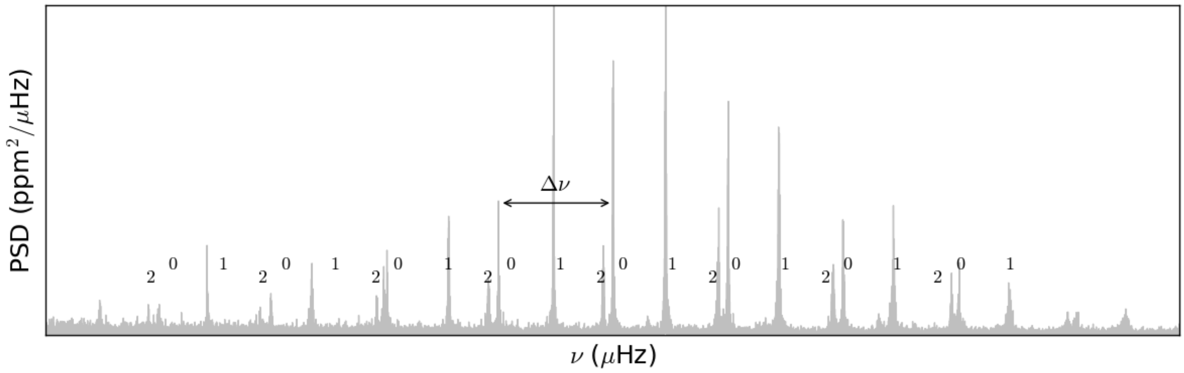


Figure 2.6: A mode oscillation spectrum of a *Kepler* solar-like star, showing the large frequency separation $\Delta\nu$. Modes are numbered according to degree ℓ.

This scales as

$$\Delta\nu \propto \rho^{0.5} \tag{2.7}$$

following Christensen-Dalsgaard (2003). This spacing is also equivalent to the inverse of the time it takes for the acoustic mode to cross the star. We can get this into the form of observable values and mass by considering that the density $\rho \propto MR^{-3}$. Substituting in luminosity and temperature for radius we recover

$$\Delta\nu = \Delta\nu_\odot \frac{(M/M_\odot)^{0.5}}{(L/L_\odot)^{0.75}} T^{3.0} \tag{2.8}$$

where $\Delta\nu_\odot = 135\mu$Hz.

Another important parameter is the maximum amplitude of the oscillations. A$_{\text{max}}$ denotes the amplitude of an $\ell = 0$ mode at the frequency of maximum oscillation power, ν_{max} (Verner et al. (2011)). The more evolved the star, the larger its radius, the lower the frequency of the oscillation envelope and the higher the amplitude of the oscillations. This amplitude can be written in the same form

$$A_{\text{max}} = A_{\text{max},\odot} \frac{M}{M_\odot}^{a} \frac{L}{L_\odot}^{b} \frac{T_{\text{eff}}}{T_{\text{eff},\odot}}^{c} \tag{2.9}$$

following Kjeldsen and Bedding (1995), for example. The exponents in Equation 2.9 depend on the assumptions adopted, and these are discussed at some length in Chapter 3, as well as other asteroseismic parameters related to the power in the modes and granulation. The above serves only as a brief introduction, in particular to $\nu_{\max}$ and $\Delta\nu$ which we will see again throughout this thesis.

2.1.6 $\ell = 1$ Modes in Evolved Stars

In main-sequence stars, g-modes, which have gravity (buoyancy) as their restoring force are confined to the inner radiative regions, due to them being evanescent in the convective outer layers of the star. In more evolved stars, due to the contraction of the core region, the frequencies of g-modes increase with the increase in the buoyancy frequency, which is the frequency at which a region of the plasma will oscillate if it is displaced radially. They are usually described by an asymptotic relation in period, rather than frequency, in the form (e.g Basu and Chaplin (in the press), Christensen-Dalsgaard (2003));

$$\Pi_{nl} = \nu_{nl}^{-1} \simeq \Delta\Pi_l(n + \epsilon_g). \tag{2.10}$$

As the frequencies of the g-modes increase with the core's contraction, they enter the frequency range where we find the p-modes, and they can interact with modes of the same degree l, causing the observed modes to display characteristics of both g- and p-modes. The frequencies of the p-modes which the g-modes couple with are "bumped" due to these interactions, and result in red giant oscillation spectra being far more complicated than that of a main-sequence star, which only displays pure p-mode behaviour. Radial ($l = 0$) p-modes are not affected by this as the buoyancy perturbations, which result in g-modes, do not support radial oscillations, and so there is no $l = 0$ g-mode to couple with the $l = 0$ p-mode. The coupling between the g- and p-modes shifts the p-mode frequencies from those defined by Equation 2.1. This coupling is strongest for the $l = 1$ modes.

This is illustrated well by échelle diagrams, examples of which are shown in Figure

2.7. These diagrams are produced by dividing the power spectrum into frequency moduli the size of the large frequency separation $\Delta\nu$ and stacking them in order of frequency. This displays modes of the same degree n as near-vertical ridges (Bedding et al., 2010). The first panel of Figure 2.7 shows such a diagram for 16 Cyg A, a solar analog star. The ridges showing the regions of mode power are well defined because they are pure p-modes and hence obey the relation in Equation 2.1. In the case of the more evolved stars, the modes depart dramatically from that pattern, such as in the final panel of Figure 2.7, due to the effect of the g-mode coupling on the oscillation frequencies.

In red giant stars, the signatures of this g-mode coupling can be used to discriminate between the core helium burning red clump stars and the hydrogen shell burning red giant branch stars. It has been shown (e.g by Bedding et al. (2011)) that the period spacing of the g-modes is substantially different in these two stages, and so provides a way of ascertaining which stage of evolution a star is at. RC and RGB stars have very similar surface properties, so this property of the g-mode spacing can be a very valuable tool.

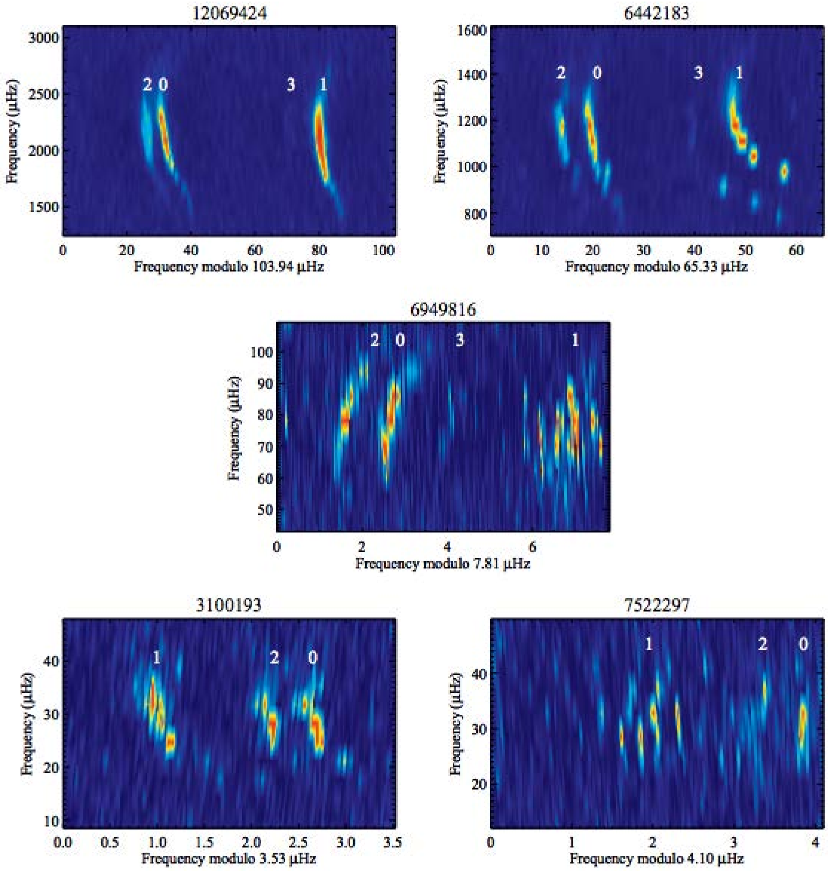


Figure 2.7: Example of échelle diagrams for stars at different evolutionary stages; from the top left, a main-sequence star 16 Cyg A, a subgiant, two RGB stars and an RC star. KIC 3100193 and KC 7522297 have very similar surface properties but have very different $l = 1$ mode structure. Image from Chaplin et al. (2013).

Elsworth et al. (2017) are currently using a method based on this principle to classify red giants as either RGB or RC. Finding the period spacing of these mixed modes usually relies on being able to isolate the individual modes and determining an accurate period spacing; something which can be difficult in the often complicated spectra of these evolved stars, while this method evaluates frequency and period differences between significant features in the spectrum. The regions of frequency containing the modes of even and odd degree l are separated by cross correlating the spectrum with the so-called "Universal Pattern" (Mosser et al., 2011) of the modes expected for red giants. The number of occurences of the different frequency differences between the features in both the regions for the odd and even modes are represented in histograms. In the even l region, the histogram is dominated by the large and small frequency spacings ($\Delta\nu$ and $\delta\nu_{\mathrm{ll+2}}(\mathrm{n})$, see

Equation 2.1.) In the odd l region, the coupling with the g-modes has an effect on the shape of this histogram by bumping the frequencies. Because the periods of the g-mode are very different in red clump stars and RGB stars, the shape of the histogram is used to discriminate between these two evolutionary states. Results from their method to separate the RGB and RC stars from each other is employed for the work presented in Chapters 4 and 6.

2.2 Observations and Data

2.2.1 *Kepler*

The NASA *Kepler* spacecraft was launched in 2009 with the objective of discovering and studying planetary systems other than our own, including determining the properties of planet-hosting stars (Koch et al., 2010). It monitored the brightness of thousands of stars to parts-per-million precision using a 0.95m aperture telescope and 21 CCD modules continuously trained on a single field of stars covering over 100-square-degrees of sky as shown in Figure 2.8. In order to keep the solar panels illuminated and the radiator which keeps the optics cool pointing away from the Sun, it was rolled 90° about it's line of sight every 93 days. Each of these 93 day blocks are referred to as a "quarter." *Kepler* had two different sampling intervals, referred to as short and long cadence, which last 58.85 seconds (Gilliland et al., 2010) and 29.4 minutes (Jenkins et al., 2010) respectively. Long cadence observations were performed on over 150,000 targets simultaneously, while the much more frequent sampling of the short cadence were limited to 512 targets.

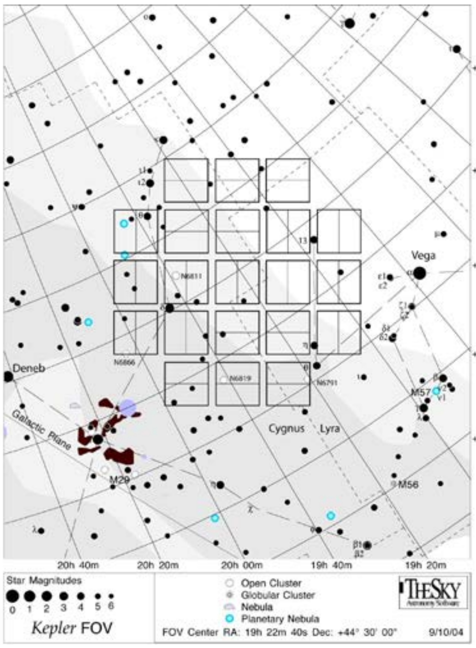


Figure 2.8: The *Kepler* field of view, showing the 42 CCDs in 21 modules. Credit NASA.

Kepler searches for exoplanets using what is known as the transit method, observing the extremely small periodic drops in brightness which occur when a planet passes across the face of the star. This relies on the plane of the orbit being aligned sufficiently close to the line of sight that the transit or eclipse which happens during the orbit of the planet is visible to us. Only a small fraction of stars will be inclined this way, so *Kepler* observes over 150,000 stars to increase the chance of detecting Earth-like planets. This method among others is also used to detect binary stars. These high precision data are excellent for studying the stars themselves in great detail, and are invaluable to the field of asteroseismology.

We use *Kepler* data from quarters Q0-Q13 which have been corrected following García et al. (2011) and Mathur et. al (in prep.). These corrections include removing outliers thought to be caused by micrometeorite impact, jumps caused by attitude tweaks to the spacecraft and drifts due to temperature changes after safe mode events. Then consideration is given to merging the observations across each quarter, since when the spacecraft rolls, the star is observed on a different CCD module with different characteristics. For the red giants, the resulting corrected light curve is also filtered to remove very low frequency trends in the CCD performance, allowing the granulation signal to be investigated. An example of these data are shown in Figure 2.9.

The power density spectra describe how the power of the light curve (a time series) is distributed over the different frequencies, up to the Nyquist frequency of the *Kepler* long cadence at 283μHz. In the example in Figure 2.9, these peaks are centred around 35μHz. In what follows, we will build artificial power spectra intended to mimic in a realistic way the real data shown here.

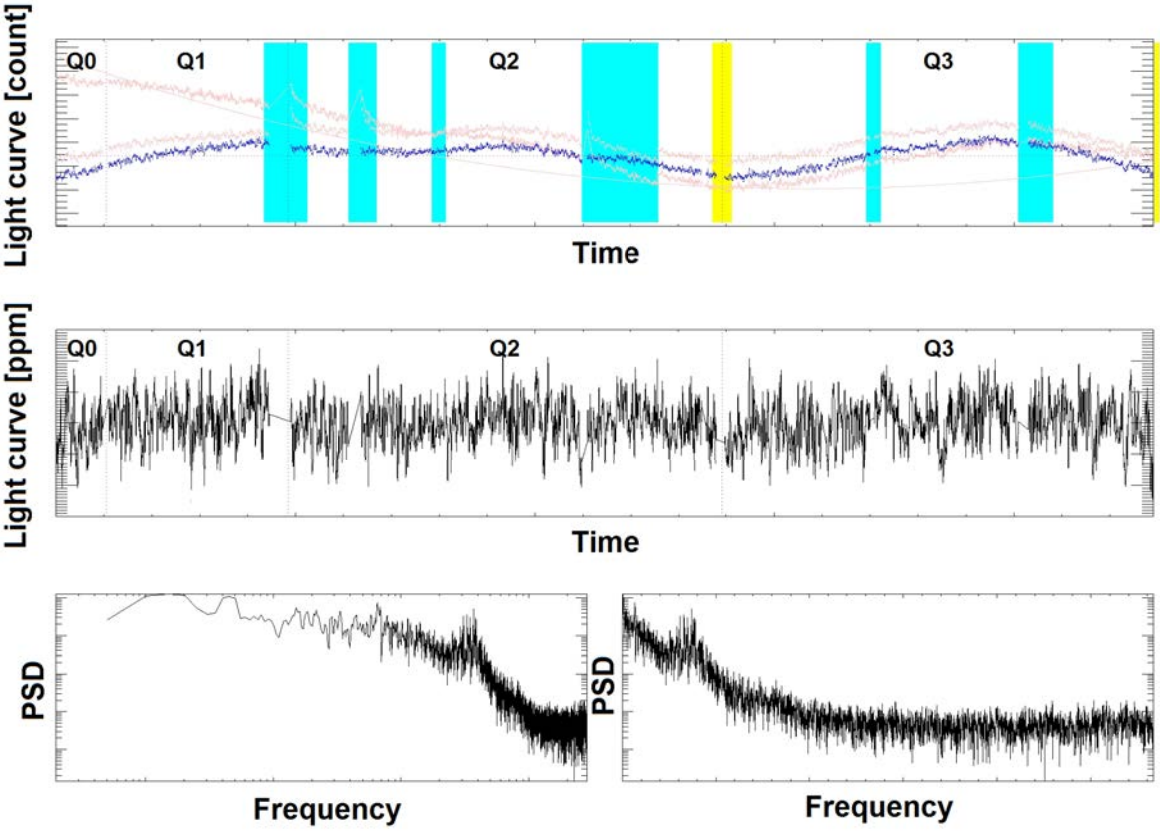


Figure 2.9: An example of a light curve used in this work. The top panel shows the raw and uncorrected light curves, the centre panel shows the filtered light curves and the bottom two are power density spectra.(García et al., 2011)

For all *Kepler* stars analysed in this work, we have taken values of the frequency of maximum oscillation power $\nu_{\max}$ and the large frequency spacing $\Delta\nu$ from work by Stello et al. (2013) and values of the *Kepler* bandpass magnitude from the Kepler Input Catalog (Brown et al. 2011.)

2.2.2 K2

The K2 mission (Howell et al., 2014) is a continuation of the *Kepler* mission. After two of the reaction wheels that the spacecraft uses to orient itself correctly failed, a new plan was presented. *Kepler* orbits in what is termed an Earth trailing orbit, falling further behind Earth over time. This means it cannot be repaired due to being simply too far away. It was decided to instead stabilise the spacecraft using solar radiation pressure. This meant, however, that it was no longer viable to observe the original *Kepler* field shown in Figure 2.8, since the new orientation of the spacecraft meant that it has to be turned every 83 days to prevent sunlight from contaminating the observations. These shorter observations combined with lower precision have changed what can be expected from the data, but K2 has continued to be a successful asteroseismology and planet hunting mission.

Figure 2.10 shows the existing and future fields of observation for K2. Each ≈ 80 day observing run is referred to as a campaign.

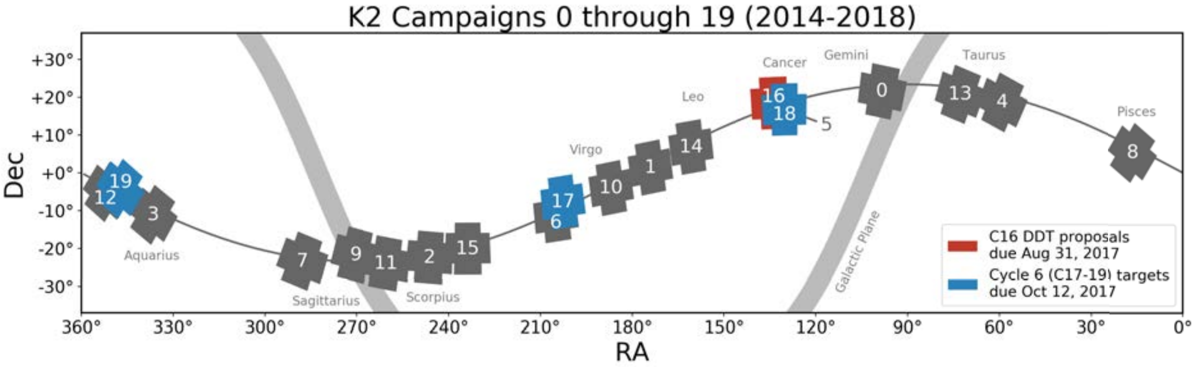


Figure 2.10: The fields of view of the K2 observing campaigns. (https://keplerscience.arc.nasa.gov/k2-fields.html)

The results presented in Chapter 5 are for stars observed during Campaign 1.

2.2.3 CoRoT

The CoRoT mission (Baglin et al., 2006) was a French-led mission designed to search for exoplanets and perform asteroseismology, and was the first mission dedicated to transit detection. It was launched in 2006 and was operational until it suffered a computer failure in 2012. It had field of view of 2.7 by 3.05 degrees and consisted of a 0.27m defocused telescope and an array of four CCD detectors.

CoRoT orbited in a low-Earth polar orbit, and to avoid the Sun entering the field of view of the telescope, it had two main different fields of view, (see Figure 2.11), each observed for 150 days at a time, and between these observing runs, shorter observations were made and allowed for a large number of stars to be observed for asteroseismology.

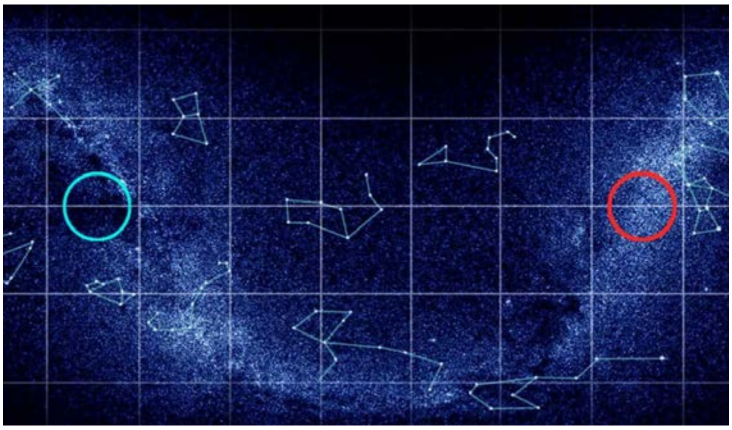


Figure 2.11: The two main fields observed by CoRoT. The region circled in blue is towards the galactic centre, and that shown in red is the anticentre. (http://sci.esa.int/corot/)

The data used in Chapter 5 are from the LRc01 observation, which was performed in 2007 and is a 150 day long run, observing in the direction of the galactic centre.

CoRoT, like *Kepler* has two cadences, and here we use the long-cadence. In the work done in Chapter 5, the search for the seismic parameters is confined below 100μHz. The reason for this is the South Atlantic Anomaly. This is a region above South America where the Van Allen radiation belt is closest to the Earth, and therefore due to the orbit of CoRoT, the spacecraft is exposed to high levels of radiation every time it passes through this region. Though the CCDs are shielded, the higher energy particles still impact the detectors. In practice, this appears as a very high peak in the power spectrum at around 110μHz, and this can bias the detection tests described in Chapter 5.

TRILEGAL stage flag.	**Stellar evolutionary stage**
0	Pre-Main-Sequence
1	Main-Sequence
2	Subgiant
3	Red Giant Branch
4	Core Helium Burning- Low Mass RC and Horizontal Branch
5 & 6	Core Helium Burning- Higher Mass RC
7	Early Asymptotic Giant Branch
8	Thermally Pulsing Asymptotic Giant Branch
9	Post Asymptotic Giant Branch
10	White Dwarf

Table 2.1: The stage flag variable output by TRILEGAL, and the corresponding evolutionary state.

2.2.4 TRILEGAL

The TRILEGAL (TRI-dimensionaL modELs of the GALaxy) code (Girardi et al., 2005) simulates the photometry and intrinsic properties of a field of stars in the Galaxy, simulating a stellar population containing stars at all stages of evolution from pre-main-sequence to White Dwarf stars, and we will use it extensively in this thesis. It also simulates non-interacting binary stars, and the desired fraction of binary stars may be varied by the user. The initial mass ratio of the binaries lies in a flat distribution between 0.7 and 1, by default. This code is described in full by Girardi et al. (2005). A useful feature for our work is the binary probability (and hence the fraction in the population) is variable, allowing for different binary fractions to be simulated and, for our purposes, tested against real data with an unknown binary fraction. The stellar properties of interest are those which the scaling relations described in Chapter 3 rely on, e.g. the stellar luminosity, mass and temperature, along with the magnitude in the *Kepler* bandpass, the metallicity and the TRILEGAL stage flag, which denotes the evolutionary stage of a given simulated star as shown in Table 2.1.

Stage 3 stars are the hydrogen shell burning red giant branch stars and stages 4-6 are core helium burning stars, commonly referred to as the red clump. Stage 4 core helium burning stars are the low mass stars which undergo the helium flash. Stages 5 and 6 are the higher mass red clump stars.

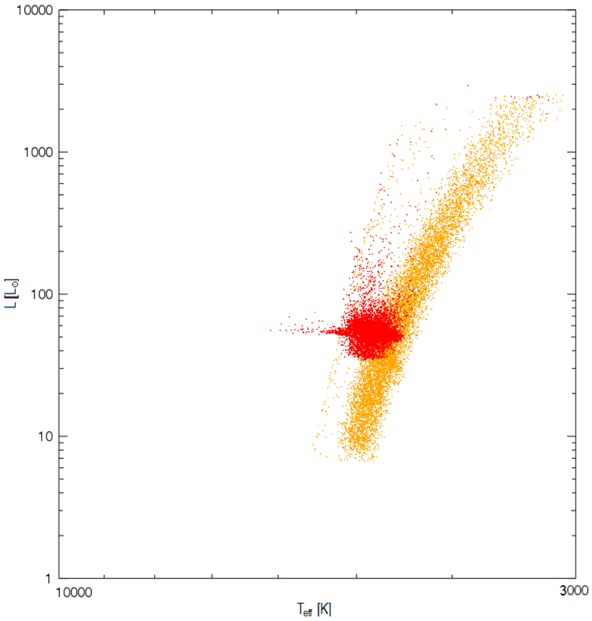


Figure 2.12: A basic Hertzsprung-Russell diagram of the synthetic stellar population considered in this work. Red Giant branch stars are shown in orange, and Core Helium Burning stars are shown in red. Stars which would lie in the classical instability strip are removed.

As the focus of this work is on red giant branch (RGB) and red clump (RC) stars, it is only these which have the variance metric simulated. In the case of binaries, RGB and RC stars which are the primary component are considered, along with their companions of any evolutionary state. Almost all the stars of these evolutionary stages are the primary component in their binary system. Of those which are the secondary component, almost all of them have an RGB or RC primary, and so are not missed. Those very few RGB/RC secondaries whose primary companions are further evolved stars are discounted from this analysis, however the number of these is so small relative to the whole cohort of stars that we choose to exclude them, as the scaling relations are not applicable to these more evolved stars since they are not thought to support solar-like oscillations.

All the red giant and red clump stars considered are shown in Figure 2.12. When applying the scaling relations, they were tested when treating the stars as all being single stars, as well as including the binarity. In the simulated data, stars are flagged according to whether they are alone, the primary (more luminous) component in a binary or the less luminous secondary component. The effect of the binarity of a star is accounted for by applying so called washout factors to the synthetic power spectrum, as described in Chapter 3.

For the simulated stars, a frequency spectrum is constructed using the ouputs from TRILEGAL and scaling relations for the expected frequency region of the modes, their amplitude and the profile of the granulation background. This forms a major part of the work undertaken, and is explained fully in Chapter 3.

2.3 Binary Star Detection

Prospects for using asteroseismology to detect binary stars are discussed at length by Miglio et al. (2014). The method described therein relies on detecting two regions of solar like oscillations (see Section 2.2.1) in the same power spectrum, implying that there are in fact two stars present in the single lightcurve observed by *Kepler* and the authors expect that around 200 binaries should be detectable in this way in the long-cadence data, most of these systems comprising two red clump stars.

The method presented in Chapter 4 uses the detection of the modes of at least one star, and considers the total power present in the region of frequency which the modes occupy. Because the changes in light intensity caused by the oscillations are measured by *Kepler* in parts-per-million, and hence relative to the total light detected, it is clear that the fact that the light from an otherwise undetected companion star would "dilute" the changes in intensity caused by the modes of the main target star with detectable oscillations. This translates into the frequency domain as lower power spectral density than expected for the frequencies of the modes as shown by the star. Following scaling relations, it is possible to make good predictions for what the power spectral density ought to be for a given star.

All the work in this thesis relies on a metric for the filtered variance of the star, defined as the total mean-squared power in the frequency range occupied by the modes. This is preferable to taking a value for the total power across the entire frequency range which the long cadence observations cover to avoid considering power in the spectrum at very low frequency, such as that from stellar activity, and also allows for good simulated power spectra to be created for the frequency range required.

Now we have taken a brief tour of stellar evolution and basic asteroseismology concepts, the remaining chapters concern the results of a series of investigations into red giant stars. First, I present the results and process of creating realistic simulated power spectra which underpins much of the analysis undertaken in the rest of the thesis. Then, a study aiming to infer the binary fraction of *Kepler* red giant stars using a simple metric, an investigation into the differences between red clump and red giant branch stars, and finally, the development of an analysis code, designed to automate the detection of oscillations and parameter estimation for the *K2* and *CoRoT* missions.

Work undertaken during the preparation of this thesis has been included in several publications, including Lund et al. (2015) and Miglio et al. (2016). In particular, work in Chapter 5 has contributed to the K2 Galactic Archeology Project by Stello et al. (2017), and the CoRoT ExoGiants project by Montalban et al. (in prep). Chapter 4 will form Jones et al. (in prep).

Chapter 3

Power Spectrum and Variance Simulation and Interpretation

3.1 Introduction

A basic property of the the power spectrum is that the variance in the time domain is equal to the total integrated power across all frequencies in the frequency domain. The variance arising from intrinsic stellar variability due to oscillations and granulation is known to scale with the frequency of maximum oscillation power ν_{max} (eg. Hekker et al. 2012) and it is this relationship that forms the basis of much of the work contained within this thesis. In order to use this relationship to make inferences about stellar populations, realistic simulated spectra are required for a synthetic population of stars with known properties, as described in Chapter 2.

The metric used throughout this work is a bandpass filtered estimate of the variance, integrating the region in frequency containing the asteroseismic modes of oscillation, and removing photon noise. The reasons for this are twofold. In taking this filtered region, the physical signal from the modes and the stellar granulation is preserved while any contribution from the stellar activity signal and high amplitude peaks from the spacecraft pointing is not considered, as these are confined to very low frequencies. Also, the granulation, modes and photon noise can be described well by existing relations in terms of the stellar properties, allowing predictions of the filtered variance for a synthetic population of stars to be as realistic as is needed. A cartoon example of the frequency region integrated to produce this metric is shown in Figure 3.1.

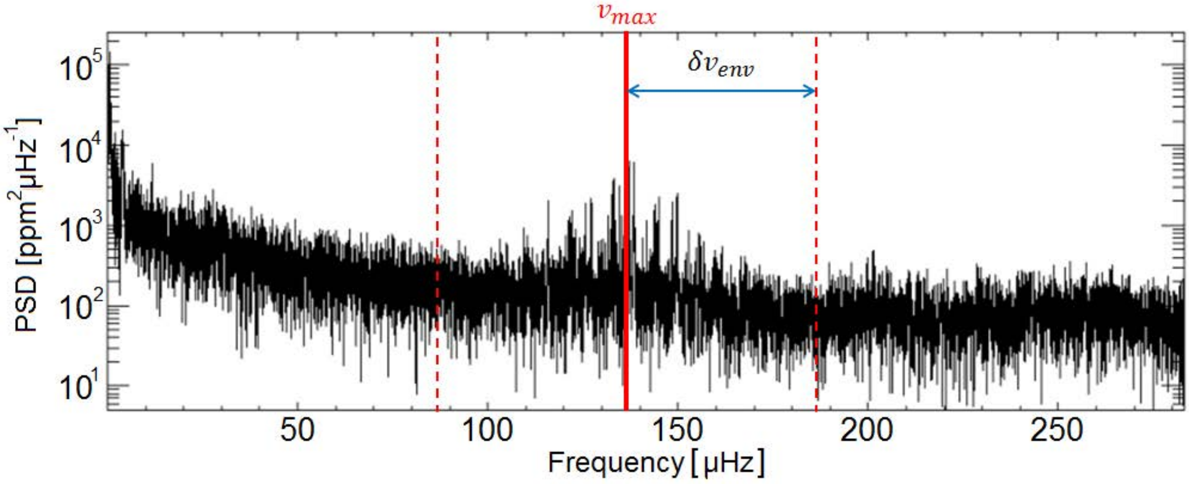


Figure 3.1: The frequency region integrated to produce the variance metric for a *Kepler* red giant. The metric encompasses all power that lies within $\nu_{\rm max} \pm \delta\nu_{\rm env}$, where $\delta\nu_{\rm env}$ is the FWHM of the Gaussian envelope modulating the mode power, discussed later in this chapter.

This metric also has the advantage that in real stars, we do not need to rely on model dependent fits for the granulation background in order to study the mode region. It is simply the total power present in the spectrum due to the modes ($P_{\rm m}$) and granulation ($P_{\rm g}$), over the frequency range $N_{\rm bins}$ wide containing the modes. The metric can be described as

$$\text{Variance} = N_{\rm bins} \sum_{\nu_{\rm max}-\delta\nu_{\rm env}}^{\nu_{\rm max}+\delta\nu_{\rm env}} \left(P_{\rm m}(\nu) + P_{\rm g}(\nu)\right) \tag{3.1}$$

This leaves us with our variance metric, a bandpass filtered estimate of the total mean-squared power in the frequency range occupied by the modes. The variance-$\nu_{\rm max}$ relationship for real stars is shown in Figure 3.2.

Here, I will first explain how the artificial spectrum is constructed. This chapter focuses on the choices of descriptions for all the constituent parts of the synthetic spectra, and justifying the level of complexity of the simulations. For the purposes of most of the work contained within this thesis, red giant branch stars only are simulated. A discussion of the red clump stars follows in Chapter 6.

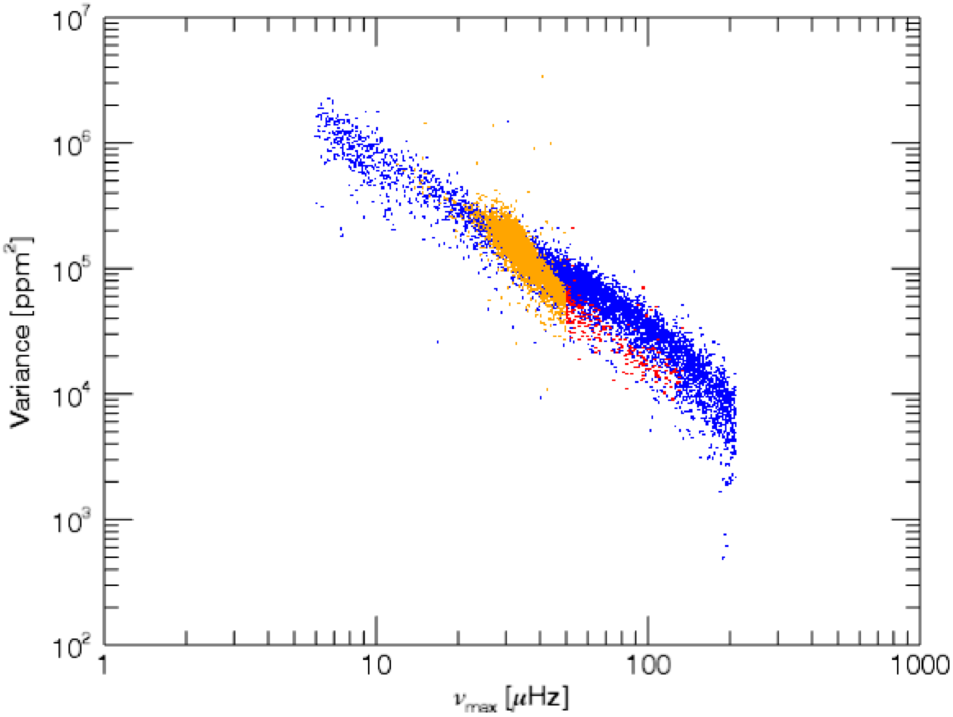


Figure 3.2: The variance metric - $\nu_{\rm max}$ relationship for the *Kepler* sample used in this work. Blue points denote red giant branch stars, while orange and red denote the red clump and secondary red clump respectively.

3.2 Asteroseismic mode power: Scaling relations

In creating the simulated data, asteroseismic scaling relations are used along with fundamental stellar properties from the TRILEGAL code to simulate the basic asteroseismic parameters $\nu_{\rm max}$, $\Delta\nu$ and $A_{\rm max}$.

The scaling relations used in this work describe three of the seismic parameters of stars; the frequency of maximum oscillation power $\nu_{\rm max}$, the large frequency separation $\Delta\nu$ and $\mathrm{A}_{\rm max}$, the amplitude of an $\ell = 0$ mode at $\nu_{\rm max}$ (Verner et al. (2011)). While the relations for $\nu_{\rm max}$ and $\Delta\nu$ are based on common assumptions and therefore largely agreed upon, there are many relations for $\mathrm{A}_{\rm max}$ with differing derivations and assumptions involved. For this work, it is necessary to test some existing $\mathrm{A}_{\mathrm{max}}$ scaling relations in order to determine which one is the most suitable to apply to the artificial data generated by the TRILEGAL code so it can be compared in a meaningful way to the real *Kepler* data. Solar values used are $T_{\mathrm{eff},\odot} = 5777\mathrm{K}$, $\nu_{\mathrm{max},\odot} = 3090\ \mu\mathrm{Hz}$ and $\Delta\nu_{\odot} = 135\mu\mathrm{Hz}$ (Huber et al., 2010). $A_{\mathrm{max},\odot}$ is taken as 3.6 ppm from Michel et al. (2009) unless otherwise stated.

The following are the eight scaling relations for A_{max} that have been tested in this work, along with a brief explanation of their justification. Depending on the method and the data used to fix the exponents in the relations, there can be need to apply a bolometric correction of the form

$$A_{\mathrm{max,\ in\ }\mathit{Kepler}\mathrm{\ bandpass}} = \frac{A_{\mathrm{max,\ bolometric}}}{c_k} \tag{3.2}$$

where

$$c_k = \left(\frac{T_{\mathrm{eff}}}{5934K}\right)^{0.8}. \tag{3.3}$$

This allows amplitudes predicted by a given relation to be converted between the value which they would take if being observed across all wavelengths (the bolometric measurement) and that when being observed in the *Kepler* bandpass. This is applicable to the relations derived theoretically, and those which have been obtained by fitting to already corrected data. Applying this correction allows us to directly compare the simulated data with the real data, which is observed in the *Kepler* bandpass.

The Kjeldsen and Bedding (1995) (KB95) relation.

Kjeldsen and Bedding (1995) informs many later studies on scaling relations. The amplitude relation they define is referred to as KB95 from now on. There were no confirmed observations of solar-like oscillations on other stars at this time, and as such it is based on linear adiabatic theory. The central assumption is that the variation in the luminosity due to the oscillation is entirely due to changes in temperature, which are in turn caused by the expansion and contraction of the surface, though the change in radius itself is negligible by comparison. Assuming that the oscillations are adiabatic, the luminosity amplitude $A_{\text{max, bol}}$ scales as

$$\mathrm{A}_{\text{max, bol}} = (\delta \mathrm{L}/\mathrm{L}) \propto (\delta \mathrm{T}/\mathrm{T}) \propto (\delta\rho/\rho) = \mathrm{v}/\mathrm{c}_{\mathrm{s}}, \tag{3.4}$$

where L, T, ρ, v and $\mathrm{c_s}$ denote luminosity, temperature, density, fluid velocity and sound speed. The velocity amplitudes of the oscillations scale as $\mathrm{A}_{\mathrm{vel}} \propto \mathrm{v}$ and estimating $\mathrm{c_s}$ using the ideal gas law leads to

$$A_{\text{max}} \propto \frac{A_{\text{vel}}}{\sqrt{T_{\text{eff}}}}. \tag{3.5}$$

Further calibration, using the very few stars with detectable oscillations that were available at the time, recovers the velocity amplitude in terms of mass and luminosity which scales as $A_{\text{vel}} \propto L/M$ and also adjusts the temperature exponent, leading to the relation referred to in Figures 3.3 and 3.4 as KB95;

$$A_{\text{max}} = \frac{17.7\text{ ppm } A_{vel}[\text{ms}^{-1}]}{c_k(T_{\text{eff}}/T_{\text{eff},\odot})} = \frac{17.7\text{ ppm } \times 0.234(L/L_\odot)}{c_k(T_{\text{eff}}/T_{\text{eff},\odot})(M/M_\odot)}, \tag{3.6}$$

to which the bolometric correction is applied to convert into the *Kepler* bandpass.

The Mosser et al. (2010) (M10) relation.

The next relation is adapted from Mosser et al. (2010) and is derived from *CoRoT* data. *CoRoT* (Baglin et al., 2006) was a French-led satellite and was a predecessor to *Kepler*. It was the first satellite dedicated to asteroseismology and exoplanet detection, and observed in a different bandpass to *Kepler*, and so has a different response when observing the variations in intensity. The relation

$$A_{\text{max}} = \frac{A_{\text{max}\odot}}{c_k}\left(\frac{L/L_\odot}{M/M_\odot}\right)^{0.89}\left(\frac{T_{\text{eff}}}{T_{\text{eff},\odot}}\right)^{-\frac{1}{2}}, \tag{3.7}$$

referred to as M10 throughout this chapter, comes from fitting the amplitudes observed on over 4600 *CoRoT* light curves to their L/M ratio and fixing the exponent of the power law relationship between them. This was included to test whether taking an amplitude derived from *CoRoT* and converting it to the *Kepler* bandpass gives consistent results.

The first Huber et al. (2010) (H11-1) relation.

Huber et al. (2010) tested scaling relations by fitting models to around 1700 *Kepler* stars ranging from the main-sequence, through red giant branch stars to the red clump. They assume the form of the relation is the same as proposed by Kjeldsen and Bedding (1995) and fit the relationship between amplitude and ν_{max} to fix the L/M exponent and find it to take the value 0.8, leading to what we refer to here as the H11-1 relation

$$A_{\mathrm{max}} = A_{\mathrm{max}\odot} \left(\frac{L/L_{\odot}}{M/M_{\odot}} \right)^{0.8} \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}} \right)^{-1.0} . \tag{3.8}$$

The second Huber et al. (2010) (H11-2) relation, with additional mass dependency.

However, in agreement with older studies, they also considered whether there may be an additional mass dependency which is not necessarily accounted for in using a single exponent for L/M. They fit the same data while accounting for the mass distribution using masses and radii determined from the observed $\Delta\nu$ and ν_{max}. This leads to H11-2

$$A_{\mathrm{max}} = A_{\mathrm{max}\odot} \frac{(L/L_{\odot})^{0.84}}{(M/M_{\odot})^{1.32}} \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}} \right)^{-1.0} \tag{3.9}$$

being the best fit, where they find model amplitudes matching the observations far more closely, accounting for the spread in amplitudes in red giants while obeying the mass-independent relationship that is present in main-sequence stars.

The Chaplin et al. (2011) (C11) relation.

The predictions made by Chaplin et al. (2011) are also based on those made by Kjeldsen and Bedding (1995). As the work is directly considering the oscillation amplitudes in the

Kepler bandpass the authors follow Kjeldsen and Bedding (1995) who, when correcting for observations at a particular bandpass as opposed to bolometric, assume a quadratic dependence on temperature leading to

$$A_{\mathrm{max}} = A_{\mathrm{max}\odot}\beta\left(\frac{L/L_{\odot}}{M/M_{\odot}}\right)\left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}}\right)^{-2.0} \tag{3.10}$$

where

$$\beta = 1 - \exp\left(\frac{-(\mathrm{T}_{\mathrm{red}} - \mathrm{T}_{\mathrm{eff}})}{1550}\right) \tag{3.11}$$

labelled as C11 here. The β correction accounts for an overestimate of the oscillation amplitude of hot solar type stars. Models are compared to observed values, and the exponent of L/M is fixed at 1.0.

The Kjeldsen and Bedding (2011) (KB11) relation.

Kjeldsen and Bedding (2011) revised their earlier work, taking into account the excitation and damping mechanism of the modes. They consider that the amplitude in velocity has an additional dependency on the mode lifetime and arrive at

$$A_{\mathrm{vel}} \propto \frac{L\tau^{0.5}}{M^{1.5}T_{\mathrm{eff}}^{2.25}} \tag{3.12}$$

as an expression for $\mathrm{A}_{\mathrm{vel}}$. Here τ here refers to the lifetime of the modes. The scaling of this quantity with the fundamental stellar parameters is uncertain, but we adopt

$$\frac{\tau}{\tau_{\odot}} = \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}}\right)^{-4.0} \tag{3.13}$$

from Chaplin et al. (2009). Combining these with the relationship between the velocity and luminosity amplitudes used by Kjeldsen and Bedding (1995) leads us to

$$A_{\mathrm{max}} \propto \frac{A_{vel}}{\sqrt{T_{\mathrm{eff}}}} \implies A_{\mathrm{max}} = \frac{A_{\mathrm{max}\odot}}{c_k}\frac{(L/L_{\odot})}{(M/M_{\odot})^{1.5}}\left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}}\right)^{-4.75} \tag{3.14}$$

as a description for $\mathrm{A}_{\mathrm{max}}$, which we refer to as KB11.

The Stello et al. (2011) (S11) relation.

The relation derived by Stello et al. (2011) is the result of fitting model amplitudes to observations with the assumption that a single exponent for L/M cannot describe the observed amplitudes. This leads to the expression

$$A_{\text{max}} = \frac{A_{\text{max}\odot}}{c_k} \frac{(L/L_\odot)^{0.9}}{(M/M_\odot)^{1.7}} \left(\frac{T_{\text{eff}}}{T_{\text{eff},\odot}}\right)^{-1.0}, \tag{3.15}$$

which the authors state works well for main sequence and subgiant stars, as well as the red giant branch.

Deriving a relation (MCon11)from Mosser et al. (2012a).

It is important to appreciate the distinction between amplitude and height. Height in power spectral density is proportional to the power of that frequency component, which is equivalent to the amplitude squared of the time domain signal in the light curve, which corresponds to that frequency. The final relation studied is from Mosser et al. (2012a), and is slightly different from the previous seven in that it is constructed from a power law relationship such that $H_{\text{env}} = \alpha \nu_{\text{max}}^{\beta}$. H_{env} is the height of a mean Gaussian centred on ν_{max} which describes the power excess of of the oscillation power above the granulation background. It relates to A_{max} as:

$$H_{\text{env}} = \frac{3.1\, A_{\text{max}}^2}{\Delta\nu} = 2.03 \times 10^7 \nu_{\text{max}}^{-2.38}. \tag{3.16}$$

The factor 3.1 represents the total mode power present in the central $\Delta\nu$ of the mode envelope, in units of the power of the central $\ell = 0$ mode ($\text{A}_{\text{max}}{}^2$), accounting for the difference in visibility of the modes of different degree ℓ. This intuitively shows that H_{env} represents the maximum power spectral density of the mode envelope.

A power law relation for $\Delta\nu$

$$\Delta\nu = 0.276\, \nu_{\text{max}}^{0.75} \tag{3.17}$$

is introduced in Mosser et al. (2012a) and in substituting this into equation 3.16, we arrive at a power law description of A$_{\rm max}$

$$A_{\rm max} = 1344.4\,\nu_{\rm max}^{-0.815}. \tag{3.18}$$

though in order to have an A$_{\rm max}$ relation into the same format as the others which can be applied to the synthetic data, recall that $\nu_{\rm max}$ scales as

$$\nu_{\rm max} = \nu_{\rm max,\odot} \left(\frac{M/M_\odot}{L/L_\odot}\right) \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{3.5} \tag{3.19}$$

and substituting this in, the testable scaling relation referred to as MCon11 is

$$A_{\rm max} = 1.89 \left(\frac{L/L_\odot}{M/M_\odot}\right)^{0.815} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{-2.85}. \tag{3.20}$$

3.2.1 Comparing scaling relations for $A_{\rm max}$

In order to decide which of these relations to use, I make a qualitative comparison with amplitude values for 1248 *Kepler* red giants for each of the eight relations.

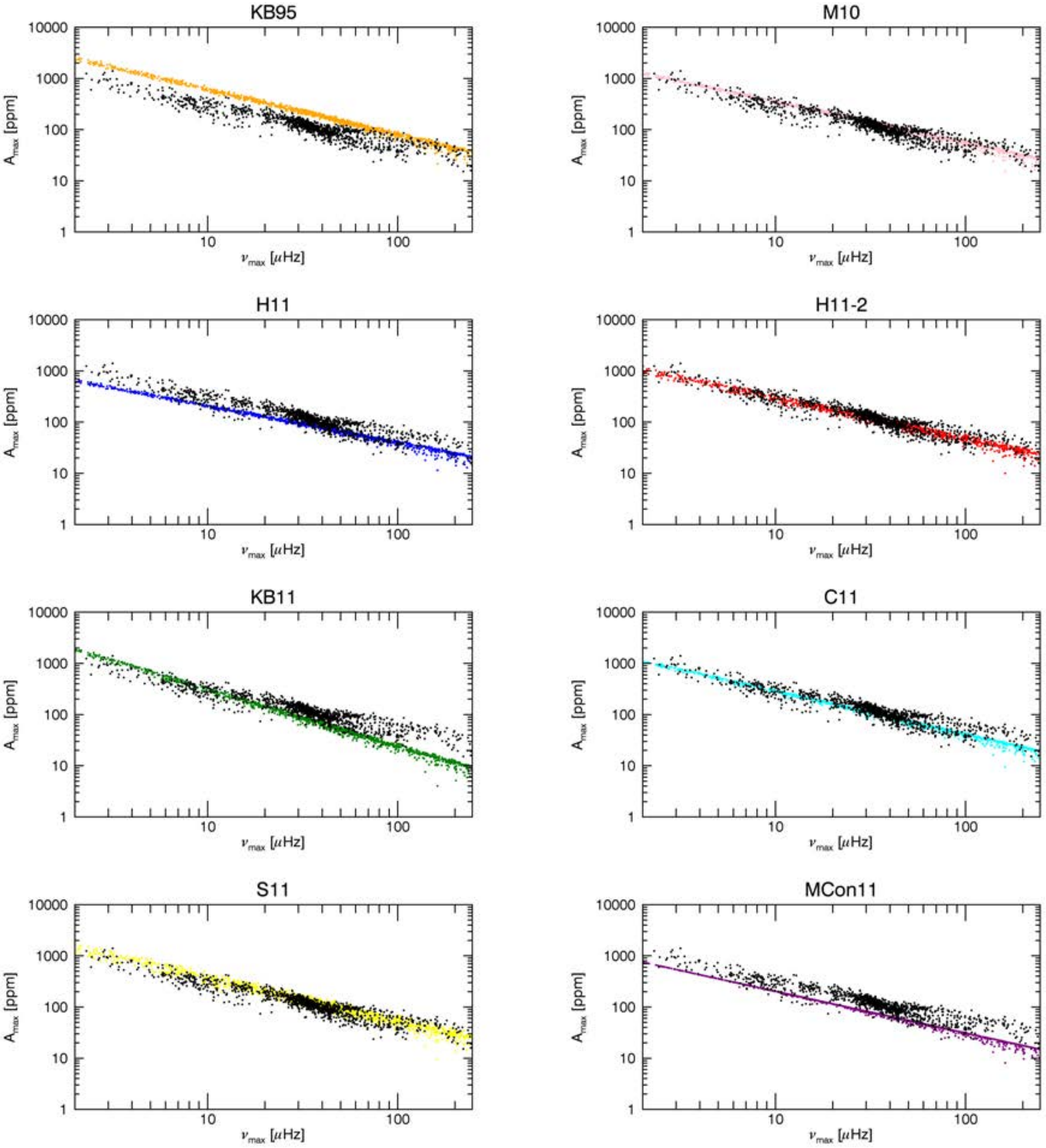


Figure 3.3: The $\nu_{\max}$- A$_{\max}$ relationships for each of the above eight relations. The grey points are the values for 1248 *Kepler* stars and the coloured points are the simulated amplitudes given by the eight relations as labelled.

Unsurpisingly, most of the relations compare reasonably well to the measured amplitudes since they are based on fits to observed amplitudes. When fitting a power law to these relations, the only simulated data whose fit parameters agree with those of the real data within a 95% confidence interval is the H11-2 data, shown above in red, however a small but appreciable offset in both intercept and slope remains. The much greater scatter present in the real data is expected to be due to uncertainties in measurement, however, as suggested in Huber et al. (2011), there is also a contribution from the mass distribution of the stars, and as such an additional mass dependency in the relations is favourable, and the H11-2 from this paper (Equation 3.9) and the S11 relation from Stello et al. (2011) (Equation 3.15) both contain such an exponent in their fits.

As another constraint, the signal-to-noise ratio can also be calculated. Chaplin et al. (2011) suggest that the value of the granulation background profile measured at $\nu_{\rm max}$ adequately describes the average background profile in the mode region to first order, allowing an estimate of the integrated background power across the frequency range containing most of the mode power to be made in a trivial way. In Chaplin et al. (2011), this region is $\nu_{\rm max} \pm \nu_{\rm max}/2$, assuming that the full-width-half-maximum of the Gaussian envelope that modulates the oscillations (as described in detail in section 3.4) is equal to $\nu_{\rm max}/2$ so

$$\int_{0.5\nu_{\rm max}}^{1.5\nu_{\rm max}} P_{\rm gran}(\nu) \approx \nu_{\rm max} P_{\rm gran}(\nu_{\rm max}). \tag{3.21}$$

Following Mosser et al. (2010), the value of the background at $\nu_{\rm max}$ can be represented as

$$P_{\rm gran}(\nu_{\rm max}) = 6.37 \times 10^{6} \nu_{\rm max}^{-2.41}. \tag{3.22}$$

for red giant stars. For the total integrated mode power over the above selected range in frequency, $\mathrm{P_{tot}}$ the Chaplin et al. (2011) form for the approximation is used

$$P_{\rm tot} \approx 1.55 \mathrm{A}_{\rm max}^{2} \frac{\nu_{\rm max}}{\Delta\nu} \tag{3.23}$$

and the photon noise $\mathrm{b_{instr}}$ as discussed in Chapter 2 is described by the relation from Jenkins et al. (2010) (see Section 3.7). This leads to a formula for the signal-to-noise ratio of

$$\mathrm{SNR} = \frac{P_{\mathrm{tot}}}{\nu_{\mathrm{max}} \mathrm{b}_{\mathrm{instr}} + P_{\mathrm{gran}}} \tag{3.24}$$

Figure 3.4 shows this quantity for the eight relations.

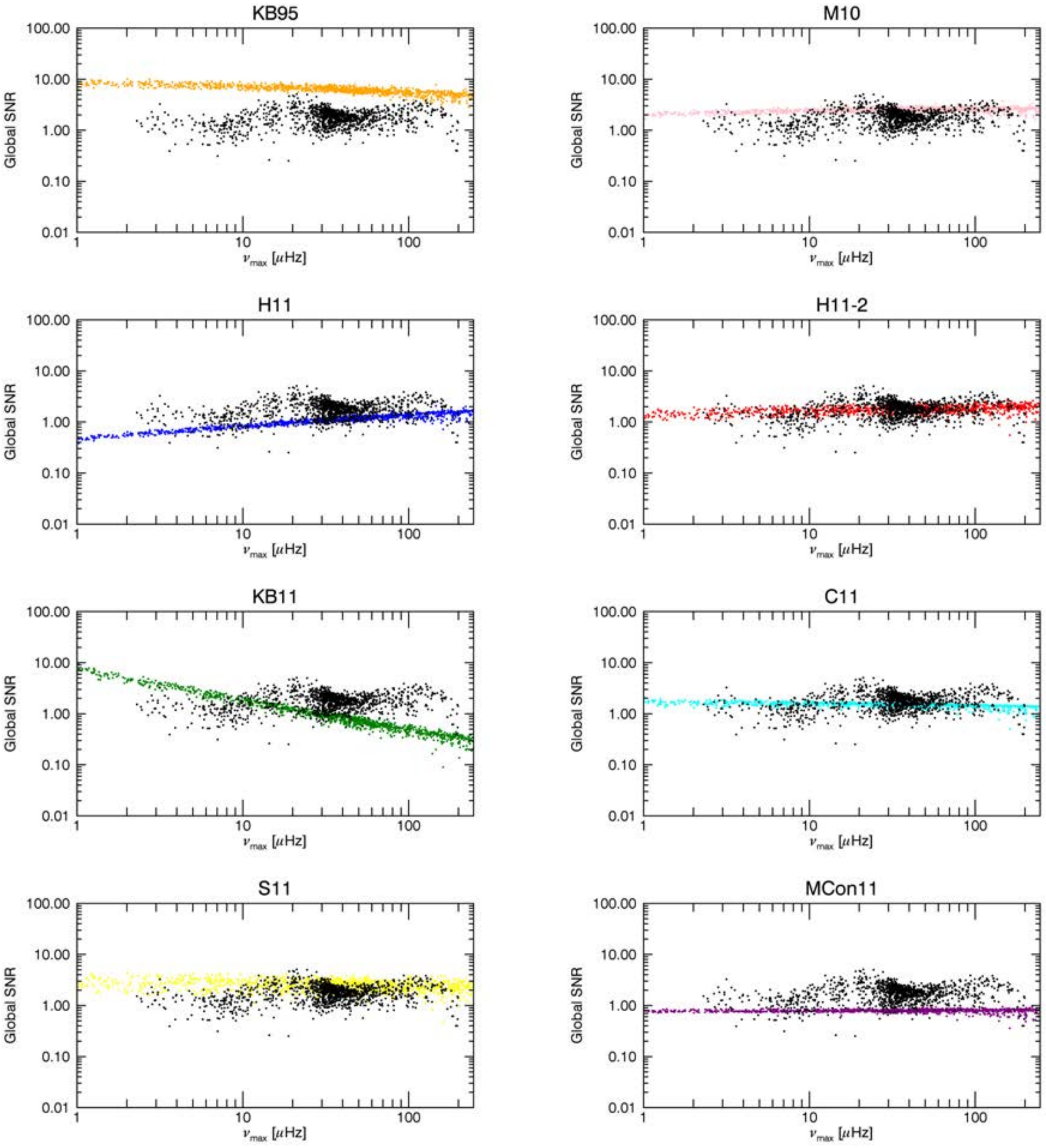


Figure 3.4: Signal-to-noise ratios calculated for each of the 8 amplitude relations according to the above description. Points in black are the long cadence signal-to-noise values from Chaplin et al. (2013)

Due to the simulated signal-to-noise ratios being a very crude approximation, it is not particularly informative to draw any statistical comparison between the signal-to noise ratios of the simulated and real data. The slope of the linear fits to the real and simulated data agree within a 95% confidence interval for the H11-2 only. Again, we expect the real data to have a far greater scatter than the simulations, due to observational uncertainty, but the relations with an additional mass dependency (H11-2 and S11) do go some way towards replicating the scatter we see in the real data for both amplitude and SNR.

The H11-2 relation from Huber et al. (2010) is used for the remaining work, as it performs best for red giant stars. Moreover it is also correct for solar-type stars of a higher ν_{max}, which is a consideration for the work contained within this thesis where we simulate binary stars, some with solar-type companions. Having chosen an amplitude relation, this chapter goes on to describe other components needed to simulate a realistic power spectrum.

3.3 Frequency spacings

As discussed in Chapter 2, the large frequency separation $\Delta\nu$ measures the average spacing between modes of the same degree ℓ in consecutive radial orders, n. In these simulations, only modes of degree up to and including $\ell = 3$ are considered, as degrees higher than this are not readily detectable in *Kepler* data. Having generated a value for ν_{max}, $\ell = 0$ modes are spaced at intervals of $\Delta\nu$ centred on ν_{max}, and $\ell = 1$ modes are anchored halfway between $\ell = 0$ ridges. The frequencies of the $\ell = 2$ and $\ell = 3$ modes are described relative to the position of the $\ell = 0$ and $\ell = 1$ frequencies respectively by

$$\delta\nu_{02} = 0.122\Delta\nu + 0.047 \tag{3.25}$$

$$\delta\nu_{13} = 0.282\Delta\nu - 0.16 \tag{3.26}$$

as determined by Huber et al. (2010) by taking the centroids of the mode ridges in an ensemble collapsed *é*chelle diagram for 470 red giant stars. The precise locations of the oscillation modes make little to no difference to the variance itself, however this level

Degree	V_ℓ
$\ell = 0$	1.0
$\ell = 1$	1.5
$\ell = 2$	0.5
$\ell = 3$	0.1

Table 3.1: Mode visibilites assumed throughout.

of detail is necessary for other uses for the simulated spectrum, including testing the detection pipeline described in Chapter 5. The frequencies of the simulated $\ell = 1$ modes are more complex, as they exhibit a dense spectrum of mixed modes.

3.4 Mode Profiles

We know that $\mathrm{A}_{\max}$ denotes the amplitude of an $\ell = 0$ mode at $\nu_{\max}$. Modes at different frequencies have their height modulated by the Gaussian envelope centred on this frequency, and modes of oscillation of different degree ℓ also have different have different relative visibilities due to geometrical cancellation as well as intrinsic properties of their damping and excitation. Therefore, modes of different degree ℓ contribute different fractions of power to the oscillations, and there are assumptions that can be made about these contributions. Following Chaplin et al. (2011) and Ballot et al. (2011), the relative visibilites in power of the modes used are given in Table 3.1.

There are of course departures from the values in Table 3.1, but in the absence of simple relations for these values in giant stars, these reasonable approximations are chosen. Recall from Chapter 2 that individual oscillations behave in time as a damped harmonic oscillator, and therefore appear as a Lorentzian in frequency. Their amplitude in the time domain is related to their height in the frequency domain by

$$H_{\max} = \frac{V_\ell A(\nu_0)^2}{\pi \Gamma(\nu_0)}, \tag{3.27}$$

where $H_{\max}$ is the maximum power spectral density of the mode at its central frequency

ν_0, with visibility V_ℓ dependent on its degree, and a linewidth $\Gamma(\nu_0)$ which is related to the lifetime of the mode. Also recall from Chapter 2 that the heights of the Lorentzian profiles of the modes are modulated by a Gaussian envelope of the form

$$P_\nu = H\exp\left(\frac{-(\nu-\nu_{\max})^2}{2\sigma^2}\right). \tag{3.28}$$

of height H, centre $\nu_{\max}$ and standard deviation σ. This Gaussian envelope which modulates the heights of the mode profiles is a phenomenological description which reflects the power distribution around the frequency of maximum power $\nu_{\max}$.

An important quantity for the variance metric itself is the full-width-half-maximum (FWHM) of this envelope, as it is this which denotes the limits of the integration, as shown in Figure 3.1. The FWHM $\delta_{\nu\mathrm{env}}$ relates to σ as $2\sqrt{2\ln 2}\sigma$ and the value of

$$\delta\nu_{\mathrm{env}} = 0.66\nu_{\max}^{0.88} \tag{3.29}$$

is adopted throughout this work, following Mosser et al. (2012a). As the power spectral density of the modes is modulated by a Gaussian, hence, the amplitude $A(\nu_0)$ of an individual mode is calculated by interpolating the value of a square-root-Gaussian (also a Gaussian, of different width) of height $\mathrm{A}_{\max}$ centred on $\nu_{\max}$ at the frequencies given by the relations in Section 3.3. The linewidth is related to the lifetime of the mode. In red giant stars, work by Corsaro et al. (2012) provides a scaling relation for the linewidth of radial modes related to the effective temperature of the star

$$\Gamma = \exp\left(\frac{T_{\mathrm{eff}} - T_{\mathrm{eff},\odot}}{T_0}\right) \tag{3.30}$$

where $T_{\mathrm{eff},\odot} = 5777K$ and $T_0 = 601K$. This is also adopted as the linewidth for the non-radial modes. In modelling the modes as a damped driven oscillator in the time domain, the modes have a Lorentzian profile in the frequency domain, (e.g. Chaplin et al. (2013)), and others, leading to a peak profile of the form

$$P(\nu) = H\left(1 + \left(\frac{\nu - \nu_0}{\Gamma/2}\right)^2\right)^{-1} \tag{3.31}$$

for pure p-modes. For mixed modes, the situation is rather more complex, which we will now explore.

3.5 Mixed Modes

In red giant stars, as discussed in Chapter 2, the frequencies of the p-mode oscillations become low enough that they are in the same range as the g-mode pulsations. This leads to coupling between p- and g-modes for the same degree ℓ , and this effect is most pronounced for the $\ell = 1$ modes. This can be incorporated into the synthetic spectrum to add an additional level of realism to the simulation. The frequencies of the mixed $\ell = 1$ modes in a given order n are a function of ν_{n_p}, the frequency of the pure $\ell = 1$ p-mode and the asymptotic g-mode spacing $\Delta\Pi$, expressed as

$$\nu = \nu_{n_p,\ell=1} + \frac{\Delta\nu}{\pi}\arctan\left[\mathrm{q}\,\tan\left(\frac{\pi}{\Delta\Pi_1\nu} - \epsilon_g\right)\right], \tag{3.32}$$

following Basu and Chaplin (in the press), and solved for ν by looking for changes of sign in

$$\frac{\pi(\nu - \nu_{n_p,\ell=1})}{\Delta\nu} - \arctan\left[\mathrm{q}\,\tan\left(\frac{\pi}{\Delta\Pi_1\nu} - \epsilon_g\right)\right]. \tag{3.33}$$

Here q is a constant dictating the level of coupling between the p- and g-modes, assumed to be q = 0.2 (e.g. Mosser et al. (2012b)), and ϵ_g provides an offset in ν, as does ϵ in the asymptotic relation for pure p-modes. This does not impact on the results of the tests carried out here, and is assumed to be zero, following Mosser et al. (2012b) and Bedding et al. (2011). The asymptotic period spacing for $\ell = 1$ g-modes $\Delta\Pi$ is taken by performing a linear fit to the data from Mosser et al. (2014) as shown in Figure 3.5. $\Delta\Pi$ relates to the $\Delta\nu$ from scaling relations as

$$\Delta\Pi = 54.16 + 2.0\Delta\nu. \tag{3.34}$$

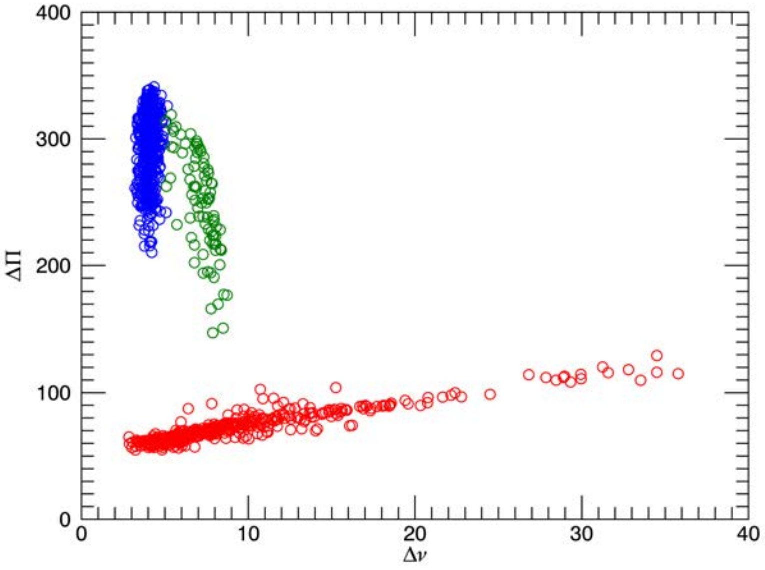


Figure 3.5: $\ell = 1$ Period spacing $\Delta\Pi$ and large frequency separation $\Delta\nu$ data from Mosser et al. (2014). Values for red giant branch stars are shown in red, red clump stars in blue and secondary clump stars in green. The $\Delta\Pi$ values used in the simulations of the RGB stars are taken from a linear fit to the values shown here.

Once the frequencies of the mixed modes have been determined by solving Equation 3.32 for all orders that fall within the range $\nu_{\rm max} \pm \delta_{\nu{\rm env}}$, their heights must be modified according to Q_1, the inertia ratio of the $\ell = 1$ mixed mode relative to that of an $\ell = 0$ mode at the same frequency. The inertia of a mode is a measure of the fraction of the star's mass which is involved in the pulsations of that mode. g-modes, which have their maximum amplitude in the core, where the density is highest, have large mode inertias. So, mixed modes, with their mixed p-mode and g-mode characteristics must have their height predicted according to

$$H_1 = \frac{V_1 T A_0(\nu)^2}{\pi T \Gamma_0(\nu) + 2Q_1} \tag{3.35}$$

following Basu and Chaplin (in the press), rather than the simplified case used for pure p-modes in Equation 3.27. Here T is the observation time (4 years for *Kepler* observations), A_0 is the amplitude a radial p-mode would have were it at the frequency of the mixed mode, obtained by interpolation, Γ_0 is the linewidth of this same hypothetical p-mode, and Q_1 is the inertia ratio. We may estimate the inertia ratio from the mode trapping parameter ζ, which quantifies the ratio between the kinetic energy present in the g-mode cavity and the total kinetic energy of the mode. The expression for ζ

$$\zeta = 1 + \frac{1}{q}\frac{\cos^2\left(\pi\left(\frac{1}{\nu\Delta\Pi_1}\right)\right)\frac{\nu^2\Delta\Pi_1}{\Delta\nu}}{\cos^2\left(\frac{\pi(\nu-\nu_p)}{\Delta\nu}\right)} \tag{3.36}$$

is taken from Deheuvels et al. (2015), and a detailed explanation of the derivation and assumptions made in this expression is present in their paper. There are simpler expressions for this parameter, but it was found that they did not perform well at lower frequencies. Q_1 is related to ζ via;

$$Q_1 = \frac{1}{1-\zeta}. \tag{3.37}$$

This is plotted for a number of simulated stars over a range of frequencies in Figure 3.6, along with the number of modes per order generated by the solution of equation 3.32 and a test that the power is consistent with a pure p-mode simulation of the same stars. Along with the mode heights, the mode widths are also modified by Q_1 according to

$$\Gamma = \Gamma_0 Q_1^{-1}. \tag{3.38}$$

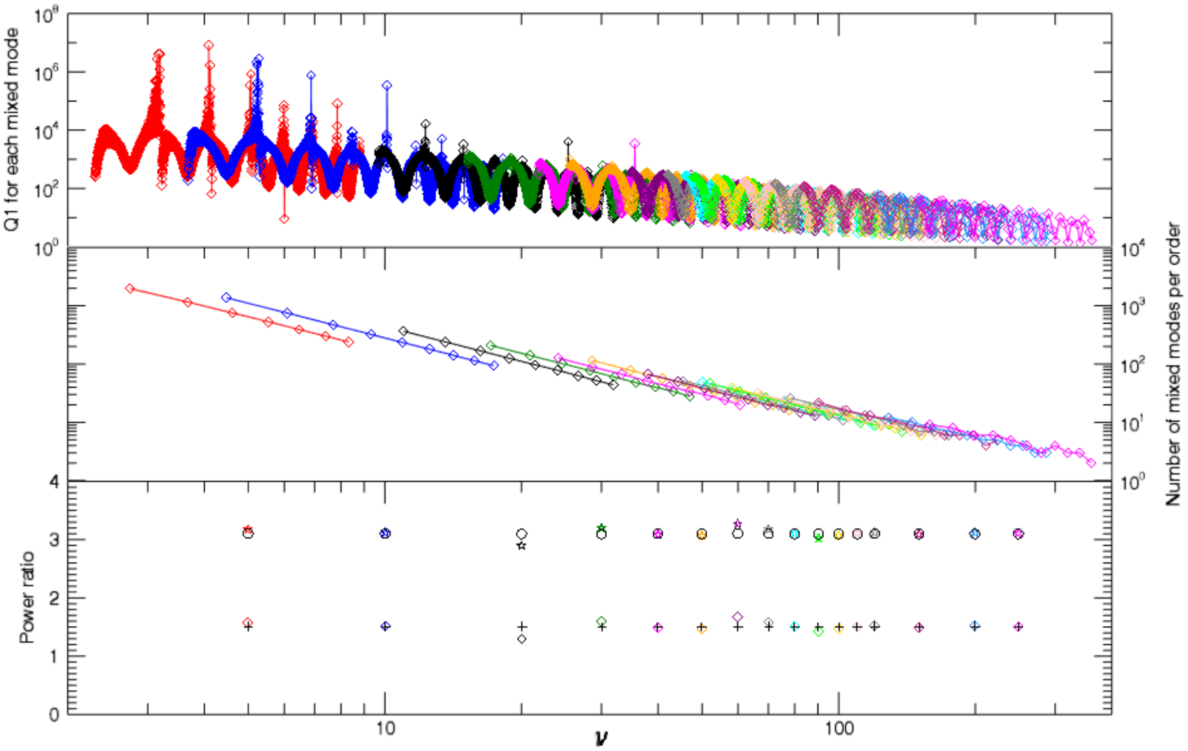


Figure 3.6: Diagnostics for the simulation of $\ell = 1$ mixed modes. Top panel: the inertia ratio Q_1 for each mixed mode simulated. The values for modes of 16 different simulated stars are shown in different colours. Middle panel: The number of mixed modes per order, plotted against nominal $\ell = 1$ frequency. Bottom panel: A sanity check for power levels. Coloured stars show the total power in units of $\ell = 0$ power for the mixed mode simulations, black circles for the pure p-mode case. Coloured diamonds and black crosses follow the same convention for the ratio of $\ell = 1$ power to $\ell = 0$ power.

These tests shown in Figure 3.6 assure us that the behaviour of mixed modes across the full range of frequency is reasonable and consistent. The number of mixed modes per order decreases steadily with frequency as predicted by theory, and the behaviour of Q_1 ensures that the heights and linewidths are modified such that the power present in the modes is consistent with that in the case of pure p-modes. The mode profiles are still described by the Lorentzian form of Equation 3.31, but with the input quantities modified as above. Figure 3.7 shows a typical simulated limit spectrum of the modes. The $\ell = 1$ mixed modes are shown in red.

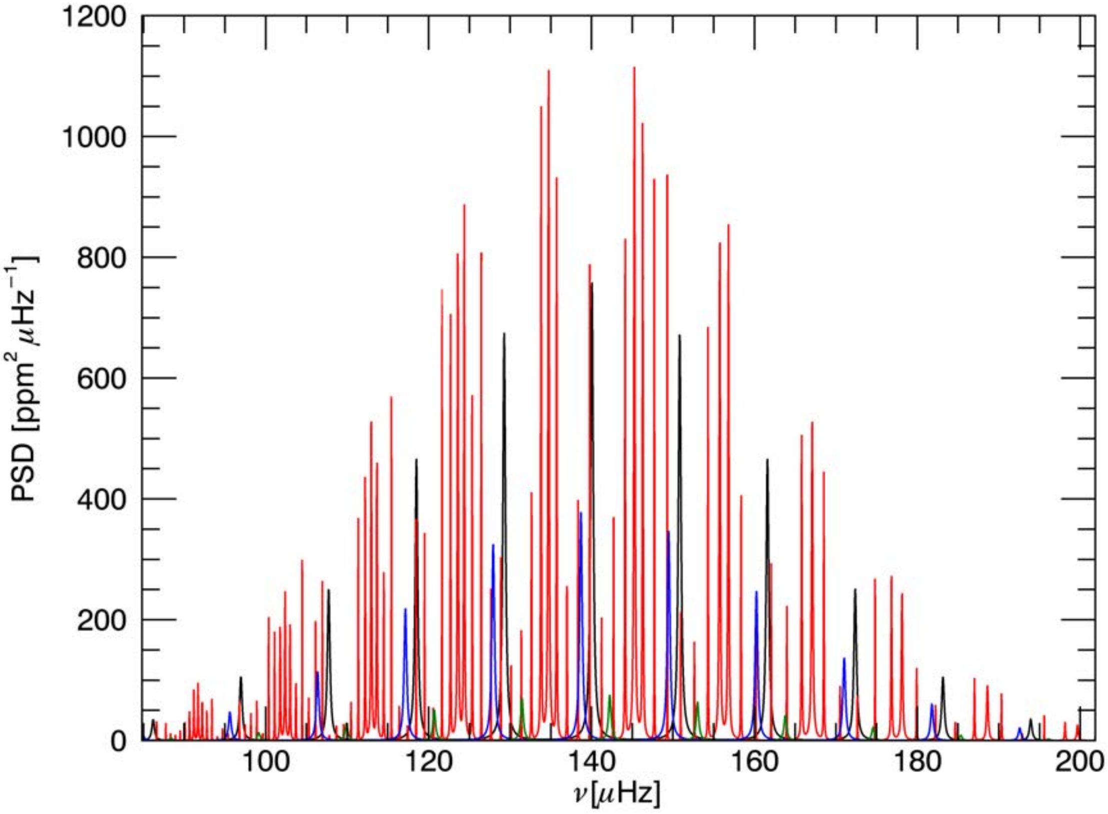


Figure 3.7: A typical simulated mode spectrum. $\ell = 0$ modes are in black, $\ell = 2$ modes in blue, $\ell = 3$ modes in green, and the $\ell = 1$ mixed modes in red.

3.6 Granulation

The granulation signal present in a power spectrum is due to brightness fluctuations associated with the evolution of convective granules at the surface of the convection zone. The granules observed on the stellar surface are formed by the rising and falling plasma. They have a characteristic timescale and brightness fluctuation associated with them which may be shown to scale with $\nu_{\max}$ and hence the evolutionary state of the star.

In the power spectrum, the granulation signature appears as a continuum background signal. It has long been described using a varying number of zero-centred-Lorentzian profiles, eg. by Harvey (1985). Kallinger et al. (2014) studied multiple models for the granulation background using a large set of *Kepler* red giant stars, and following that work, their "model F", a two-component description for the granulation background of the form

$$P_{\text{bg}}(\nu) = \sum \frac{\xi_i a_i^2/b_i}{1+(\nu/b_i)^{c_i}} = \frac{\xi a^2/b_1}{1+(\nu/b_1)^c} + \frac{\xi a^2/b_2}{1+(\nu/b_2)^c} \tag{3.39}$$

where $\mathrm{a} = 3710\ \nu_{\text{max}}^{-0.613}\mathrm{M}^{-0.26}$, $\mathrm{b}_1 = 0.317\ \nu_{\text{max}}^{0.970}$, $\mathrm{b}_2 = 0.948\ \nu_{\text{max}}^{0.992}$, $\mathrm{c} = 4$ and $\xi = 2\sqrt{2}/\pi$ is used in this study. Note the fixing of the exponent $c = 4$, so the two components are in fact "super Lorentzian", since a true Lorentzian profile has an exponent of 2. This is justified by the behaviour of the granules in the time domain. In order for the signature in the frequency domain to be a Lorentzian, the signal in the time domain would have to be an instant brightening phase followed by an exponential decay. In reality, the brightness fluctuation signature in the time domain is arguably better described by an exponential increase phase, then exponential decay. Figure 3.8 shows a toy example of these two time domain signatures and their frequency transforms. Note the more rapid decay of the super-Lorentzian profile shown in red, owing to its $1/\nu^4$ dependency.

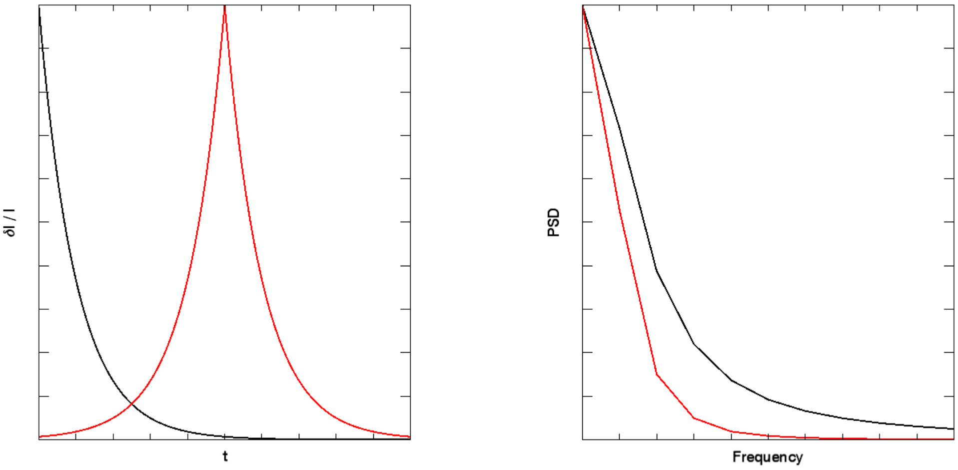


Figure 3.8: A schematic of the behaviour of the granulation signature. The left hand panel shows the time domain signature, black shows exponential decay only, and red shows the more realistic exponential brightening and decay. The right hand panel shows their transformations into the frequency domain.

3.7 Shot Noise

As discussed in Chapter 2, there are various contributions to the noise properties of *Kepler*. The photon noise contributes a frequency independent background term to the power spectrum, and this is a function of the brightness in the *Kepler* bandpass. Jenkins et al. (2010) provide an estimate of this for the long cadence data. It is common to take the mean value of the power spectrum close to the Nyquist frequency as an estimate of the background shot noise. It was found that the value given by the Jenkins et al. (2010) paper is lower than the median-estimated value for the *Kepler* stars used here (see Figure 3.9) so in order to keep the simulations consistent with the observations, an estimation of this noise made by fitting the upper and lower bounds of the noise values is used instead. Figure 3.9 shows fits to the minimum and maximum white noise values at each *Kepler* magnitude for the real stars, then for the simulated stars. The white noise value is estimated by a uniform random value between these two limits. This does not have an impact on the variance itself, as the shot noise background is removed, however as described in the following section, the realisation noise affects the limit spectrum, so it was thought that it suited the purpose of the simulation to increase the shot noise level to make it more realistic.

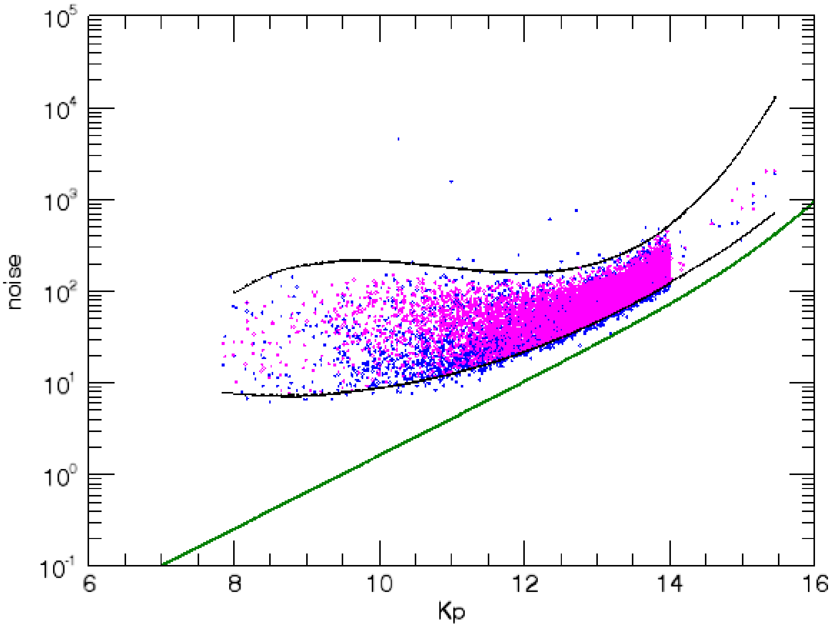


Figure 3.9: The green line shows the minimum term noise model from Jenkins et al. (2010), and the black lines are fits to the upper and lower bounds of the distribution of noise values. In blue is the mean value of the spectrum between 270-280 μHz for the real stars, and in magenta is a simulation of the noise for the simulated spectra, randomly distributed between the upper and lower limits.

3.8 The Limit Spectrum and Realisation Noise

Due to observations being finite in time, random fluctuations are introduced into the power spectrum estimated from the techniques described in Chapter 2. All the components of the power spectrum discussed in this chapter are effectively noise free "limit' values of the various properties, in the sense of what one would expect to see on average over an infinite number of realisations. As described by Anderson et al. (1990), the observed values are distributed according to χ^2 2 degrees of freedom noise about some limit spectrum, which the idealised descriptions here simulate. Following Anderson et al. (1990), the limit spectrum is perturbed using

$$S_{\text{sim}} = -S_{\text{limit}}\ln(\text{y}_\text{i}) \tag{3.40}$$

where S_{sim} is the simulated spectrum, S_{limit} is the limit spectrum, and y_i is a random number uniformly distributed from 0 to 1. Figures 3.10 and 3.11 show the components of the limit spectrum and the final simulated spectrum with this realisation noise added for a mid-frequency red giant.

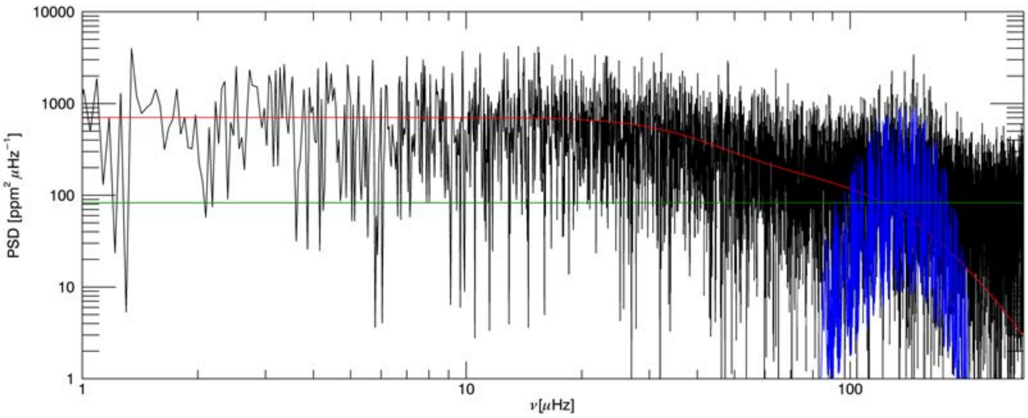


Figure 3.10: The limit spectrum components, shot noise, granulation noise and modes are shown in green, red and blue respectively, and the full spectrum with realisation noise in black.

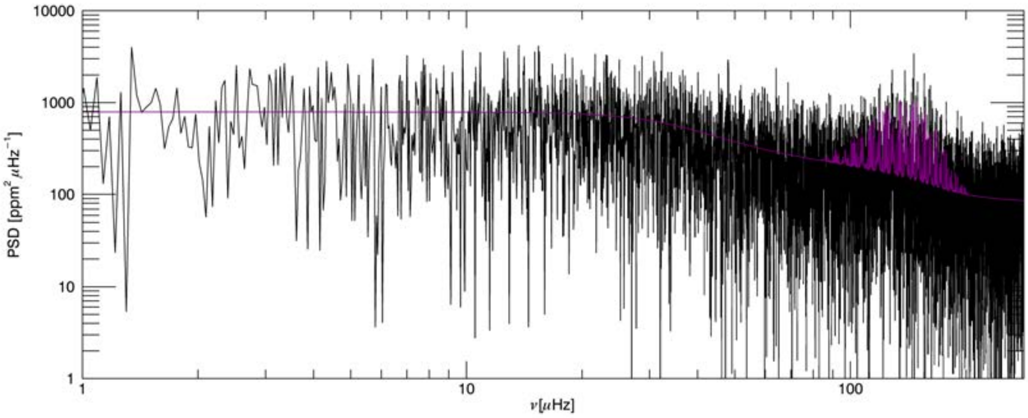


Figure 3.11: The limit spectrum in purple and the single realisation in black.

Additionally, the signal in the power spectrum is affected by apodisation due to the sampling frequency of the observations. This effect is larger the closer the frequency is to the Nyquist frequency, so the whole spectrum except the shot noise, is modified by the factor

$$\eta = \mathrm{sinc}\left(\pi/2\left(\frac{\nu_{\mathrm{max}}}{283\mu\mathrm{Hz}}\right)\right). \tag{3.41}$$

3.9 Accounting for Binarity

The variance metric would seem to be an intrinsic stellar property, dependent on the physical properties of the star, as is clear from the relations used for the granulation background and solar-like oscillations. However, the presence of the light from a second star in the light curve will reduce the observed variance of the star, in a phenomenon we will refer to as "washout." When we deal with the amplitude of the granulation or the oscillation modes, they are measured in parts-per-million, or ppm, a measurement of the fluctuation of the brightness of the star, relative to its mean level. When there is a second star present in the light curve, these fluctuations in brightness due to the oscillations and granulation are reduced, as the mean level is higher. The degree to which the amplitudes are diluted can be described by

$$k_{\mathrm{washout,\ 1}} = \frac{c_{\mathrm{k\ -\ bol,\ 1}} L_1}{c_{\mathrm{k\ -\ bol,\ 1}} L_1 + c_{\mathrm{k\ -\ bol,\ 2}} L_2} \tag{3.42}$$

where $c_{\mathrm{k\ -\ bol}}$ is a factor (Ballot et al., 2011) to convert the bolometric luminosity L to the luminosity in the *Kepler* bandpass. The above formula is therefore is a simple ratio of the luminosity in the *Kepler* bandpass of the oscillating star to that of the whole binary system. Note that in the case of a single star, there is no dilution, L_2 is zero and therefore $k_{\mathrm{washout,\ 1}}$ is unity. In a system where L_2 is a significant fraction of L_1, $k_{\mathrm{washout,\ 1}}$ is small and the amplitudes are significantly diluted, for example in a binary system composed of two red giants. The diluted amplitude is related to the intrinsic amplitude A_{max} by

$$A_{\mathrm{max,\ in\ binary}} = k_{\mathrm{washout,1}} A_{\mathrm{max}}. \tag{3.43}$$

This washout effect of course applies to all the features in the light curve, not only the modes of oscillation. As the simulations described here are only in the power spectrum, the effect of this washout as it affects light curve amplitudes must be squared in power, and so and so we modify the power spectrum components or simply the estimate of the variance by k^2_{washout}. With this accounted for, the effect of a population of binaries on the distribution of the variance-ν_{max} relationship can be tested and studied, and this follows in Chapter 4.

3.10 Results

Applying all of these assumptions and relations, we arrive at realistic power spectra for a population of synthetic red giants. Figure 3.12 shows 6 synthetic spectra for stars across the frequency range observed in the *Kepler* long cadence. It is very noticeable how the bulk properties of the spectra change as stars evolve towards a lower ν_{max}. The power spectral density of the oscillations and granulation background increases, and the envelope containing the oscillations gets narrower.

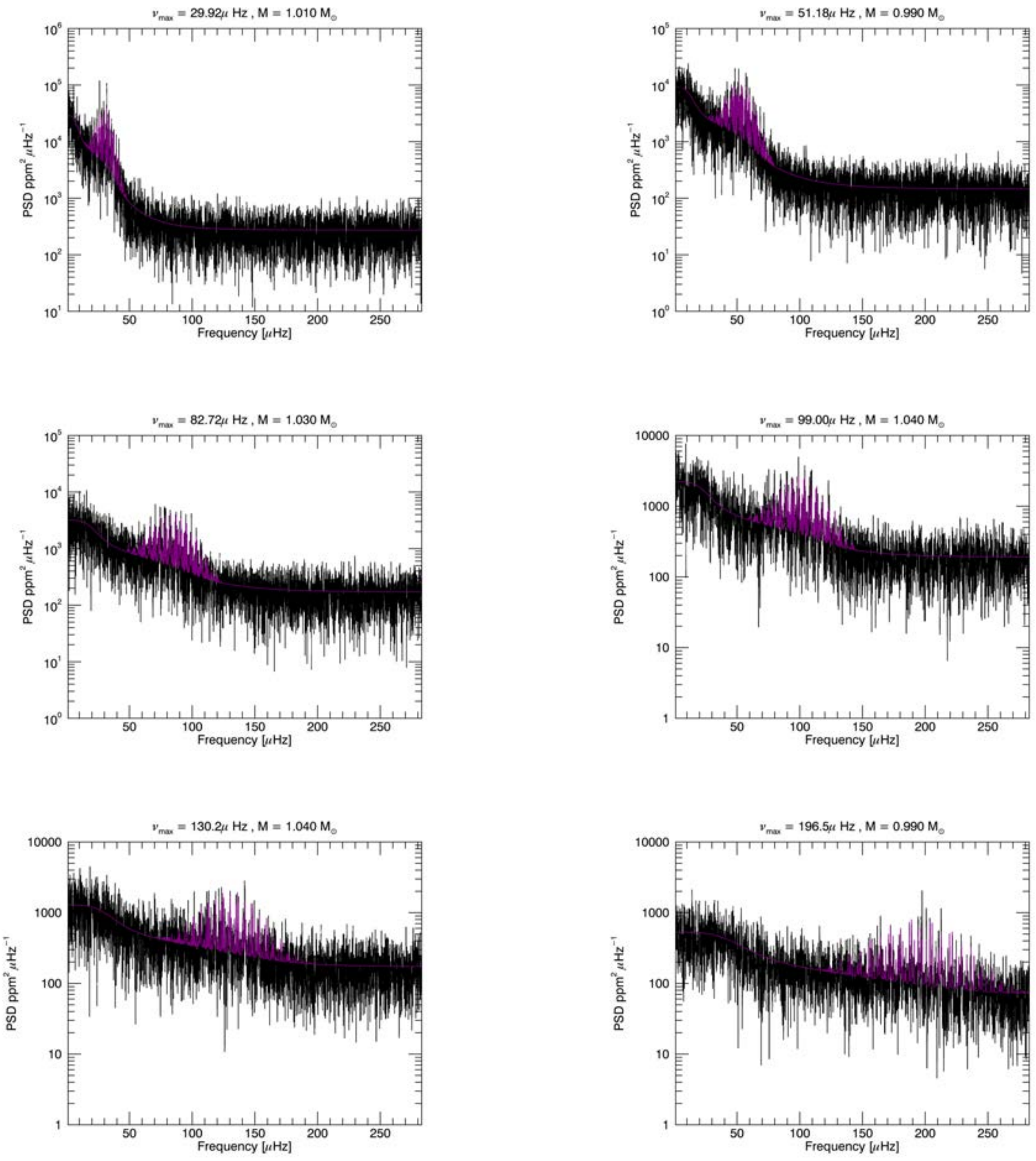


Figure 3.12: 6 simulated spectra of red giants of $\approx 1M_\odot$. The limit spectrum is in purple, and the single noisy realisation is in black.

The complexity of the mode structure can begin to be appreciated in Figure 3.13. The dominant spacing $\Delta\nu$ is clearly visible, as is this quantities evolution with $\nu_{\max}$, similarly decreasing as the star evolves. The evolution of mixed mode behaviour is perhaps a little hard to appreciate in this figure, however the general trend towards more mixed modes at lower frequencies is visible.

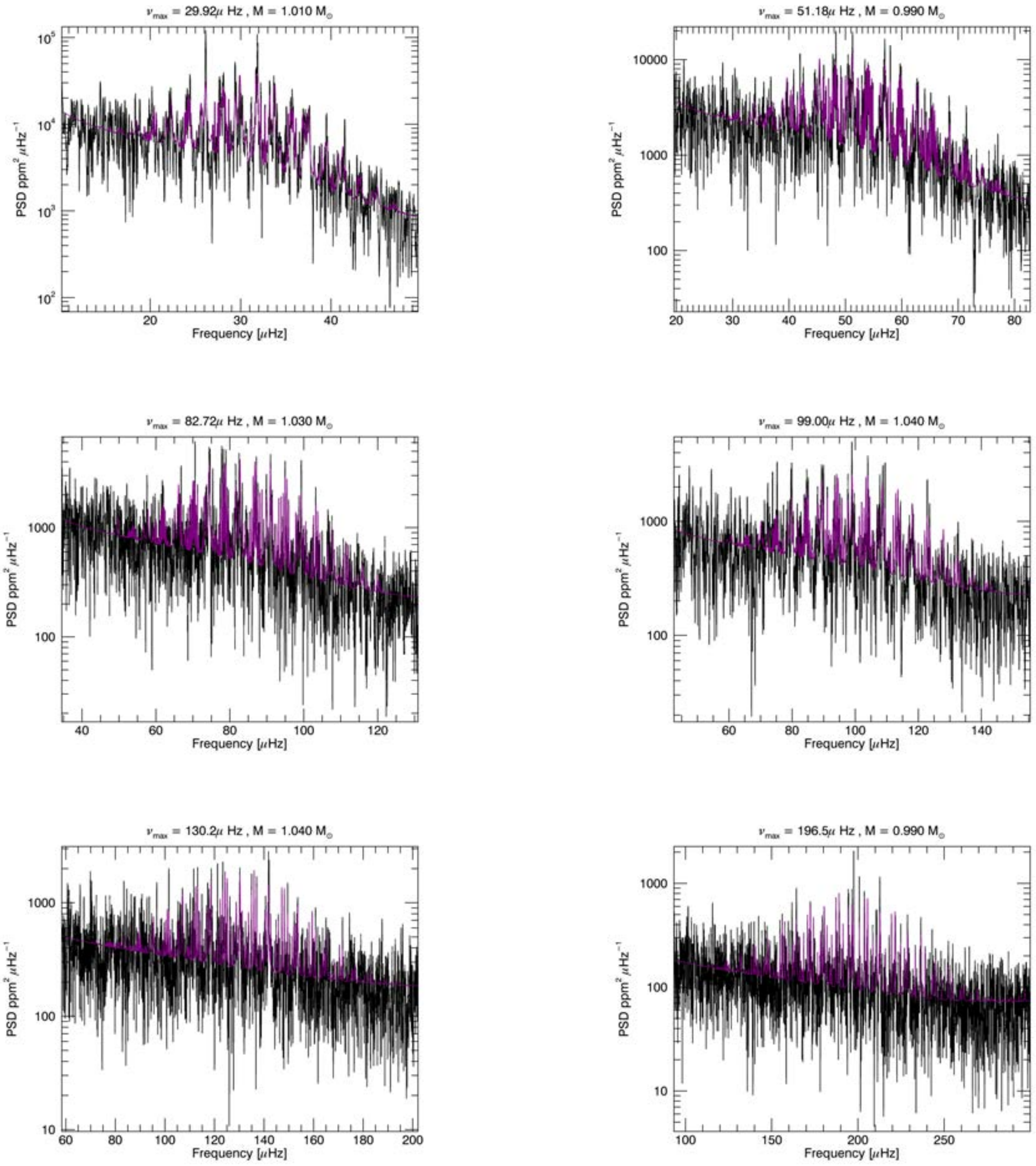


Figure 3.13: A zoomed in view of the modes of the same simulated stars as shown in Figure 3.12.

The mixed mode evolution is evident in an échelle diagram, Figure 3.14 shows the distinctive departure of $\ell = 1$ modes from the pure p-mode values. The pure p-mode $\ell = 0, 2$ and 3 modes appear as straight vertical ridges in this parameter space, while the $\ell = 1$ modes have a far more complex structure.

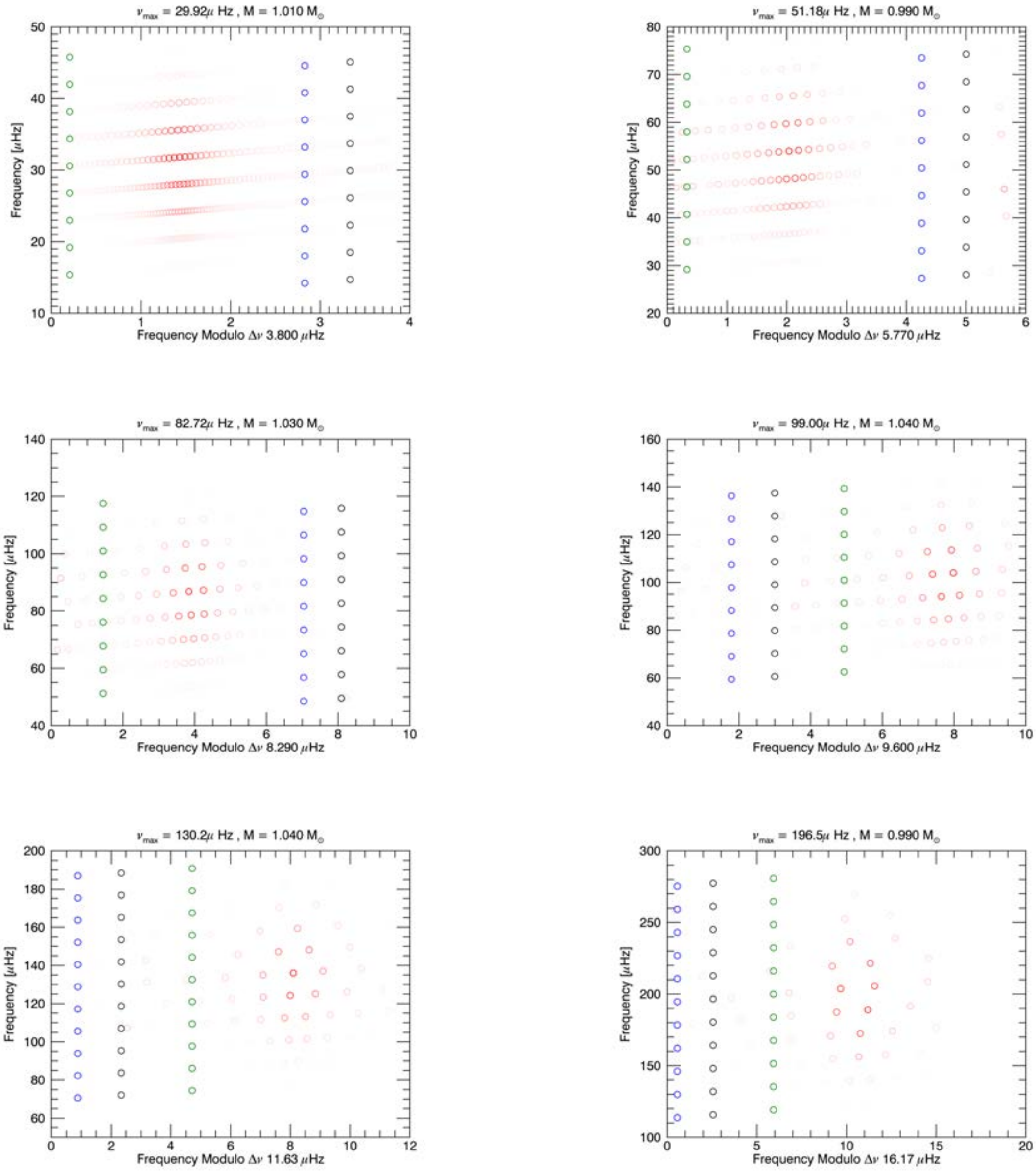


Figure 3.14: Échelle diagrams of the 6 simulated stars. Modes of degree $\ell = 0 - 3$ are shown, and are in black, red, blue and green respectively. $\ell = 1$ modes are shown on a colour scale where the most intense colour shows the more powerful modes.

It is these spectra that the variance metric is taken from, the power spectral density within the frequency range shown in Figure 3.1 integrated to give our bandpass filtered estimate of the light curve variance for the simulated stars. This metric informs the binary fraction inference in Chapter 4 and automated detection of oscillations in Chapter 5. Figure 3.15 shows the scaling of this quantity with $\nu_{\max}$ for the simulated population and the *Kepler* stars, again showing the very good agreement between the simulations and observations.

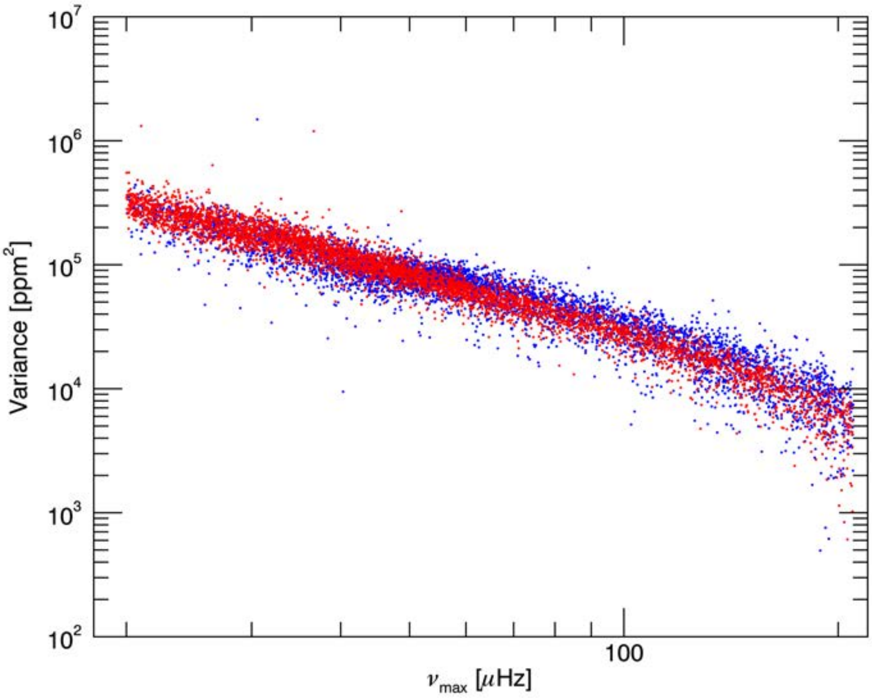


Figure 3.15: The variance for *Kepler* red giant branch stars is shown in blue, and the simulated values for the simulated TRILEGAL RGB stars is shown in red. The TRILEGAL population here has a binary fraction of 50%.

Chapter 4

Binary Fraction Inference From an Ensemble Study

While Chapter 5 will discuss the usefulness of predicting and measuring the variance metric for individual power spectra, this chapter will discuss the application of the variance metric introduced in Chapter 3 to an ensemble study of *Kepler* red giant stars in order to determine the fraction of the population which exist in binary systems. The inferred binary fraction is presented here, along with an explanation and justification of the methods and assumptions employed, and the advantages and limitations of this ensemble study. This will be proceeded by a brief review of binary star detection methods.

The washout effect induced by companion stars as introduced in Chapter 3 can allow inferences to be made about the fraction of stars in the population which have a companion, subject to certain assumptions. An introduction to the data and simulations used follows, along with an explanation of the methods and assumptions made in this study.

4.1 An Introduction to Binary Star Detection

A pair of gravitationally bound stars orbiting their common centre of mass are referred to as a binary star system. It is thought that a large fraction of stars in our Galaxy reside in these types of systems, and a significant amount of research has gone into devising methods of detecting them. Binary stars are classified depending on how the second star is detected. Visual binary stars are those which have a large enough angular separation to allow them to be resolved separately, and relative levels of brightness such that glare from the primary does not make the secondary impossible to detect. These binary systems usually have a fairly wide separation and consequently long ($\sim$ years - centuries) period orbits. These observations allow the orbits of these systems to be solved, that is the semi-major axes of the orbital ellipses, the eccentricity and inclination of the orbits can be determined, and through applying Kepler's third law and by knowing the parallax of and hence distance to the system, the masses of the stars can be determined.

Spectroscopic observations can reveal the presence of a secondary star. In the case that only the spectral lines of one star are visible, they may oscillate periodically in wavelength due to being Doppler shifted during the orbit around the centre of mass of the binary system. If both sets of spectral lines are visible, then due to the Doppler shift, the lines will appear as double or single depending on the motion of the stars relative to the observer. As a star moves towards or away from the observer, the spectral line will be blue or red shifted, and hence appear as a double line in the spectrum, while when there is no motion relative to the observer the line will appear as a single. If the stars are of two different spectral types, the presence of an additional star can be inferred from the unusual shape and intensity of the spectral line, relative to that of a single star. The Doppler shifting of the lines in spectroscopic binaries allows the product of the semi major axis and the angle of inclination of the orbital plane of the system to be determined, but information from another type of observation (such as an eclipse) is needed to fully solve the orbit.

Astrometry is the precise measurement of the position of a star. A star with a companion which is not visible will not follow a straight path as it moves across the sky, instead the centre of mass of the system follows a straight path while the star will "wobble" periodically back and forth across this trajectory. This method has historically been applied to nearby stars with a high proper motion, however the Gaia mission (Gaia Collaboration, 2016) performs astrometry to unprecedented levels of accuracy, and so at far greater distances. It is expected that it could discover millions of binary systems in this way (Eyer et al., 2012).

The transit method used by *Kepler* to detect exoplanets is pre-dated by eclipse observations of binaries, though the concerns of making these kinds of observations are similar. As with exoplanets, the system must be oriented such that the orbital plane is inclined edge on to the line-of-sight so it is possible to observe the secondary star eclipsing the primary. The secondary star of course emits it's own light so the periodic drops in brightness will have different depths depending on whether the secondary is in front of or behind the primary, or whether both stars are fully visible, distinguishing it from a planetary transit. This is only possible if the orbit of the system is inclined very close to 90° relative to the plane of the sky. The large number of *Kepler* targets means that although this alignment is rare, over 2500 eclipsing binaries have already been discovered by *Kepler* (Slawson et al., 2011).

While all of these methods can provide sensitive information on the masses and orbits of the stars, they are all limited in their detection of the binary by the location, orientation and particularly the orbital period of the stars or through requiring other information such as parallax. In using asteroseismology to detect binary stars, these additional constraints are not necessary, and so the data from *Kepler* are sufficient in isolation for binary detection. *Kepler* has detected solar-like oscillations in hundreds of main-sequence and subgiant stars, and in over ten-thousand red giants (Miglio et al. (2014) and references). Red giant oscillations are of a much lower frequency than those of less evolved stars and so are detectable by the long cadence observations, for which data are available on many stars, while the oscillations of main-sequence and subgiant stars require short cadence observations, which can only be performed on a select number of targets. For individual cases, the presence of two oscillation spectra in a single lightcurve is a clear signature of binarity, referred to as a seismic binary. Figure 4.1 shows a seismic binary, clearly showing the signatures of 2 oscillating stars. While this can also be due to a foreground and background star which are in close proximity as observed by *Kepler* while not being gravitationally bound, Miglio et al. (2014) find that the fraction of apparent seismic binaries that suffer this contamination is extremely small.

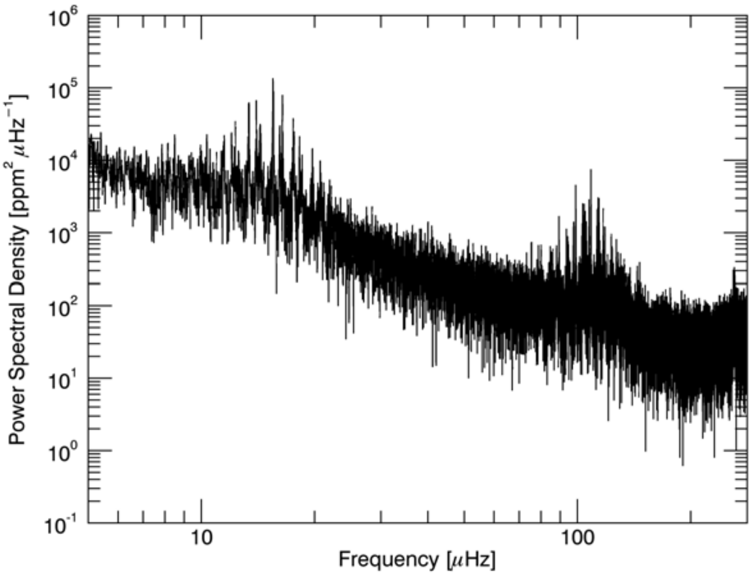


Figure 4.1: A power spectrum showing the oscillations of two red giant stars observed by *Kepler* in the same lightcurve.

Though stellar population modelling codes give consideration to the binary fraction, and the formation and evolution of binary stars have implications for our understanding of Galactic history, no consensus seems to exist on the fraction of stars in binary systems. The large number of light curves we have at our disposal thanks to *Kepler* and the fact that the asteroseismic variance is insensitive to the orientation of the binary system means it has the potential to be a valuable tool in inferring the binary fraction of the population.

4.2 The Role of Variance in Binary Fraction Inference.

As shown in Chapter 3, the variance metric is ostensibly an intrinsic stellar property, however the washout factor described by Equation 3.42 reduces the power present in the power spectrum, and hence the variance. Figure 4.2 shows an artificial asteroseismic binary created by combining the light curves of two *Kepler* red giant stars and producing a power spectrum where both sets of oscillation spectra can be seen.

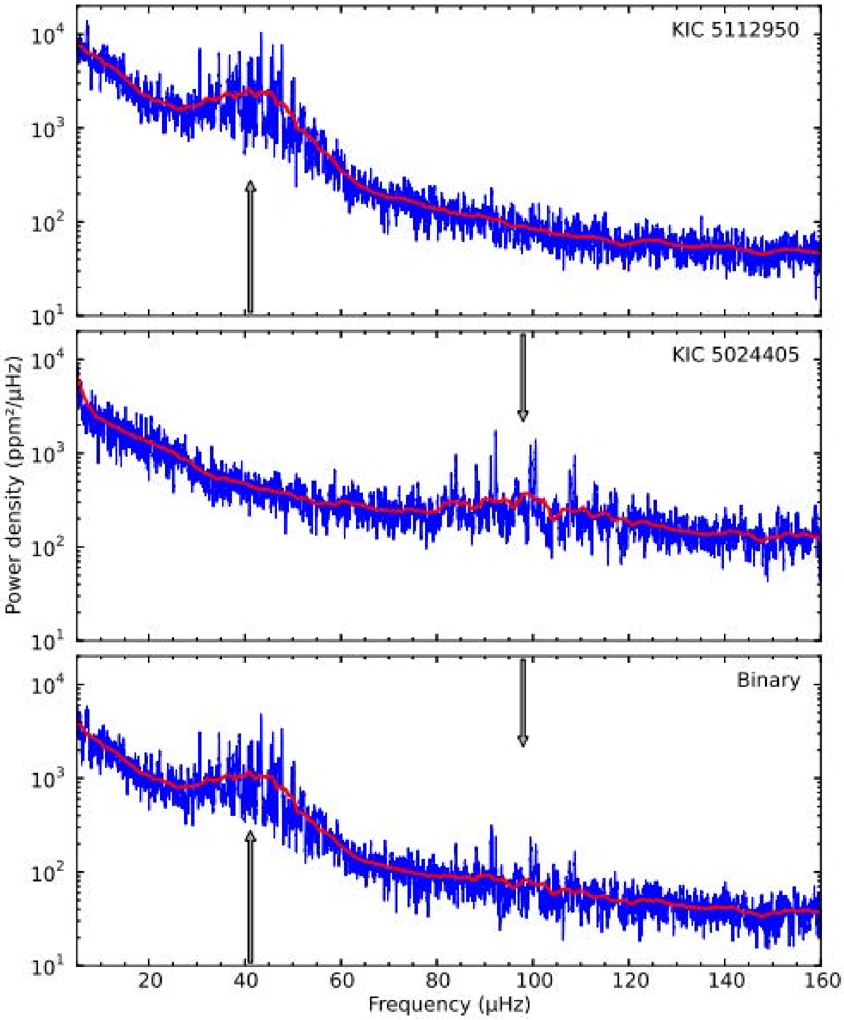


Figure 4.2: The bottom panel shows the power spectrum of an artificial seismic binary taken from Miglio et al. (2014) created from combining the light curves of the two red giants whose power spectra are shown in the top two panels. Note the reduction in power of the oscillations in the combined spectrum.

The power present in both sets of oscillations is appreciably reduced in the combined power spectrum. For true binaries where the luminosities of both stellar components are comparable, this effect is large, as in the case above, while in the situation that the secondary star is much fainter, for example in the case of a red giant with a main sequence companion, this effect is more subtle. It is therefore sensitive to the mass ratio of the components of the binary.

Figure 4.3 shows the scaling of the variance with ν_{max} for a sample of *Kepler* red giants, a relationship which approximates a power law. As the washout factor only depends on the magnitude in the *Kepler* bandpass, we can expect stars with a companion to have lower than expected variance for their ν_{max}. Of course, there is scatter intrinsic to the variance$-\nu_{\text{max}}$ relationship regardless of the binary fraction. Since the relationship shown in Figure 4.3 may not be an exact power law, a median fit is performed by taking the logarithm of both the frequency and variance values and binning the data by log-frequency. The median frequency and variance values for each bin are taken, then a cubic spline interpolation gives a median value across the entire range of frequencies, shown in Figure 4.3 as a red line. The residuals in dex are then given by

$$\log_{10}\text{FV} - \text{fit}. \tag{4.1}$$

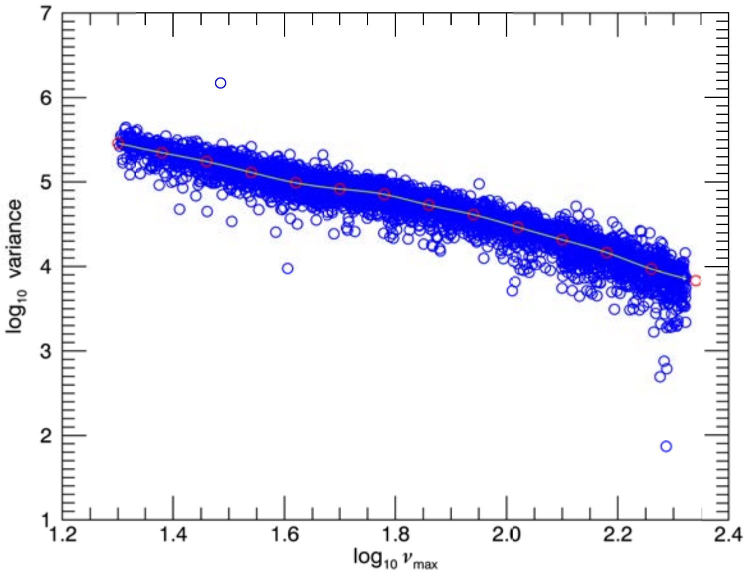


Figure 4.3: Filtered variance - ν_{max} relationship for 5283 *Kepler* red giant branch stars. Median values of log-frequency bins are shown in red and a cubic spline interpolation fit line in green.

A histogram of these residuals about this fit are shown in Figure 4.4. Notice that the distribution of residuals is decidedly asymmetrical, the "tail" of the distribution on the left suggests a proportion of the stars have a lower variance than their $\nu_{\rm max}$ suggests they should have. Under the hypothesis investigated in this work, the presence of this asymmetry is due to contamination of the lightcurves by companion stars.

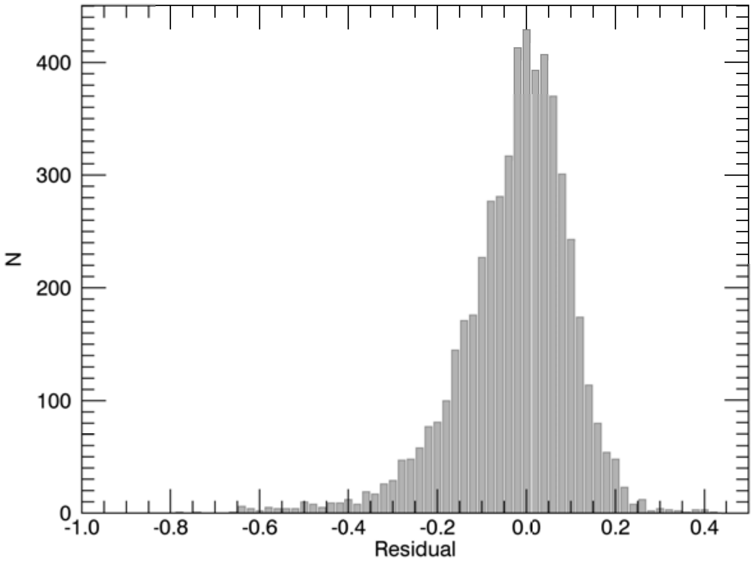


Figure 4.4: A histogram of the residuals of the data about the fit shown in Figure 4.3 obtained following Equation 4.1.

To assess the impact of the binary population on the distribution of residuals in real stars, we need multiple simulated populations with known binary fraction. Chapter 3 describes the process of simulating power spectra in detail.

4.3 Using Simulated Populations

Following the procedure outlined in Chapter 3, the fundamental properties generated by stellar population synthesis code TRILEGAL (see Chapter 2) are used as input to construct a realisation of the artificial 4-year power spectrum for each star in the simulated population, and the variance estimate is then extracted as per the real data, by integrating the spectrum over the range of frequency occupied by the modes of oscillation, as defined by the width of the envelope modulating the modes of oscillation thus;

$$\nu_{\mathrm{max}} \pm \delta\nu_{\mathrm{env}} \approx \nu_{\mathrm{max}} \pm 0.66\nu_{\mathrm{max}}^{0.88}. \tag{4.2}$$

Cuts are made in various parameter spaces to most closely match the real data, as described in Chapter 2. A simple false alarm detection test as described by Chaplin et al. (2011) is applied to determine whether the simulated oscillations ought to be detectable by *Kepler*, however in practice all of the simulated stars pass this test. Given the high amplitude oscillations exhibited by red giant stars, this is not surprising. Hekker et al. (2011) finds that oscillations were detected in 71% of red giants observed by *Kepler*. Of the remaining stars, ≈ 90% have a predicted ν_{max} which lies outside of the frequency range which can be explored with the long cadence observations, the upper limit of which is set by the Nyquist frequency of ≈ 283μHz. The lower limit of around 10μHz is due to the high amplitude granulation signal at these frequencies, along with the frequency resolution of the observations. Explanations for the few stars whose oscillations are predicted to lie within the explored frequency range but are not detected are given by Hekker et al. (2011) to include instrumental effects and high shot noise. Since this number is so small, we do not consider selection effects within the *Kepler* sample to pose a problem in comparing them to our simulated populations.

We will here exclude stars with ν_{max} less than 20μHz, as there is a significant excess of these stars in the simulated populations as compared to the real stars, again due to the difficulty of detecting these low frequency oscillations. This is a very conservative measure, in case this has the potential to introduce bias into the binary fraction estimation.

A further consideration needed to match the observations as well as possible is the uncertainty on the observed seismic parameters, $\Delta\nu$ and ν_{max}. Following empirical relations, derived by Chaplin (private communication) and discussed in Miglio et al. (in prep.), which provide an excellent match to the errors on the real data, we scatter the "true" values obtained from the scaling relations (see equations 2.6 and 2.8) using

$$\text{error}, \nu_{\max} = \frac{\nu_{\max}}{53.75}(10.0^{1.2-0.71\log_{10}\nu_{\max}+0.35x} + 1.0) \tag{4.3}$$

and

$$\text{error}, \Delta\nu = \frac{\nu_{\max}}{750.0}(10.0^{1.7-0.99\log_{10}\nu_{\max}+0.15x} + 1.0). \tag{4.4}$$

The variance and $\nu_{\max}$ relationship is simulated for binary fractions from 20% to 80% at intervals of 10%, following exactly the same process to obtain the residuals as used for the real stars. These fits are shown in Figure 4.5. We will only look at fractions from 20% to 80% as current estimates suggest that 20% would seem to be too low for a reasonable estimate, and above 80% would suggest that single stars are much more rare than is reasonable for the mass range of stars we consider. Having realistic simulated populations with a range of binary fractions will be our guide to the interpretation of the real stars, and care is taken to make the same assumptions of both the real and simulated data.

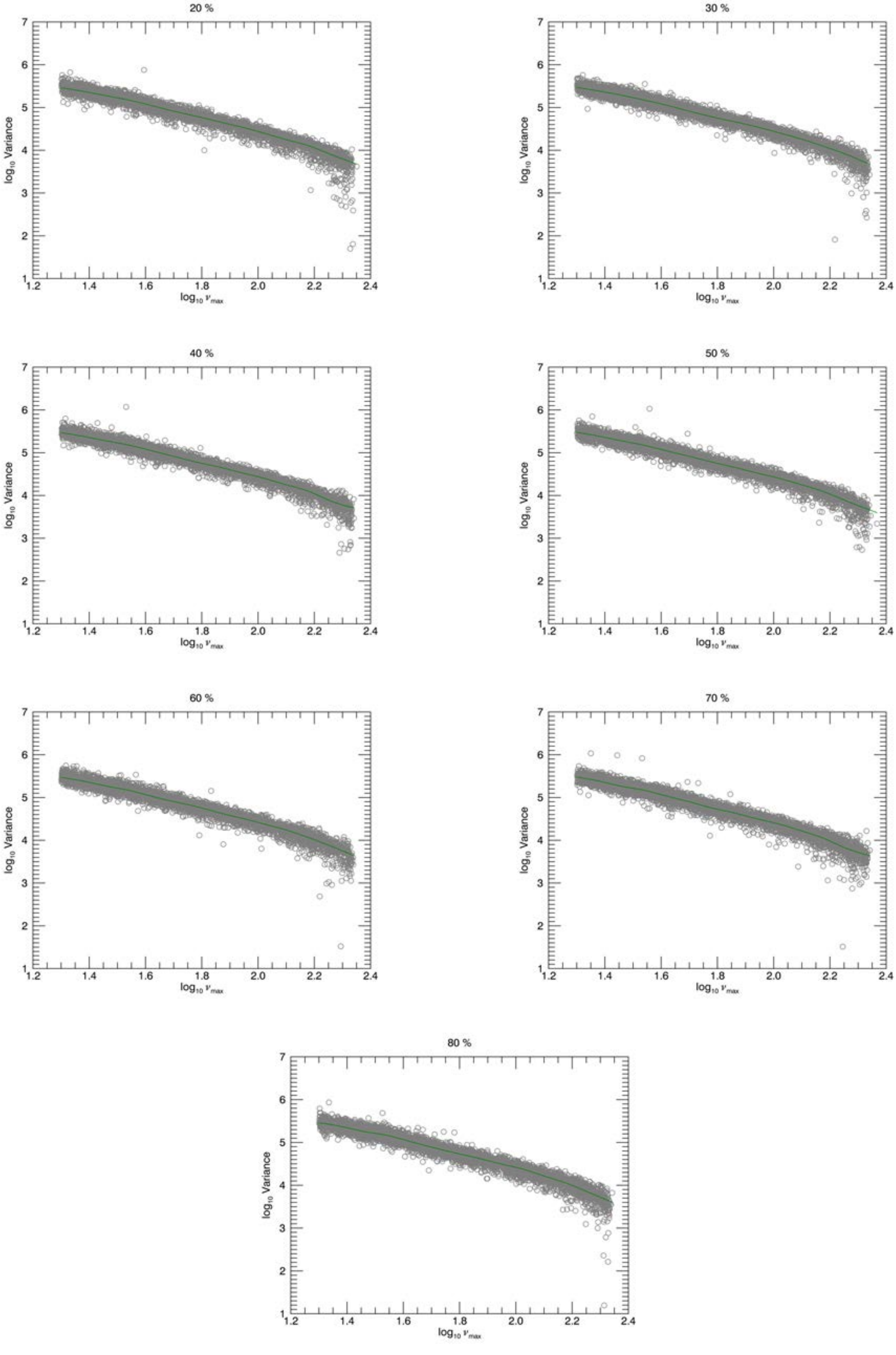


Figure 4.5: Simulated $\nu_{\rm max}$ vs. Filtered Variance relationships with binary fractions from 20% to 80%.

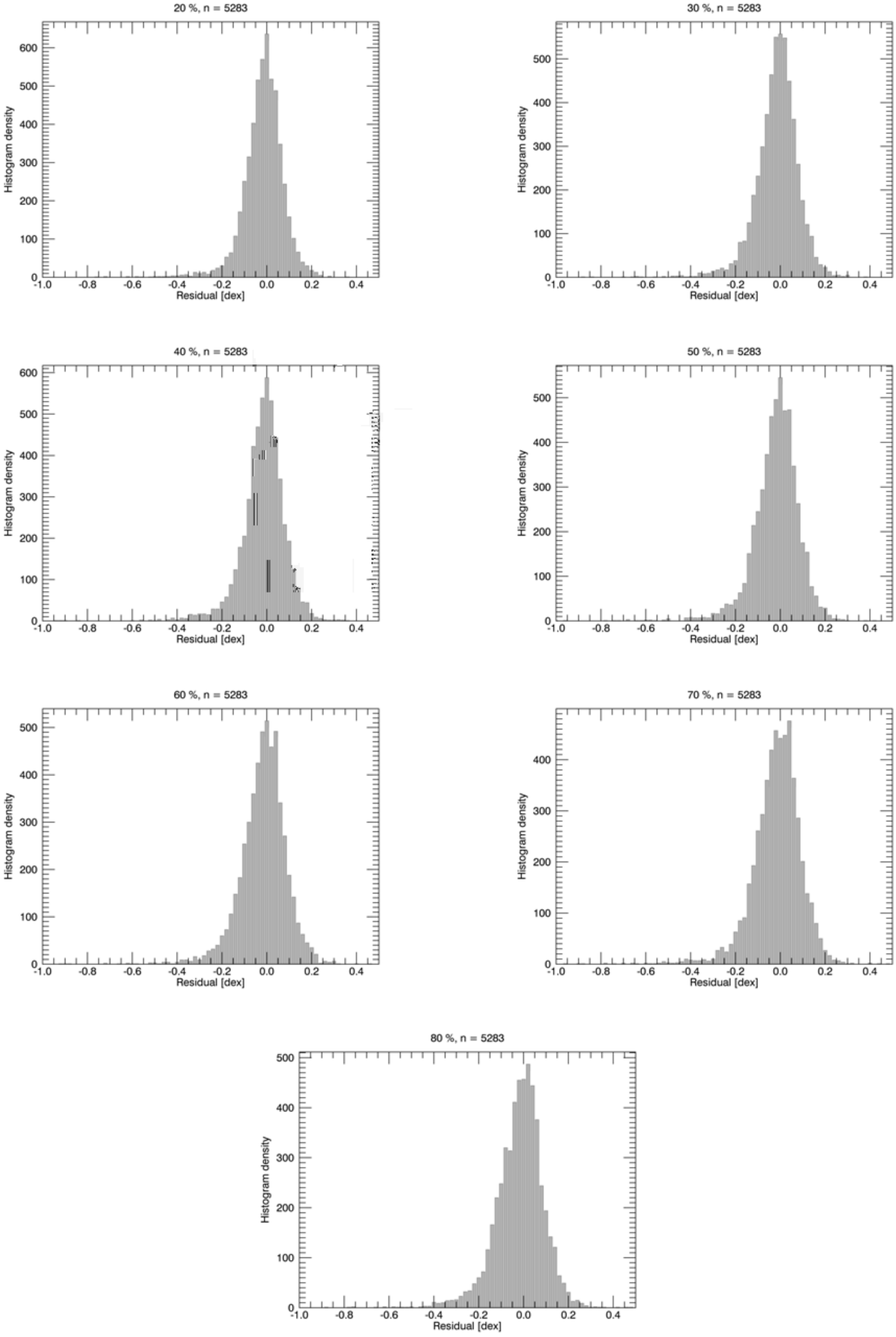


Figure 4.6: Distributions of residuals from fits to ν_{max} vs. Filtered Variance relationships for binary fractions from 20% to 80%.

Figure 4.6 shows the distributions of residuals for the simulated fields for seven different binary fractions. The subtle departure from symmetry is appreciable visually for all the fractions, due to the varying proportion of binary stars experiencing washout and therefore displaying a lower filtered variance and residual. We must consider the possibility that other factors may have an effect on the shape of the distributions, and that in case of the real stars, there may be yet more influences which are not included in the simulations.

4.4 Other Potential Contributors to Asymmetry

Due consideration must be given to other potential causes for the appearance of the lower variance stars, and how other factors may affect the distributions, or there is a risk of biasing the inferred fraction. These can be due to properties of the population as a whole, or properties of individual stars. The phenomena considered, and what effect they may have on the distributions are discussed below.

4.4.1 Binary Mass Ratio Distributions

Figure 4.7 shows the mass distribution and washout factors for 3 simulated sets of stars, with the same cuts and error assumptions as described in the previous section and Chapter 2. The left panel shows the ratio of secondary to primary mass for three simulated populations with uniform initial mass ratio distributions between [0.7-1.0], [0.8-1.0] and [0.9-1.0] respectively, serving only to confirm the mass ratio distribution for red giants is consistent with the distribution of the whole population.

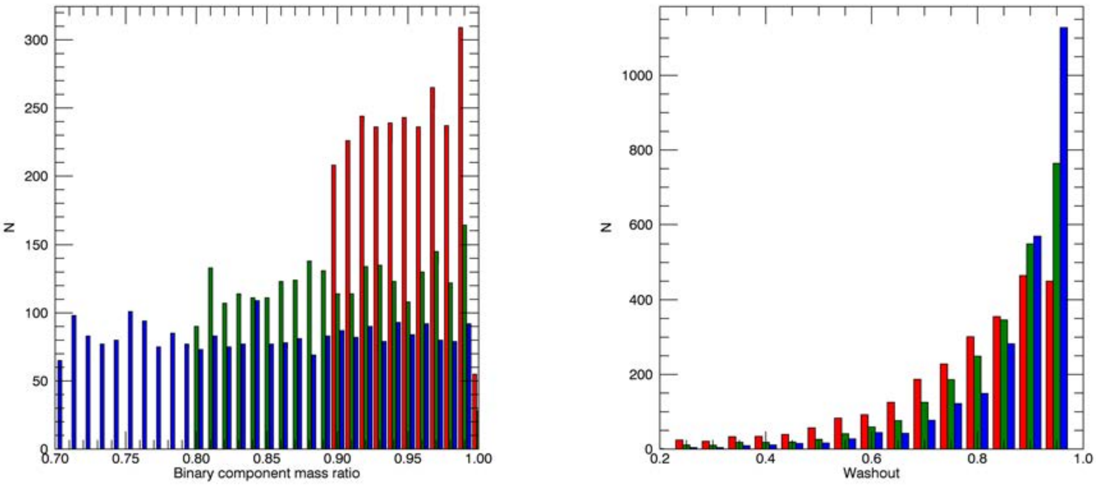


Figure 4.7: Left hand panel: The mass ratio distribution for three different populations. Right hand panel: The distribution of washout factors associated with these mass ratios. Blue, green and red show these for populations with uniform mass distributions on the range [0.7-1.0], [0.8-1.0] and [0.9-1.0] respectively.

The right hand panel shows the washout factor $\mathrm{k_{washout}}$ for the same stars. It is clear that for all three mass ratio distributions, the majority of the stars experience a very modest level of washout. Recall that the washout factor is multiplicative, so a value close to unity implies low dilution of the power spectrum (see explanation in section 3.9). The average washout factor is lowest, and hence the levels of dilution greatest, for the mass ratio distribution closest to unity.

Figure 4.8 shows the effect of the average level of washout on the distribution of the residuals. The effect is subtle, but the separation of the two underlying distributions is greater for the cases which have a larger average washout.

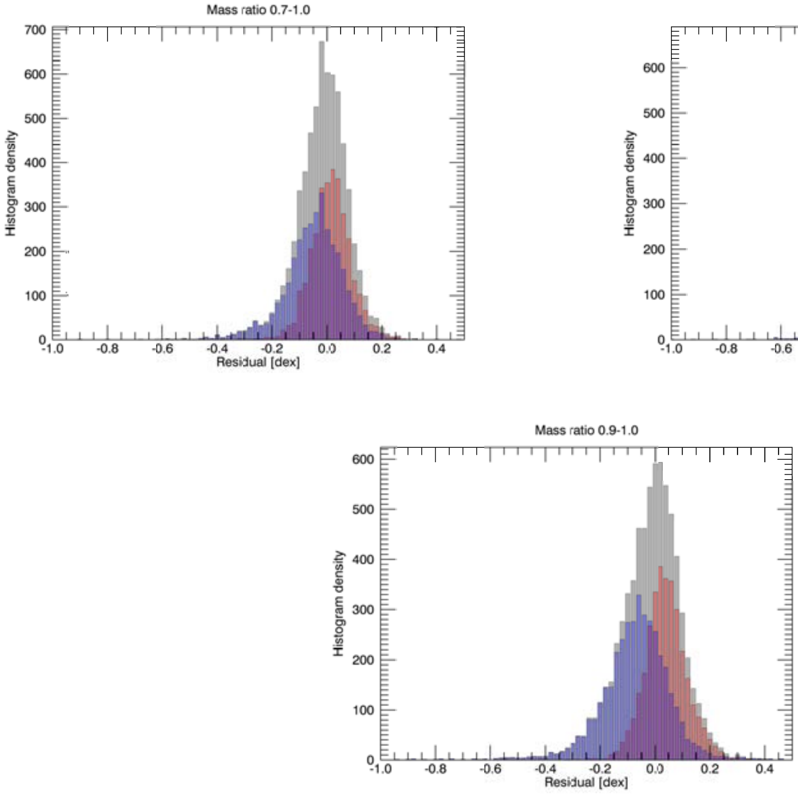


Figure 4.8: Histograms of residuals for the three above mass ratio distributions.

The mass ratio distribution is not something we can know for sure for the real stars, however earlier work (eg. Miglio et al. (2014)) suggests that a mass ratio near unity is reasonable, and so a uniform mass ratio distribution between 0.9 and 1.0 is adopted for the simulations used. As discussed in the results section of this chapter, this seems to be a reasonable assumption for the real data.

4.4.2 Scaling Relation Dependencies on Mass

Chapter 3 discussed in some detail the approach used to simulate the power spectra we use in this work. The seismic parameters which govern the amount of power in the frequency region we are interested in for the filtered variance metric all have a dependency on mass, and for the amplitude parameter, this can vary quite dramatically between relations. To satisfy ourselves that this is not the cause of the deviation from a normal distribution, we can look back at the scaling relations we tested in Chapter 3 and determine whether, in the simulated case of single stars only, the different dependence on mass can induce any asymmetry.

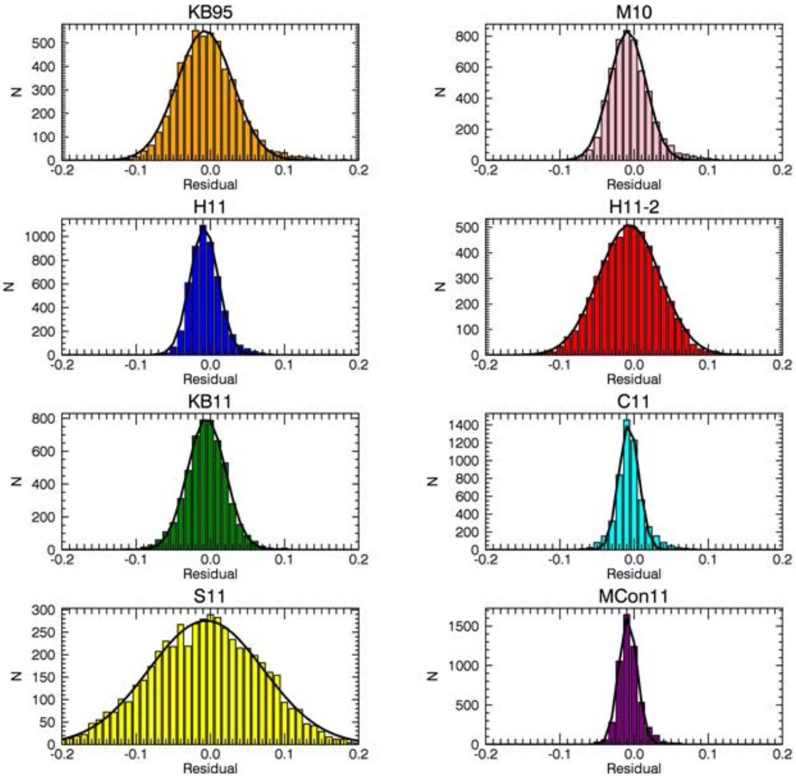


Figure 4.9: Following the same labelling and colour coding as in Chapter 3, these histograms show the distribution of residuals for the single-star-only case for the estimated variance simulated using each of the eight amplitude scaling relations. The black curves are a Gaussian fit to the histograms, serving as a guide to the eye.

For this simple sanity check, we use only the basic estimation for the variance, treating the mode power as the smoothed Gaussian envelope only, and assuming the background power at $\nu_{\rm max}$ adequately captures the background power across the frequency range of interest. This leads to the distributions being generally narrower than those obtained for those produced by simulating the power spectra with all the features, noise and uncertainties they contain. This test does however show that the differing mass dependency of the amplitude scaling relations does not induce any significant asymmetry into any of the above distributions. Recall that in Chapter 3, the H11-2 relation (shown in red above) was chosen for the simulations, and for the 0% binary case shown here, there is no asymmetry, and hence no reason to suggest that the mass dependency is contributing to the non-Gaussian appearance of the distributions shown in Figure 4.6.

4.4.3 Metallicity

It has been suggested that the mode amplitudes are expected to change with metal abundance (eg. Houdek et al. (1999), Samadi et al. (2010)). To assess whether this would have an impact on the asymmetry of the distribution, it is helpful to look for a relationship between the residuals about the fit and the metallicity. Figure 4.10 shows this for the real stars for which APOGEE abundances (Pinsonneault et al., 2014) are available.

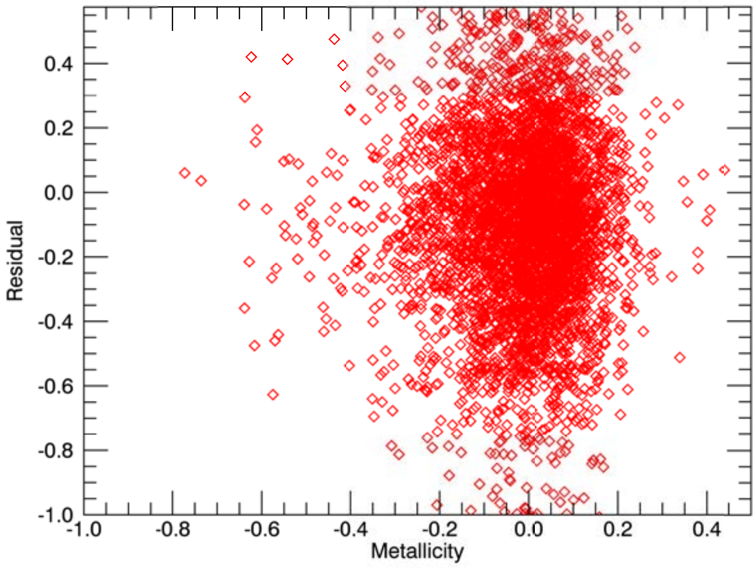
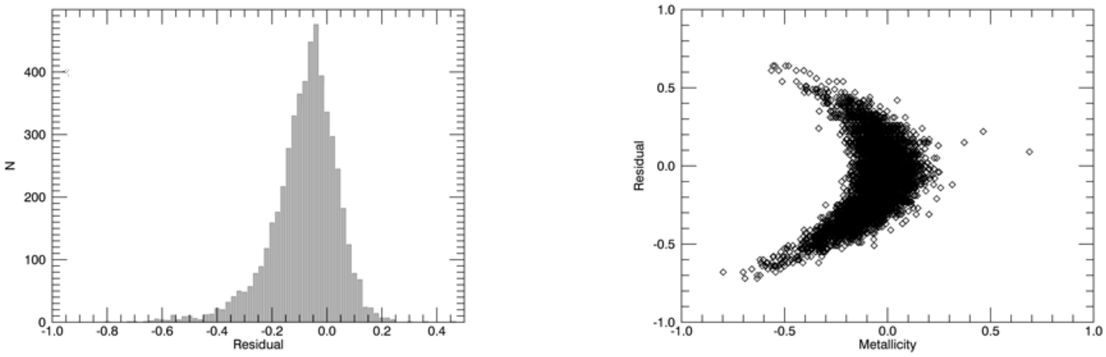


Figure 4.10: APOKASC metallicities and residuals for *Kepler* red giants.

Attempting to fit these data shows no relationship, the above data having a Pearson correlation coefficient of only 0.002, implying no correlation between metallicity and residuals. In attempting to duplicate the asymmetry by including a metallicity dependence in the scaling relations for amplitude, it was found that in order to produce a significant asymmetry, the effect of abundances on the amplitudes would have to be so large such that it would have been readily obvious from the work of the APOKASC (Pinsonneault et al., 2014) collaboration.



(a) A distribution of single-only simulated stars with an asymmetry induced by metallicity.

(b) The metallicity-residual relationship for the single star only case shown in the left hand panel.

Figure 4.11: The effect of a strong metallicity dependence on the residual distribution of a synthetic population with no binary stars.

Figure 4.11 shows a simulation of purely single stars with a strong dependence on metallicity introduced into the residuals. Notice that the asymmetry in the histogram of residuals appears similar to that for real stars (see Figure 4.4), however it is appreciable in the scatter plot that such a strong metallicity dependence would induce an obvious correlation between these parameters, which is simply not present in the real sample. It is therefore unrealistic for metallicity effects to be the cause of the asymmetry we see in the real stars. In this work, we hence choose to include no such dependence on metallicity, and therefore assume that it does not contribute the shape of the distribution.

4.4.4 $\ell = 1$ suppression

In recent years, the appearance of missing or suppressed $\ell = 1$ modes in intermediate mass red giant stars has been attributed to strong interior magnetic fields (eg. Mosser et al. (2012a), García et al. (2014) & Stello et al. (2016)). Clearly, if there is reduced mode power due to this phenomenon, the variance will be reduced. However, since the red giant branch stars used for this work are distinguished from the red clump by the method described in Elsworth et al. (2017), stars that have very highly suppressed or entirely missing $\ell = 1$ modes are removed from the sample as the method relies on the detection of enough significant peaks in the power. Hence, the remaining stars that are classifiable by this method do not a significant amount of $\ell = 1$ suppression.

4.5 Beyond Asymmetry

Various measures were used to try and capture the asymmetry of the distribution. These include Pearson's moment coefficient of skewness, derived from the normalised third-moment of the data, defined as

$$\gamma = \frac{1}{n_x \sigma^3} \sum_{i=0}^{i=n_x} (x_i - \bar{x})^3 \tag{4.5}$$

(e.g Barlow (1989)), where we have n_x data with standard deviation σ. An alternative measure is based on finding the narrowest interval of the distribution which contains 68% of the data, and comparing this to what we would expect for a normal distribution. In the case of a symmetrical, normal distribution, we would expect 68% of the data to lie between $\pm 1\sigma$, with the same number of data lying above zero as lies below. Theoretically, assessing the range over which 68% of the data lie in our asymmetric distribution and whether there are more data above or below zero ought to allow some measure of asymmetry to be made. In practice however, these tended to be degenerate, have no clear trend with binary fraction or, due to the reliance on the standard deviation σ, were extremely sensitive to outliers.

An alternative solution considered is fitting the distributions with a skew normal distribution, defined by

$$f(x) = \left[1 + \mathrm{erf}\left(\frac{\alpha x}{\sqrt{2}}\right)\right] \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}. \tag{4.6}$$

This is a generalisation of the Gaussian case, with the additional shape parameter α to account for the skew. As can be appreciated from Figure 4.12, an increase in binary fraction does increase weight in the tail of the distribution, which on its own would increase the skewness, but it also causes the peak to become broader, hence leading to the overall "shape" to appear less skewed, also excluding this as a useful way of measuring asymmetry as a function of binary fraction.

It was decided that, rather than thinking of the distribution as a single skewed distribution, it is helpful to view it as the sum of two distributions which are described by two sets of parameters; one consisting of single, non-diluted stars, and the other of the stars with a companion, all experiencing some level of dilution. This is hinted at in Figure 4.7. In modelling our distributions in this way, we can now fit the overall distribution of real stars with an appropriate model and determine the binary fraction with uncertainties in a much more robust way. What follows is the method and results from this approach.

4.6 Bayesian Approach

It is helpful to think of the distribution of the residuals, as shown in Figure 4.4, not as a single asymmetric distribution, but a mixture of two distributions with different properties, one for the single stars, and one for binaries. As we have at our disposal simulations of which we understand and can control all of the underlying properties, it is possible to place constraints on what form those distributions are likely to take in the real data. This idea of having prior knowledge about the data lends itself well to the application of techniques based on conditional probability; measuring the likelihood of an event, here a model with parameters $\mathbf{m}$ of a certain form fitting some observed data (sometimes termed evidence) D, given some prior information I about the model, here constraints on the values of model parameters based on our realistic simulations. Bayesian inference computes the probability of the model being reasonable given all the information we have. This can be represented in terms of other probabilities using Bayes' Theorem, following the convention from Basu and Chaplin (in the press):

$$p(\mathbf{m} \mid D, I) = \frac{p(\mathbf{m} \mid I) p(D \mid \mathbf{m}, I)}{p(D \mid I)} \tag{4.7}$$

$\mathrm{p}(\mathbf{m} \mid \mathrm{D}, \mathrm{I})$ is the posterior probability, meaning the probability of the model with parameters m being appropriate for our observed data, D. This is the what we seek to determine.

$\mathrm{p}(\mathbf{m} \mid \mathrm{I})$ is the prior probability. This contains information on what we believe about the parameters of the model, before considering the evidence. This can be used to place common sense constraints on parameters, eg. disallowing negative heights when fitting Gaussians to our distribution, and can also contain information on what values of parameters are sensible, here taken from our simulated populations. The exact priors used are discussed below.

$\mathrm{p}(\mathrm{D} \mid \mathbf{m}, \mathrm{I})$ is the likelihood function. This essentially is the probability of observing the evidence given the model is correct. The form of this is dependent on the statistics we assume the observations to be distributed about the model values with. This is also commonly written as $\mathcal{L}(m)$.

$\mathrm{p}(\mathrm{D} \mid \mathrm{I})$ is the global likelihood. As it is not a function of the model parameters m it does not impact on assessing one set of model parameters against another and therefore we do not consider this term.

Without the global likelihood, we take a Maximum A Posteriori approach, seeking to find the set of parameters which will maximise

$$\mathrm{p}(\mathbf{m} \mid \mathrm{I})\mathcal{L}(m) \tag{4.8}$$

Using the simulations, we decide on a sensible form for a model, and place priors on the parameters of the model. Using the *emcee* algorithm (Foreman-Mackey et al., 2013) to implement a Markov Chain Monte Carlo method, we are able to sample the various parameter spaces and estimate the posterior distribution of these and maximise the likelihood function, subject to the priors we adopt. The form of the model, likelihood and priors are discussed below.

4.6.1 The Model

The model we fit to the real stars will be directly informed by those that fit the simulated data. Figure 4.12 shows histograms of the residuals for the simulated data, with the two underlying distributions for the single and binary stars shown, both of which approximate Gaussian distributions reasonably well. The fit to the ν_{max}-variance relationship and calculating the residuals is performed in exactly in the same way described for the real stars in Section 4.2. Then the residuals are separated according to whether they are singles or binaries and made into two histograms. Gaussians are then fitted to them separately. This suggests that a good model form for the distribution of the overall distribution of these data could be described simply as

$$a = \frac{H_b}{\sigma_b\sqrt{2\pi}} e^{-(x-\mu_b)^2/2\sigma_b^2} + \frac{H_s}{\sigma_s\sqrt{2\pi}} e^{-(x-\mu_s)^2/2\sigma_s^2}. \tag{4.9}$$

In this model, the binary fraction is then simply given by

$$\frac{H_b\sigma_b}{H_s\sigma s}. \tag{4.10}$$

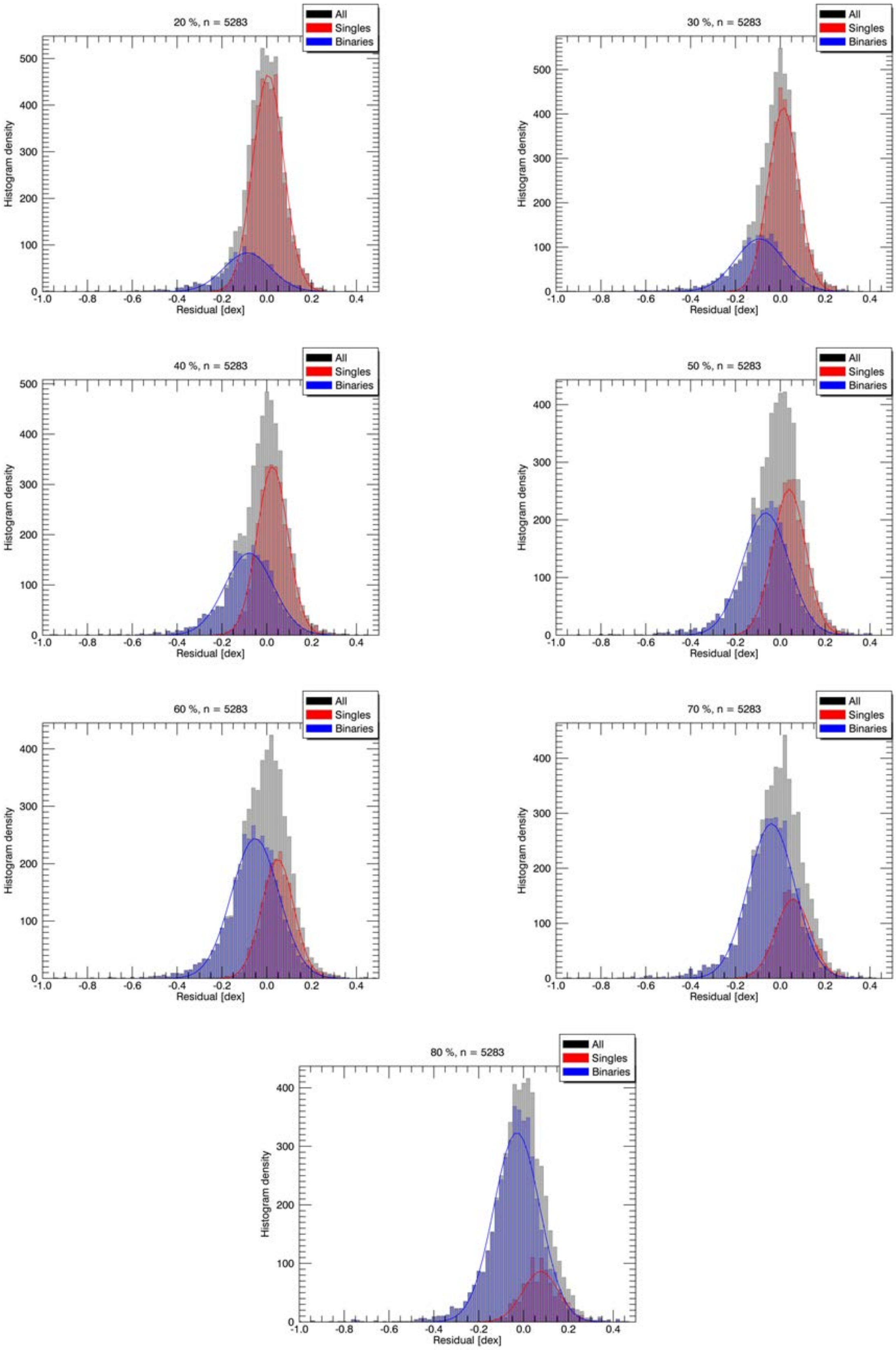


Figure 4.12: Histograms of residuals for the simulated populations showing the two underlying distributions corresponding to the single and binary stars and Gaussian fits to them.

This model form means we have six parameters to sample using the MCMC algorithm. We now adopt priors that place constraints on the values these parameters can take, based on the simulated data.

4.6.2 Priors

Assuming as we must that the simulations are a realistic representation of realistic populations of the given binary fractions, we can place constraints on the values the parameters of the model can take, and include these in the Bayesian formulation as the prior probability. These priors can take any form, but for this problem, it is sufficient to use flat priors, purely setting limits on the values that the parameters can take, and rejecting any model that requires any parameter to fall outside of these limits. Some of these are purely intuitive: not allowing the height or width of either Gaussian to be negative, for example, and some come directly from the simple fits to each Gaussian shown in red and blue in Figure 4.12. Figure 4.13 shows the values of the parameters of the fits to the simulated distributions, and it is these we use to set our priors.

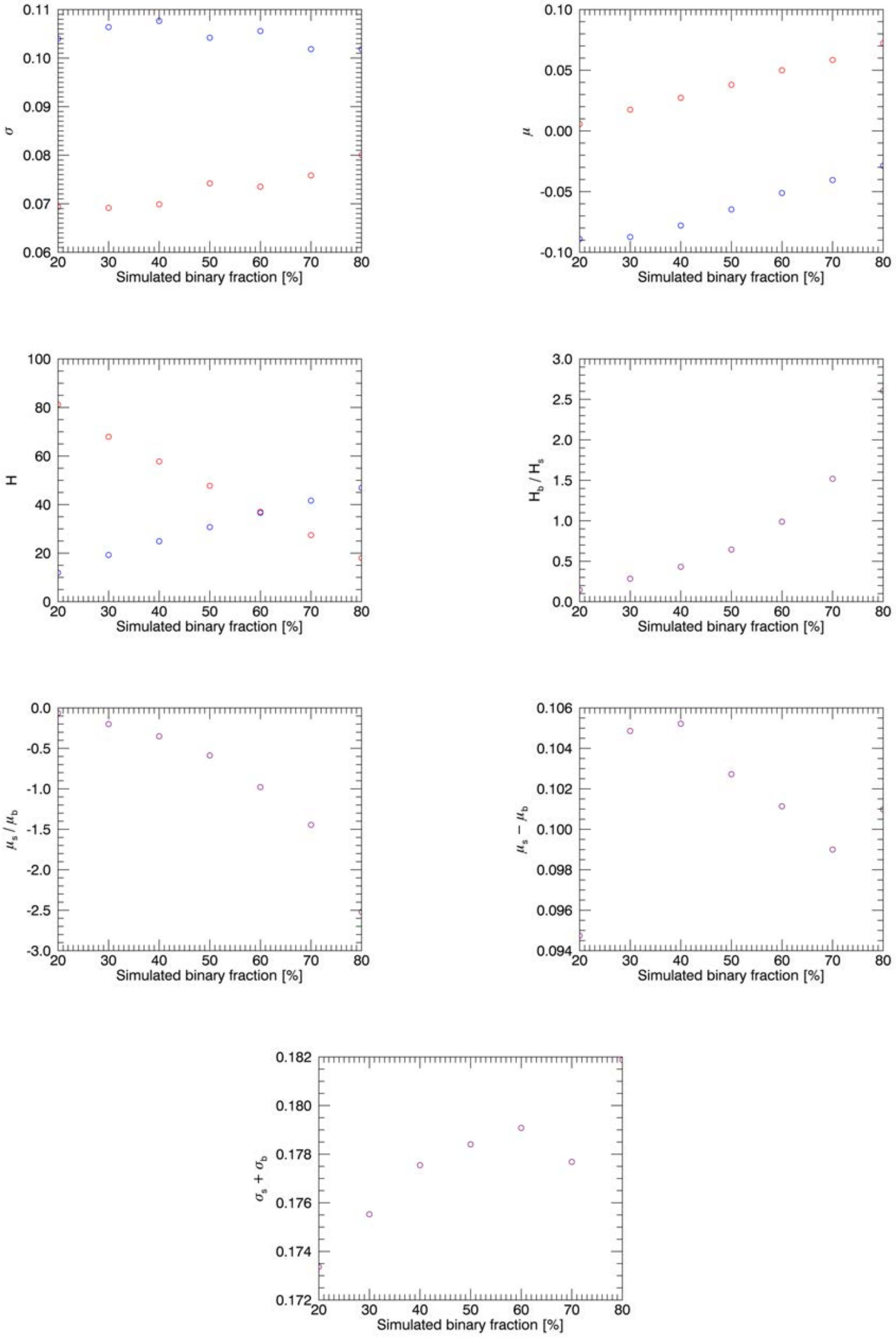


Figure 4.13: Values of the six model parameters, ratios and differences as labelled. Red denotes quantities for the distribution of single stars, blue that of binary stars and purple are quantities which relate properties of the single and binary distributions.

Parameter(s)	Limits
H_s	>0 <80
H_b	>0 <60
σ_s	>0.065 <0.085
σ_b	>0.1 <0.11
μ_s	>0.0 <0.1
μ_b	>-0.1 <0.0
$\frac{H_s}{H_b}$	<3.0
$\sigma_s + \sigma_b$	>0.172 <0.182
$\mu_s - \mu_b$	>0.094 <0.106
$\frac{\mu_s}{\mu_b}$	>-3.0

Table 4.1: Priors for simulated stars with mass ratio distribution [0.9-1.0]

Limits are placed on the values of the six individual parameters, as well as the ratio of the heights, preventing either the singles or binaries from dominating, the ratio of the centres and the ratio and sum of the standard deviations, ensuring that the calculated overlap of the distributions and their relative widths are consistent with the actual values. These priors do not favour any particular binary fraction.

In the MCMC algorithm, as we do not impose a certain form on the priors (that is the say that any allowed value is as likely as any other), we simply say that any model which requires values of parameters outside of these does not fit the data. Some tests were carried out with relaxed priors; the ranges allowed for the height and width of of the two Gaussians that we fit can be allowed to be a little wider than those we show in Table 4.1 and still produce acceptable results, however the results are very sensitive to the ratios and differences between them. In these tests we therefore choose to restrict the priors to those in Table 4.1. The last component of our Bayesian formulation is the likelihood function.

4.6.3 The Likelihood Function

We choose to maximise the likelihood using a least squares likelihood. Our histogram data take a set of (x,y) values; the bin centres and the number of data in each bin, with some uncertainty. The model, for the same y values is some function of x, and our six model parameters $\mathbf{m}$. The number of data in the bins are large enough that we can invoke the central limit theorem and assume that our measured values x_i are distributed about the ideal model value $f(x_i, \mathbf{m})$ with Gaussian statistics. The uncertainty σ_i on the number of data per bin is taken as the standard Poisson error $\sqrt{y_i}$.

Taking all of this, the probability of measuring a particular number of values y_i in a bin centred on x_i given model parameters $\mathbf{m}$ is

$$P(y_i|\mathbf{m}) = \frac{1}{\sigma_i\sqrt{2\pi}} e^{-[y_i - f(x_i,\mathbf{m})]^2/2\sigma_i^2} \tag{4.11}$$

We work with the logarithm of the likelihood for convenience

$$-\frac{1}{2}\sum\left[\frac{y_i - f(x_i,\mathbf{m})}{\sigma_i}\right]^2 - \sum \ln\, \sigma_i\sqrt{2\pi} \tag{4.12}$$

so to maximise the likelihood, we use the emcee algorithm to maximise

$$-\frac{1}{2}\sum\left[\frac{y_i - f(x_i,\mathbf{m})}{\sigma_i}\right]^2 . \tag{4.13}$$

This is a simple least squares likelihood (e.g Barlow (1989)), where we seek the set of parameters $\mathbf{m}$ which minimises the weighted sum of the squared differences between the model $f(x_i, \mathbf{m})$ and the data.

The *emcee* algorithm (Foreman-Mackey et al., 2013) implements a Markov Chain Monte Carlo method. The algorithm samples the parameter space and provides estimates of the posterior distributions of the various parameters. This is used to quickly test many sets of model parameters and allows us to fully sample the allowed parameter space and derive both the most likely values for the parameters as well as their uncertainties. What follows are the results of using this method to estimate the binary fraction for both our simulated and real data.

4.7 Results

The results of this investigation naturally fall into two broad categories. In creating our simulated fields, we must be able to trust that they are an accurate representation of a real field, were it to have the same binary fraction, so we test the method described in above on the simulated fields, and ensure that the algorithm recovers their binary fraction. Only then do we implement the same method on the real stars to infer its binary fraction.

4.7.1 Recovering the Binary Fraction of Simulated Distributions.

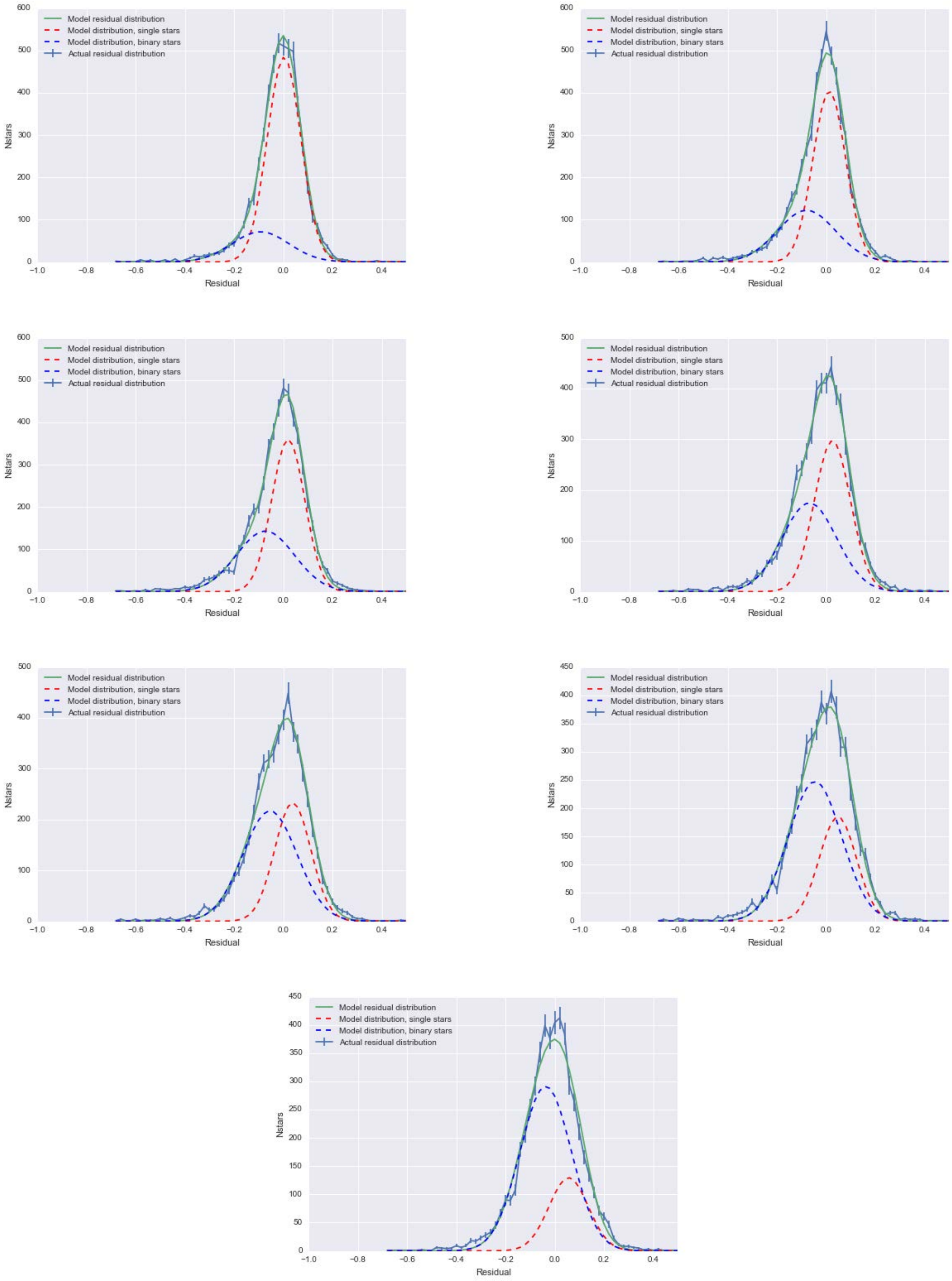


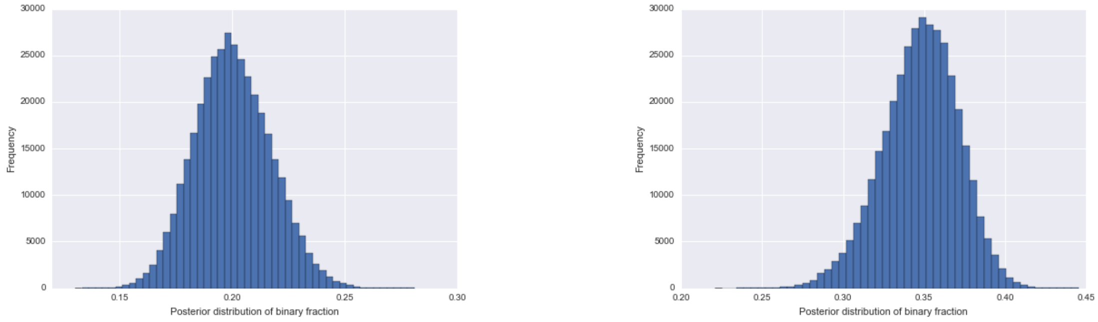
Figure 4.14: Optimum model for distributions of the simulated residuals of the binary fractions from 20% − 80% determined by *emcee*.

Figure 4.14 shows the optimum model determined by emcee for the distributions of the simulated data according to the model form and priors described in the above section. The data fitted by *emcee* are the same histograms as in Figure 4.6 with the $\sqrt{N}$ error bars shown on the count in each bin.

Since *emcee* samples the parameter spaces of the model, we obtain estimates of the posterior distribution of each parameter, including the binary fraction, as described by

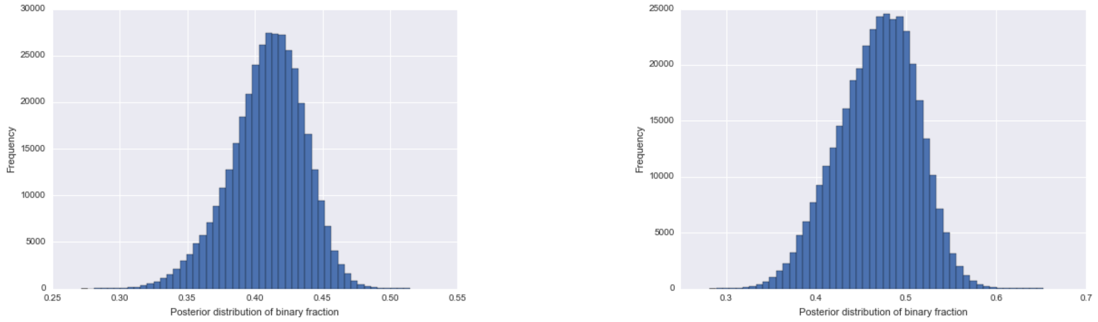
$$\frac{H_b \sigma_b}{H_s \sigma s} \tag{4.14}$$

and the posterior distributions of the binary fraction for each input simulation are shown in Figure 4.15. To ensure the parameter spaces are well sampled, *emcee* uses 500 random walkers taking 750 steps to sample each parameter space. Ideally, these posterior distributions are normally distributed.



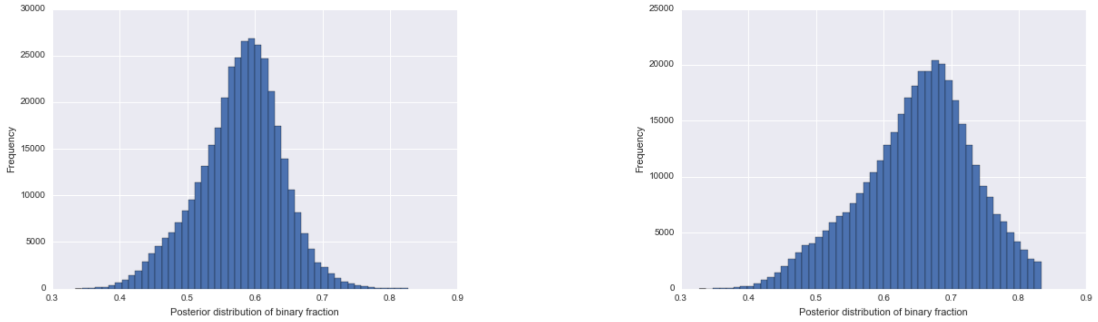
(a) 20%

(b) 30%



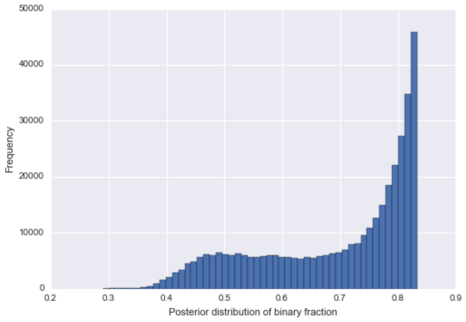
(c) 40%

(d) 50%



(e) 60%

(f) 70%



(g) 80%

Figure 4.15: Posterior distributions for the binary fraction of simulated populations with binary fractions from 20% – 80%

The posteriors shown in Figures 4.15f and 4.15g are not as convincing as the distributions for the other simulated fractions. The posterior distribution of the 70% binary fraction simulation is truncated at the high tail of the distribution. This implies that the priors are imposing a boundary on the sampling which is disallowing models which would in fact be realistic for the data, however the result of the sampling still produces a clear result on the fraction. The distribution for 80% is less convincing, however, being quite degenerate across a wide range of possible fractions. Despite this, the median recovered fractions do match the input values reasonably well, as we see below.

Table 4.2 shows the recovered binary fraction for each of the above simulated fields, and their agreement with the actual input fraction. The recovered fraction is the median value of the posterior distribution in Figure 4.15, and the uncertainty is the 68% confidence interval.

Actual fraction	Recovered Fraction	Uncertainty	Agreement
20	19.99	1.68	$< 1\sigma$
30	34.87	2.33	2.1σ
40	41.12	2.75	$< 1\sigma$
50	47.10	4.38	$< 1\sigma$
60	58.31	6.12	$< 1\sigma$
70	65.76	8.18	$< 1\sigma$
80	74.78	12.99	$< 1\sigma$

Table 4.2: Summarised results for simulated stars.

The agreement between the input and recovered binary fraction is generally good, all bar one lie within 1σ of the actual value. Figure 4.16 shows the comparison of these quantities more readily.

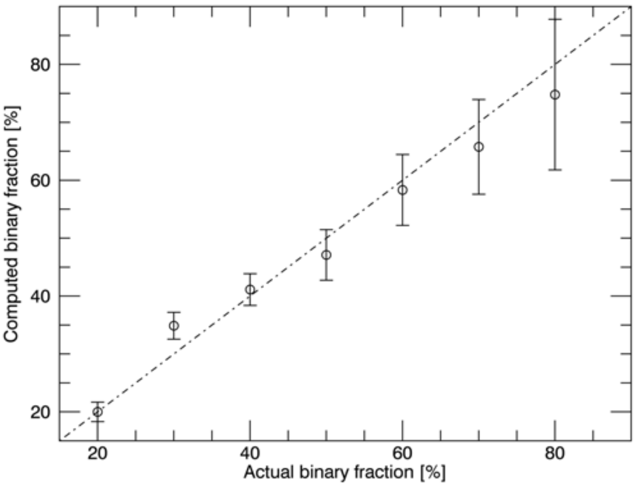


Figure 4.16: Input binary fraction vs emcee determined binary fraction for the simulated data. Error bars are the uncertainties for the 68% confidence interval and the dashed line shows $1-1$ agreement.

Now that we are convinced that the algorithm can recover the correct binary fraction reasonably well for the simulations, we can move on to applying the same method to the real stars.

4.7.2 Inferred Binary Fraction for ***Kepler*** Stars

Having used the simulated fractions to inform the model, priors and to ensure that the emcee routine recovers the true underlying fraction from the simulations reasonably reliably, it is then a case of running the routine with the same priors and assumptions on the real stars. Figures 4.18 and 4.18 show the most probable model for the distribution of the data, and the posterior distribution of the binary fraction.

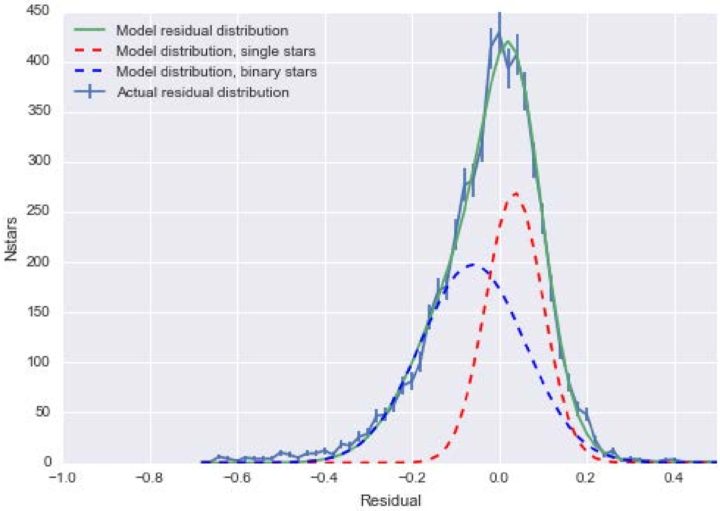


Figure 4.17: Optimum model distribution for residuals of *Kepler* stars.

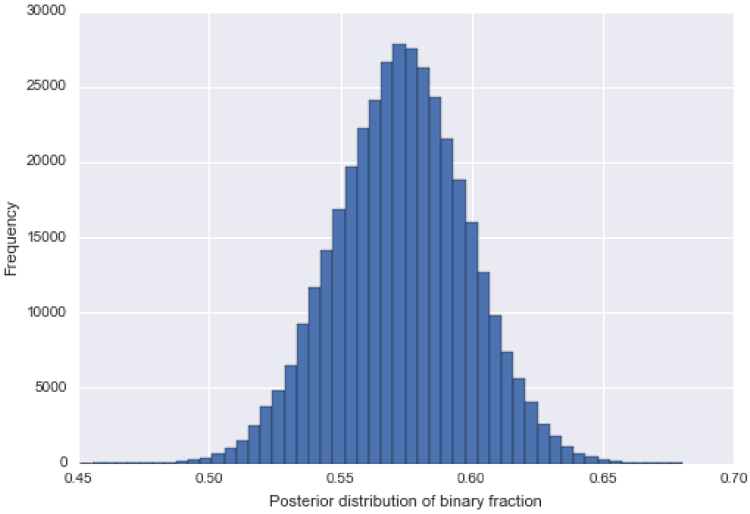
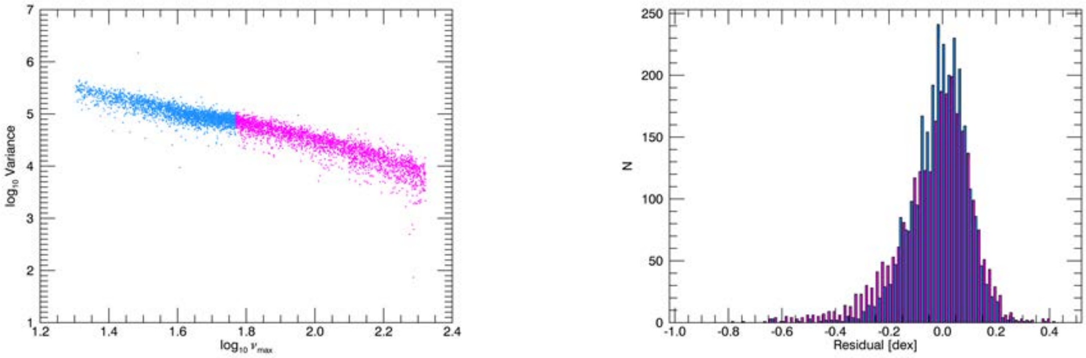


Figure 4.18: Binary fraction posterior distribution for *Kepler* stars.

The inferred binary fraction for this sample of *Kepler* red giants is 57.4% ± 2.5%. This assumes a uniform binary mass ratio distribution between 0.9 and 1.0, and no contamination from suppressed $\ell = 1$ modes or background stars.

ν_{max} Dependence

We can also determine whether there is any significant effect on the inferred fraction as the population evolves. ν_{max} can act as a rough evolutionary proxy, since we know that more evolved stars will tend to have a lower ν_{max}. Here we split the sample in half about the median ν_{max}, Figure 4.19 shows the variance and sample residuals color coded for each cohort. The distributions are different shapes, the higher ν_{max} cohort is generally broader and the lower ν_{max} cohort is more peaked.

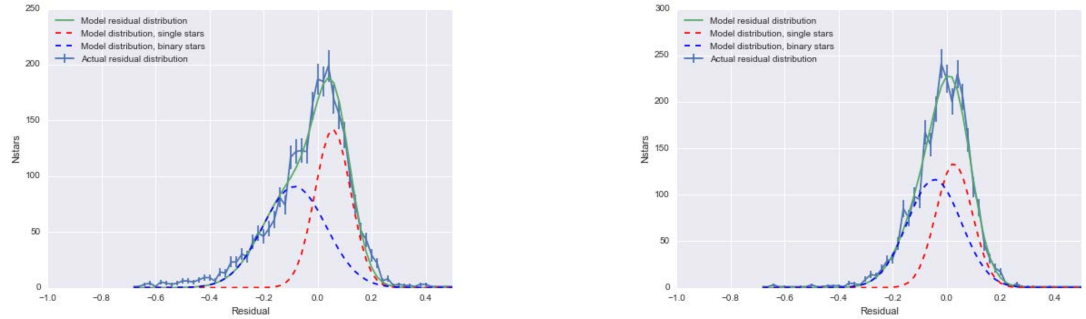


(a) Cut in ν_{max} about median.

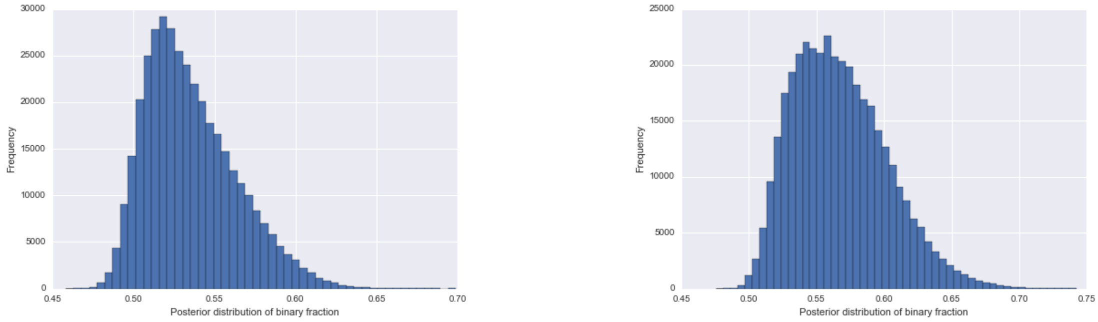
(b) Residuals for two cohorts from panel a.

Figure 4.19: The Variance-ν_{max} relationship and residuals colour coded according to ν_{max}. Blue points denote stars with a ν_{max} less than the median, and magenta points denote those with ν_{max} greater than the median.

In repeating the same procedure used for the whole sample; obtaining priors from the simulated fields with the same cut in ν_{max} in place and running the MCMC sampler, the binary fraction for each subset of stars is calculated, allowing us to test whether there is any change in the binary fraction as stars evolve up the red giant branch. Figure 4.20 shows the fitted distributions for the low and high ν_{max} groups, and the posterior distributions of the binary fraction for each.



(a) Model distribution, $\nu_{\max} >$ Median $\nu_{\max}$ (b) Model distribution, $\nu_{\max} <$ Median $\nu_{\max}$



(c) Binary fraction posterior distribution, $\nu_{\max} >$ Median $\nu_{\max}$ (d) Binary fraction posterior distribution, $\nu_{\max} <$ Median $\nu_{\max}$

Figure 4.20: Results for the sub-cohorts with the cut about the median $\nu_{\max}$.

The results for the two sub-populations dividing into lower and higher $\nu_{\max}$ are in broad agreement with each other and the fraction for the whole population, as shown in Figure 4.21.

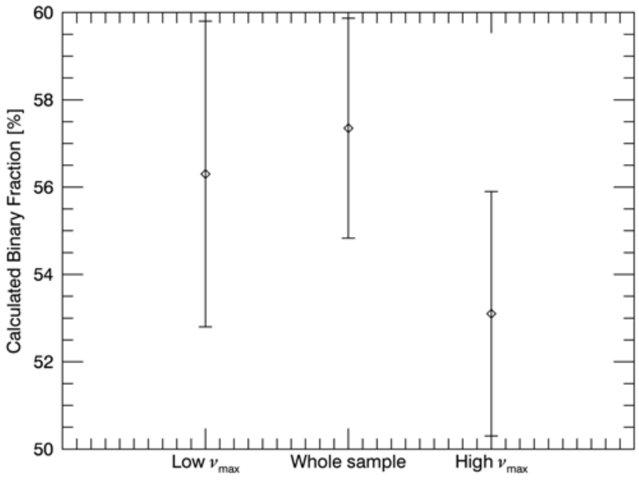
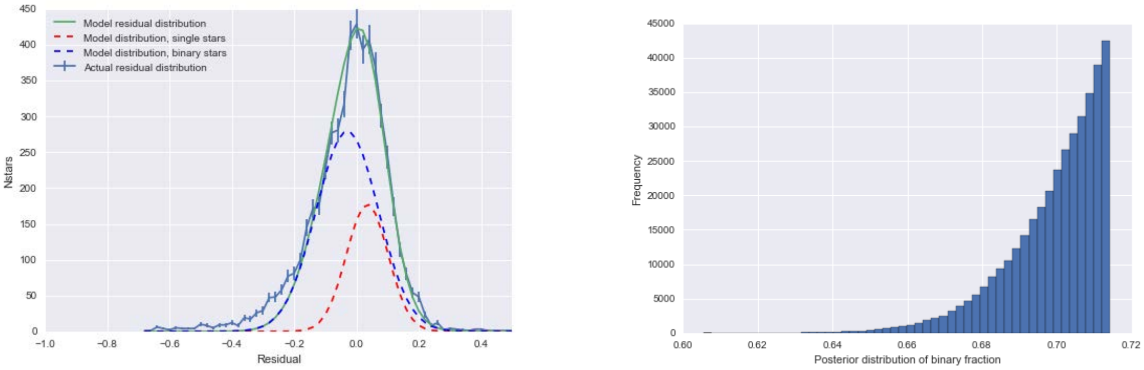


Figure 4.21: Binary fractions determined for real stars, for the whole sample and the lower and higher ν_{max} sub-cohorts.

A Note on The Binary Mass Ratio Distribution

As discussed in section 4.4.1 above, the mass ratio distribution of the binary population has its own impact on the distribution, namely on the average amount of washout experienced by a given star and hence the average offset between the two components of the distribution. This is a quantity that is outside of the scope of this work to seek to determine, however it may be possible to place some level of constraint on it. In obtaining priors from simulations with a binary mass ratio distribution uniform between $[0.7-1.0]$, we can apply these to the real stars to determine if this has any effect on the binary fraction we can infer, as the distribution from $[0.9-1.0]$ is an assumption we cannot be sure of.



(a) Model distribution.

(b) Binary mass ratio posterior.

Figure 4.22: .
]Model distribution and posterior distribution recovered binary fraction for real stars, assuming a binary mass ratio distribution uniform between [0.7-1.0].

Figure 4.22 shows the resulting model and binary fraction posterior distribution for the real stars using this assumption for the mass ratio distribution. Focusing first on the posterior distribution in Figure 4.22b, it is clear that the distribution is extremely truncated, the priors disallowing models which may actually be good fits to the data. This is also clear from the most likely model for the data as determined by emcee, shown in Figure 4.22a. While this is a subtle effect,notice that the best model allowed by the priors assuming a mass ratio between [0.7-1.0] does not capture the low tail of the distribution at all well. This may be interpreted to mean that to model this distribution, priors which allow for a wider separation between the single and binary distribution are needed, and hence suggesting that the [0.7-1.0] uniform mass ratio distribution is perhaps unrealistic for these stars, and that the mass ratio distribution must be closer to unity.

4.7.3 Conclusion

It has been suggested that a majority of main sequence stars are in multiple systems, while late type stars are far more likely to be in single star systems (eg. Lada (2006)). Most estimates of binary fraction come from spectroscopic surveys. These surveys tend to focus on finding the single star fraction, as they can detect triple and quadruple systems, which is outside of the scope of the simulations we use. In the study by Duquennoy and Mayor

(1991), a spectroscopic survey of F7-G9 main-sequence stars in the solar neighbourhood, once their various detection biases are accounted for, they find that 57% of systems have at least one stellar companion; in excellent agreement with our figure for red giants, though the assumptions about the binary mass ratios do differ.

Lada (2006) provides a discussion of the dependence on spectral type, concluding that across all spectral types, most stars are likely single, however this is due to M dwarfs being extremely common. Obtaining an estimated binary fraction that is broadly in line with estimates from other methods is encouraging, though it is important to keep in mind the limitations and assumptions of this study.

Having an estimate of the underlying binary fraction of the population is helpful in large scale surveys such as *Kepler*. Population models are routinely used in predicting and modelling the evolutionary, chemical and spatial properties of the Galaxy, and binary stars are excellent tests of physics, having formed at the same time and having the same chemical composition, so being able to estimate how many of these systems may be detectable with these surveys may guide the development of future research plans. Similarly for the exoplanet detection community, knowing the fraction of the population expected to be in binary systems may effect the predicted haul of exoplanets; while there are examples of circumbinary planets, the dynamical effect of a binary star system on an orbiting planet must be a consideration. Few detections of transiting exoplanets have yet been made around giant planets, and work is ongoing into the thresholds of detection for planets around giants (eg. North et al. (2017)). This is sensitive to the stellar noise, which is impacted by the presence of a companion star. Any predicted number of possible detections ought to include consideration of the underlying binary population, and the method described in this chapter is effective and not reliant on any other observation than the data used for transit detection.

Chapter 5

Detection of Oscillations and Determination of Seismic Parameters

The *CoRoT*, *Kepler* and *K2* missions have yielded asteroseismic detections in thousands of stars, and extracting the seismic information quickly, accurately and with minimal manual input has become a neccessity. In this chapter, an oscillation detection method relying on the now familiar variance metric is outlined, along with several methods of estimating ν_{max} and $\Delta\nu$ and their uncertainties. Results for red giant stars will be presented on data obtained by K2 and CoRoT as well as simulated power spectra created following the method in Chapter 3.

5.1 Oscillation Detection

Recall from Chapter 3 that predictions of the oscillation power present in a star of a given ν_{max} are relatively trivial to make. In combination with a good estimate of the background signal from the median filter described later in this chapter to estimate the global signal-to-noise, it is possible to implement robust statistical tests to search for the signatures of oscillations.

5.1.1 Odds Ratio Test

We begin with a basic false alarm test. The frequency range in which most of the mode power is contained is defined by $\nu_{\mathrm{max}} \pm \delta\nu_{\mathrm{env}}$ as discussed in Chapter 3. The global signal-to-noise $\mathrm{SNR_{tot}}$ is the ratio of the total mode power to the total background power in the mode envelope range. This $\mathrm{SNR_{tot}}$ can be tested against χ^2 $2\mathrm{N_s}$ - d.o.f statistics, meaning we assume that the observed spectrum is distributed about the limit spectrum (discussed in Chapter 3) with negative exponential statistics. There are $\mathrm{N_s}$ bins across the frequency range which contributes to $\mathrm{SNR_{tot}}$. The false-alarm probability, or H0, is computed, following Basu and Chaplin (in the press) according to:

$$\mathrm{P(SNR(\nu_s), N_s \mid H0)} = \int_x^\infty \frac{\mathrm{exp(x')}}{\gamma(\mathrm{N_s})} \mathrm{x'^{(N_s-1)} dx'} \tag{5.1}$$

where $\mathrm{x'} = 1 + \mathrm{SNR}(\nu_\mathrm{s})$. A low H0 value implies a implies a low probability that any potential signal is a statistical fluctuation, and is therefore more likely to be an actual power excess rather than a noise fluctuation. However, oscillations are not the only possible source of power excess, artefacts due to transits or instrumental effects can have very high power spectral density above the background and thus suggest a detection according to the H0 hypothesis. We can however improve on this. Since we can make a good prediction of the signal-to-noise expected from a solar-like oscillator, knowing its apparent magnitude (and hence expected shot noise power) as well as the other components of the power spectrum, we can compare this to the SNR we observe in our power spectrum. The H1 probability decides whether the observed SNR is consistent with the expected SNR:

$$\mathrm{P(SNR(\nu_s), SNR_{pred}, N_s \mid H1)} = \int_x^\infty \frac{\mathrm{exp(x')}}{\gamma(\mathrm{N_s})} \mathrm{x'^{(N_s-1)} dx'} \tag{5.2}$$

where $\mathrm{x'} = (1 + \mathrm{SNR}(\nu_\mathrm{s}))/(1 + \mathrm{SNR_{pred}})$. Combining this with the H0 test, we can compute the posterior probability of any power excess being due to modes, including our prior knowledge of the expected variance and SNR, thus:

$$\mathrm{P}(\mathrm{H1} \mid \mathrm{SNR}(\nu_{\mathrm{s}}), \mathrm{SNR}_{\mathrm{pred}}, \mathrm{N_s}) = \left[1 + \frac{\mathrm{P}(\mathrm{SNR}(\nu_{\mathrm{s}}), \mathrm{N_s} \mid \mathrm{H0})}{\mathrm{P}(\mathrm{SNR}(\nu_{\mathrm{s}}), \mathrm{SNR}_{\mathrm{pred}}, \mathrm{N_s} \mid \mathrm{H1})}\right]^{-1}. \tag{5.3}$$

Providing that the estimate of the background power is good, this is a robust test for oscillation power. In testing for oscillations, it is helpful to think of running a window through the power spectrum. For each frequency ν_s, we have an associated width of the window $\delta\nu_{\mathrm{env}}$; this is the FWHM of the mode envelope centred on ν_s

$$\delta\nu_{\mathrm{env}} = 0.66\nu_s^{0.88}, \tag{5.4}$$

and the observed SNR for this range is described by

$$\mathrm{SNR_{obs}} = \frac{\sum\limits_{\nu_{\max}-\delta\nu_{\mathrm{env}}}^{\nu_{\max}-\delta\nu_{\mathrm{env}}} P(\nu)}{N_{\mathrm{bins}}} - 1.0 \tag{5.5}$$

where P(ν) is the median-normalised spectrum and N_{bins} is the number of bins in the frequency window. Our predicted SNR comes from expressions given in Chapter 3 for the smoothed Gaussian we can use to represent the mode power, an expression for the granulation background, and the white noise measure near the Nyquist frequency. The predicted SNR is simply the ratio of integrated predicted mode power

$$\mathrm{P_{modes}} = \sum_{\nu_{\max}-\delta\nu_{\mathrm{env}}}^{\nu_{\max}-\delta\nu_{\mathrm{env}}} 2 \times 10^7 \nu_{\mathrm{s}} e^{\dfrac{-(\nu-\nu_s)^2}{2\sigma^2}} \tag{5.6}$$

to the that of the background, made up of the granulation:

$$\mathrm{P_{granulation}} = \sum_{\nu_{\max}-\delta\nu_{\mathrm{env}}}^{\nu_{\max}-\delta\nu_{\mathrm{env}}} \frac{2.475 \times (3090.0/{\nu_s}^2)}{(1 + (2\pi\tau\nu)^2)} \tag{5.7}$$

and the white noise component, taken as the mean value of the power between 270μHz and 280μHz. Then

$$\mathrm{SNR}_{\mathrm{pred}} = \frac{P_{\mathrm{modes}}}{P_{\mathrm{granulation}} + N_{\mathrm{bins}} P_{\mathrm{noise}}} \tag{5.8}$$

where ν_s is the centre of the window being searched for oscillations. We can compute the predicted and observed SNR for each frequency step we increase ν_s by. In repeating this for the whole frequency range, the probability computed in equation 5.3 can be visualised as shown in the lower panels of Figure 5.1.

Where this test statistic has a value close to unity is a clear indicator of the region of oscillations. This automated test provides a quick and robust way to determine whether oscillations are present, and also provides constraints on the regions of frequency in which the pipeline will search for the values of ν_{max} and $\Delta\nu$. Since this relies on an accurate determination of the observed signal to noise ratio, we must be sure that our background estimation is good in order to preserve the mode signal and hence an accurate signal to noise.

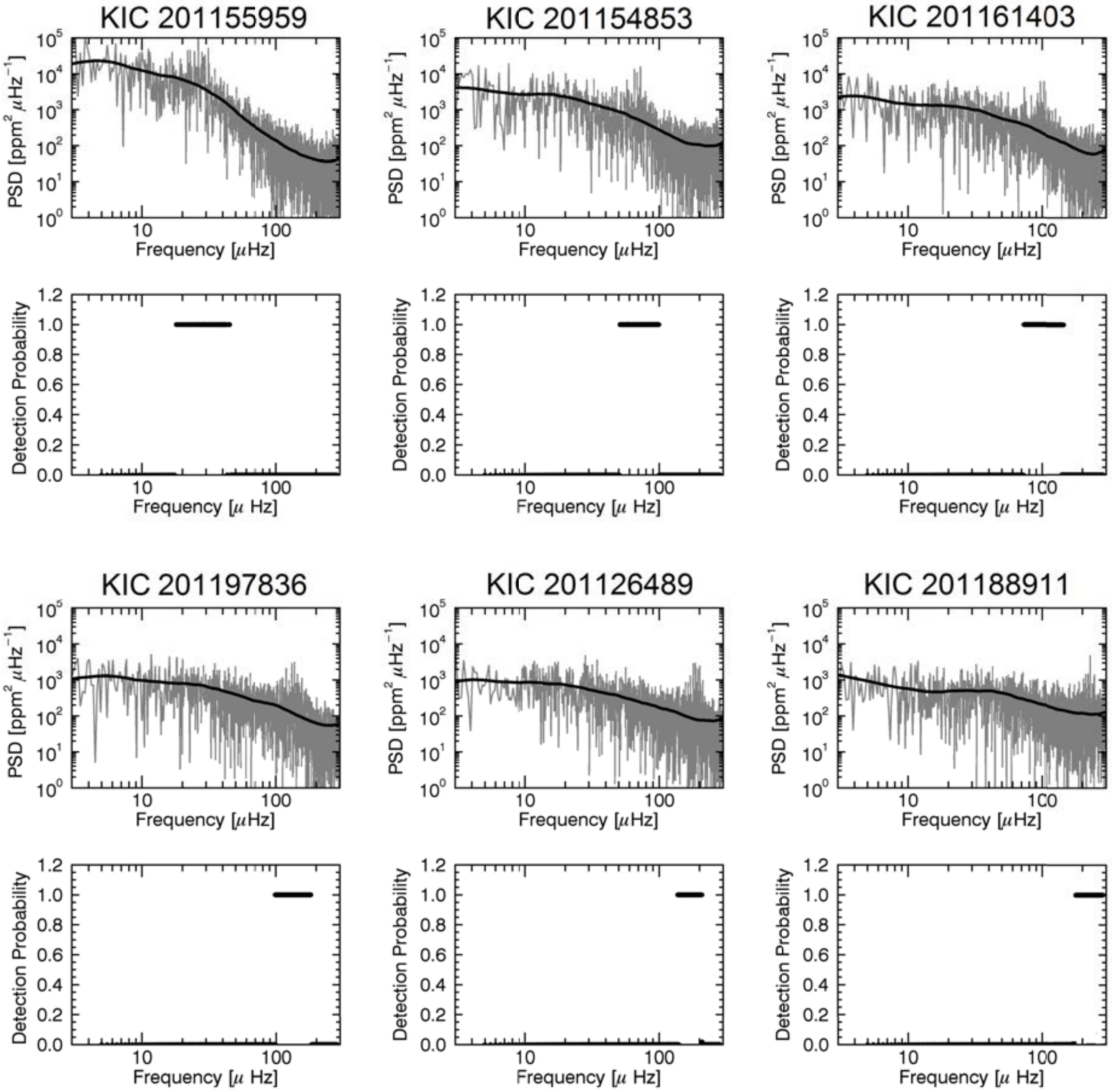


Figure 5.1: Upper panels show the power spectra of our six K2 stars with the background profiles in black. The lower panels show the detection statistic which very clearly shows the region where the oscillations lie.

5.2 Background Estimation

While the obvious approach may be to fit the power spectrum in order to extract the background, for this quick automated test we would prefer a quicker, less computationally expensive option. It is crucial that the method chosen successfully separates the modes from the background.

5.2.1 Moving-Median Filter

An inexpensive way to estimate the background signal is using a moving median filter of a width such that the mode signal is essentially removed, allowing for the mode power and background signal to be separated, and estimates of the global signal-to-noise ratio that are required for the above detection test to be made simply. The moving-median method to find the mean background power is preferable to a moving-mean filter, so it is relatively unaffected by the high peaks in the power spectrum over the region of frequency containing the oscillation power. The mean power can be estimated from the moving median with the simple normalisation

$$B_{\mathrm{mean}} = \frac{B_{\mathrm{median}}}{\ln 2} \tag{5.9}$$

assuming that the data are distributed with χ^2 2-degrees of freedom. This filter is applied in log-log space, and interpolated back to the linear. Figure 5.2 shows the background obtained from this moving median filter for 6 K2 stars with oscillations across the *Kepler* long cadence frequency range.

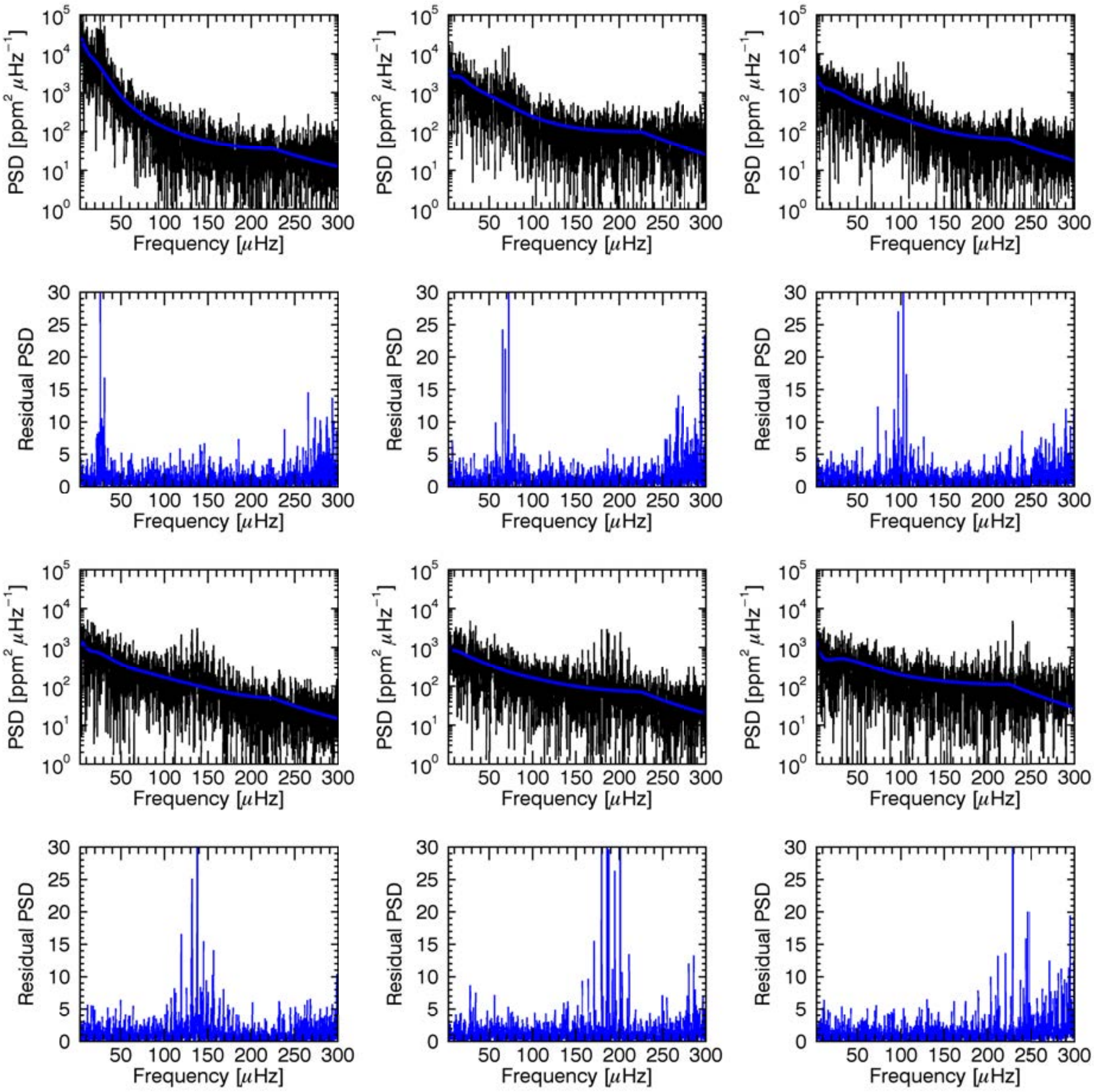


Figure 5.2: Upper panels show the power spectra of the same six K2 stars as in Figure 5.1, with the background profiles obtained by the fixed-width moving median filter, shown in blue. Lower panels are the median-normalised power spectra for these same stars. The power excesses due to the oscillation modes are clearly visible, however note the spurious excess at high frequency for the fixed-width filter, shown in blue.

This particular filter has a width of 0.4 dex, and suffers from an issue at high frequency when it runs out of bins to perform the moving-median. This is problematic for the detection step described above, since the incorrect background leaves a region of high signal-to-noise close to the Nyquist frequency, and this in turn impacts the test statistic, and gives an incorrect estimate for where we ought to search for $\nu_{\rm max}$ and $\Delta\nu$. Figure 5.3

shows the test statistic for the background profile as determined using the fixed width filter. While the oscillations of this star are in reality centred at around 30 μHz, there also appears another region where the statistic is unity, close to the Nyquist frequency. This is a false detection due to the poor background approximation made using the fixed width filter.

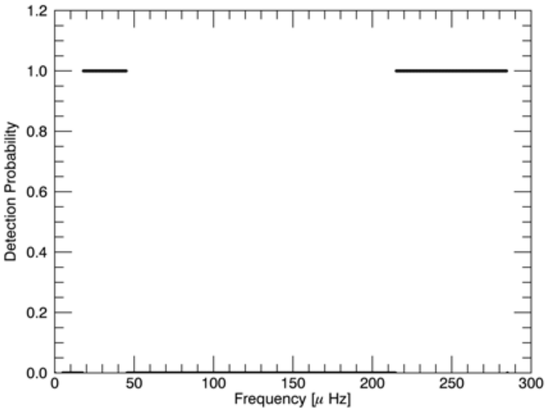


Figure 5.3: The same statistic as shown in the lower panels of Figure 5.1 using the fixed width moving-median filter. The region of high frequency where this statistic is unity is due to a poor background approximation, as shown in Figure 5.2

Clearly, this causes false-positive detections, and incorrect identification of the frequency of the oscillations. This can be avoided however.

5.2.2 Variable Width Filter

An improvement to this filter is to decrease the filter width with increasing frequency. This minimises any rise in filtered power due to the presence of the modes, as well as succeeding in obtaining a good background estimate at high frequency. The form of this filter width with frequency is fairly arbitrary, by testing different fixed filter widths on power spectra of stars with a range of frequencies and requiring there to be a smooth change with frequency, we arrive at the form shown in orange in Figure 5.4, which proves to work well.

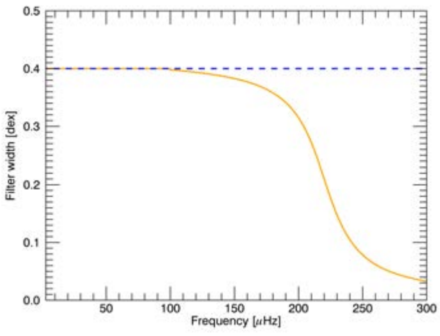


Figure 5.4: Filter widths in dex across the K2 frequency range. The blue profile shows the fixed-width filter, and the orange the variable-width filter. The form of this is fairly arbitrary, but performs well for stars exhibiting oscillations across the entire range.

Figure 5.5 shows the resulting background profile for our six *K2* stars using the variable width moving median filter. The lower panels show the median-normalised power spectrum for both of these filters. Now, this residual power spectrum contains the oscillation power excess above the background with minimal reduction of the power due to the filter, and the remaining background power consists of only the random noise component (Anderson et al. (1990), see Chapter 3) outside of frequency region of the oscillations. The excess at high frequency resulting from the fixed-width filter is now removed, in contrast to the fixed width filter case shown in Figure 5.2.

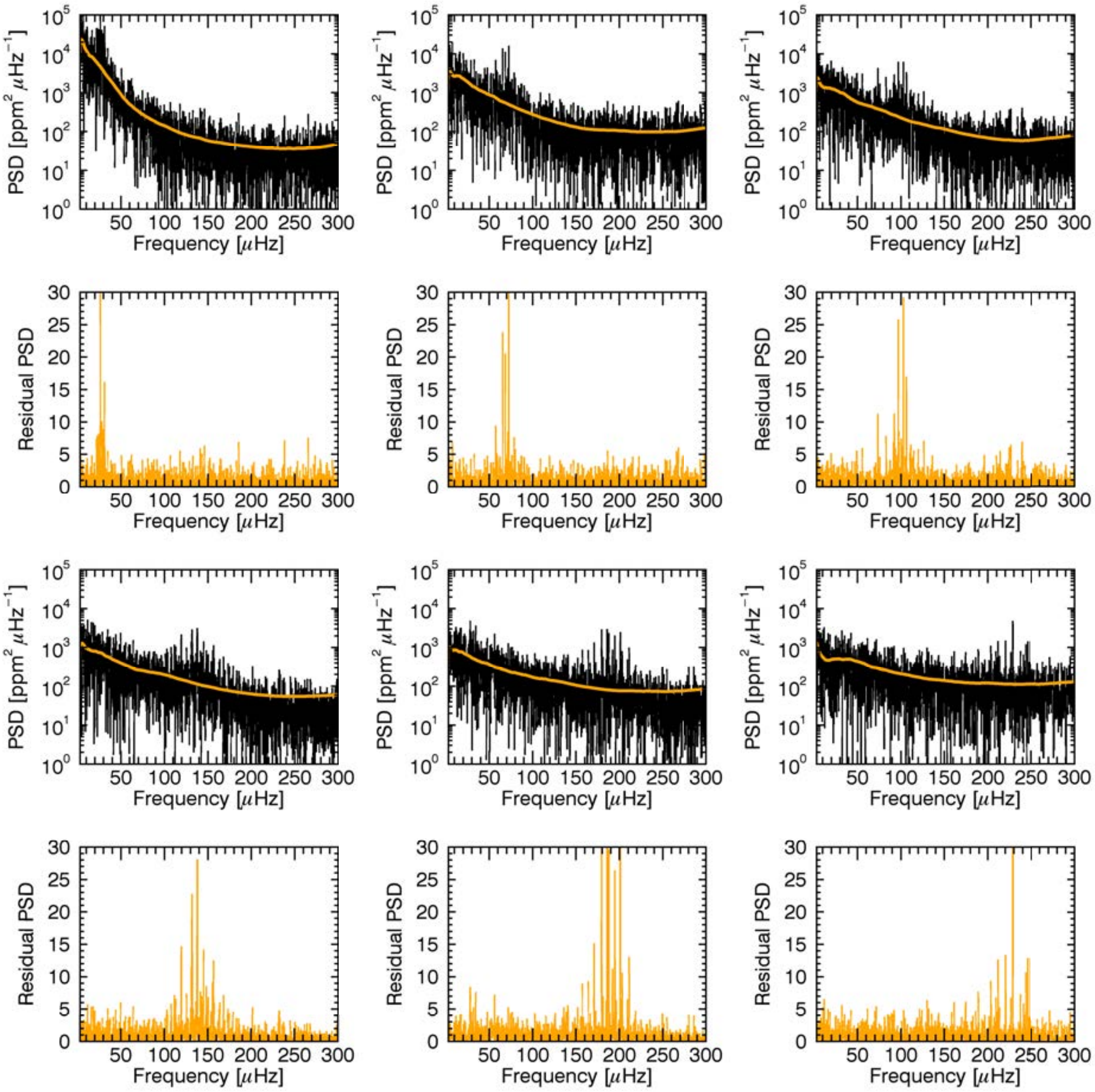


Figure 5.5: Upper panels show the same six K2 stars as in Figure 5.1, with the background profiles obtained by the variable-width moving median filter, shown in orange. Lower panels are the median-normalised power spectra for these same stars. The power excesses due to the oscillation modes are clearly visible.

With an identified region of frequency we now trust to contain the oscillations, we can go on to determine some global seismic parameters.

5.3 $\Delta\nu$ Estimation

One of the basic asteroseismic quantities we are now familiar with is the average large separation $\Delta\nu$, and extracting this from observations is essential in determining the mass and radius of the star. This section will describe two methods implemented in this pipeline for determining $\Delta\nu$ with uncertainties.

5.3.1 PSPS

We are accustomed to the power spectrum showing how power is distributed across frequencies, and we can take the power spectrum of the power spectrum (referred to as PSPS here) to find the frequency spacings between repeated features in the power spectrum. As we know, $\Delta\nu$ is the average spacing between modes of the same degree ℓ and consecutive order n. The dominant spacing in the power spectrum of solar-like oscillators is in fact $\Delta\nu/2$, since $\ell = 1$ modes lie approximately halfway between the frequencies of the $\ell = 0$ modes.

Since our detection test statistic not only determines whether or not there is a signature of oscillations present, but also tells us where in frequency we expect the oscillations to lie, we can use this to constrain the frequency range we search for $\Delta\nu$. We define some frequency ν_{cen}, as the frequency at the median bin where the test statistic is greater than 0.95. This ν_{cen} acts as our rough estimate for $\nu_{\max}$, and we then compute the PSPS over the frequency range centred on ν_{cen}, which is defined as

$$\nu_{\text{cen}} \pm \delta\nu_{\text{env}} \tag{5.10}$$

where

$$\delta\nu_{\text{env}} = 0.66\nu_{\text{cen}}^{0.88}. \tag{5.11}$$

following Mosser et al. (2012a). For this portion of the power spectrum, we compute an oversampled PSPS and choose a small region over which to search for the peak at $2/\Delta\nu$

based on the expected $\Delta\nu$ for the centre of the region of frequency, following the Mosser et al. (2012a) definition

$$\Delta\nu_{\text{guess}} = 0.276\nu_{\text{mid}}^{0.751}. \tag{5.12}$$

Searching within the range of 0.8-1.2 $\times 2/\Delta\nu_{\text{guess}}$, the frequency at which the PSPS is greatest is marked and we can then estimate $\Delta\nu$. Figure 5.6 shows the result of applying this method to the same six K2 stars as previous plots. It is clear that there are other peaks within the PSPS, owing to the presence of other repeating features in the power spectrum, for example the $\Delta\nu$ spacing between consecutive modes of the same degree ℓ, the spacings between higher order modes and to a lesser extent, any mixed mode signatures. However, the $2/\Delta\nu$ spacing dominates.

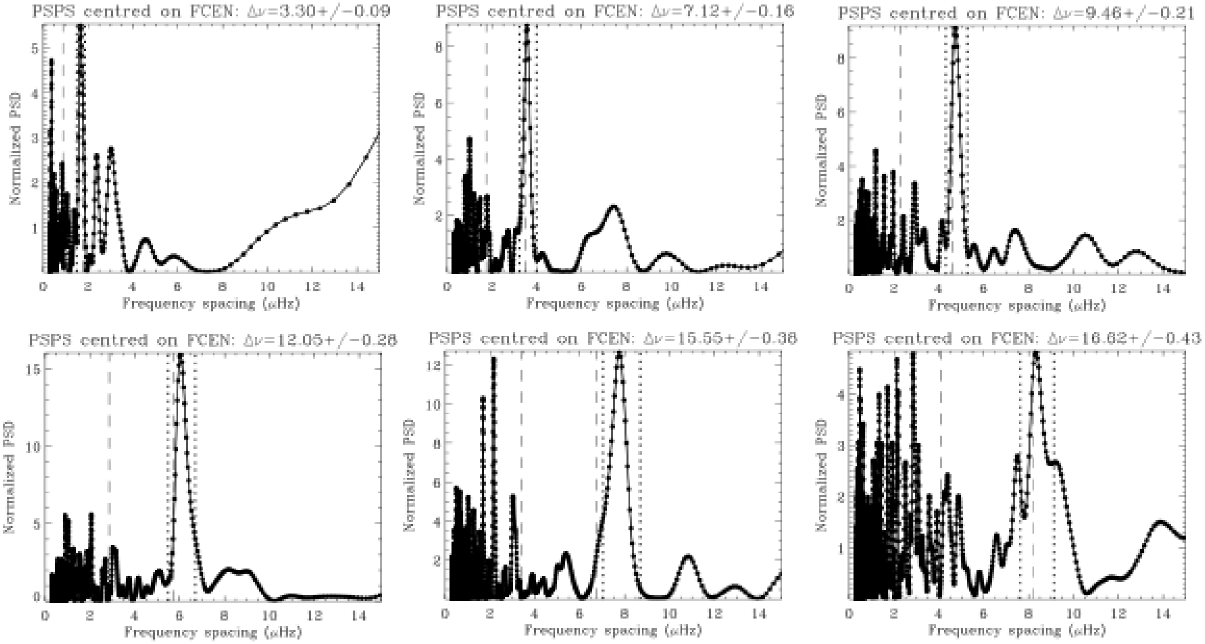


Figure 5.6: The PSPS of the six example stars. The dashed lines show the range over which we search for the maximum PSPS.

In obtaining the uncertainties on this value, the error on calculating the centroid of the peak is used. This is essentially the standard error on the mean;

$$SE_{\Delta\nu} = \sqrt{\frac{1}{N}\sum\left(\nu - \frac{\sum \nu_i PSD_i}{\sum I_i}\right)^2} \tag{5.13}$$

Another option for using the PSPS to provide our determination of $\Delta\nu$ is to fit the above peak as a Gaussian and take the centre of this to be the estimate of $\Delta\nu/2$. In practice this agrees extremely well with the estimate given by the maximum PSPS.

5.3.2 The “Red or Blue” Mask Method

Aside from methods based on using the PSPS to search for the frequency intervals of repeated features, we can search for peaks due to indivdual modes directly in the residual power spectrum based on some prediction of where we expect them to be. Familiar as we now are with the structure of a solar-like oscillation spectrum, we know that $\ell = 0$ modes, centred on $\nu_{\rm max}$ are separated by $\Delta\nu$, and halfway between each of these modes lies a region of $\ell = 1$ modes. We can use this to identify frequencies of peaks that lie close to where we expect.

Relying again on the identification of $\nu_{\rm cen}$ from the detection test, we can define a range of frequency in which to search for peaks. It is reasonable to search over the same range over which we define our variance metric and signal to noise measures, defined by a central frequency, and the expected full width at half maximum of the Gaussian envelope modulating the modes

$$\nu_{\rm cen} \pm \delta\nu_{\rm env} \tag{5.14}$$

Over this range, we identify the highest peak, and assume that it is an oscillation mode. We then anchor our priors; our predictions for where the modes ought to be, to the location of this peak. We would expect to find further modes at every $0.5\Delta\nu$ around the location of the maximum, corresponding to the locations of the $\ell = 0$ and $\ell = 1$ modes. We will in fact consider two separate sets of priors, one for the $\ell = 0$ regions and another for the $\ell = 1$ regions, both within $\pm 4\Delta\nu$ of the maximum.

The first set of priors are centred on the frequencies n-times-$\Delta\nu$ away from the highest peak, and the second set are offset by $0.5\Delta\nu$. Which one corresponds to modes of $\ell = 0$ of $\ell = 1$ is not important at this point. These priors inform where we expect to detect the modes, and take a Gaussian form with a width of a small fraction (here 0.2) of $\Delta\nu$

$$P_{\text{prior},i} = e^{\dfrac{-(\nu_{i\Delta\nu} - \nu)^2}{0.4\Delta\nu^2}} \tag{5.15}$$

and these are capped at unity. The false alarm probability depends on the number of bins being considered. Here we follow Basu and Chaplin (in the press) and decide it is reasonable to search for modes in $\Delta\nu$-sized ranges of frequency, since we expect to be able to detect several modes in a range this size. Therefore we consider the false alarm probability given we are searching $N = \Delta\nu T$ bins, where T is the length of the observation. The false alarm probability is, as in Equation 5.1, tested against χ^2 2-degrees-of-freedom statistics, that is to say that for a spike with power $P(\nu)$ relative to a mean background level $\langle P(\nu)\rangle$,the probability of detecting that spike by chance is

$$p(s_\nu) = e^{-s_\nu} \tag{5.16}$$

where

$$s_\nu = \frac{P(\nu)}{\langle P(\nu)\rangle}. \tag{5.17}$$

The probability we do not detect a spike in any one bin is $1 - p(s_\nu)$, and, as we are searching N bins, the probability we do not detect any such spike in any of our N bins is simply $[1 - p(s_\nu)]^N$, and therefore our chance of observing any such spike in any of our N bins, and our false alarm probability is given by

$$P_{\text{false}} = 1 - [1 - p(s_\nu)]^N. \tag{5.18}$$

Putting all of this together, the criterion for detecting a mode is simply

$$P_{\mathrm{prior}} P_{\mathrm{false}} \geq P_{\mathrm{thresh}} \tag{5.19}$$

that the product of the prior and the false alarm probabilities meet some threshold, here we choose 0.8. This threshold was chosen to allow the detection of marginal cases, which are more common in the lower signal-to-noise datasets of K2 and CoRoT as compared to *Kepler*, while not producing an unacceptable number of unreliable detections.

For ease of reference, the two sets of priors are treated separately and referred to as the "red" and "blue" masks. Figures 5.7 and 5.8 show, for the same six stars, the locations of the centres of the two sets of priors, and the modes identified by this method.

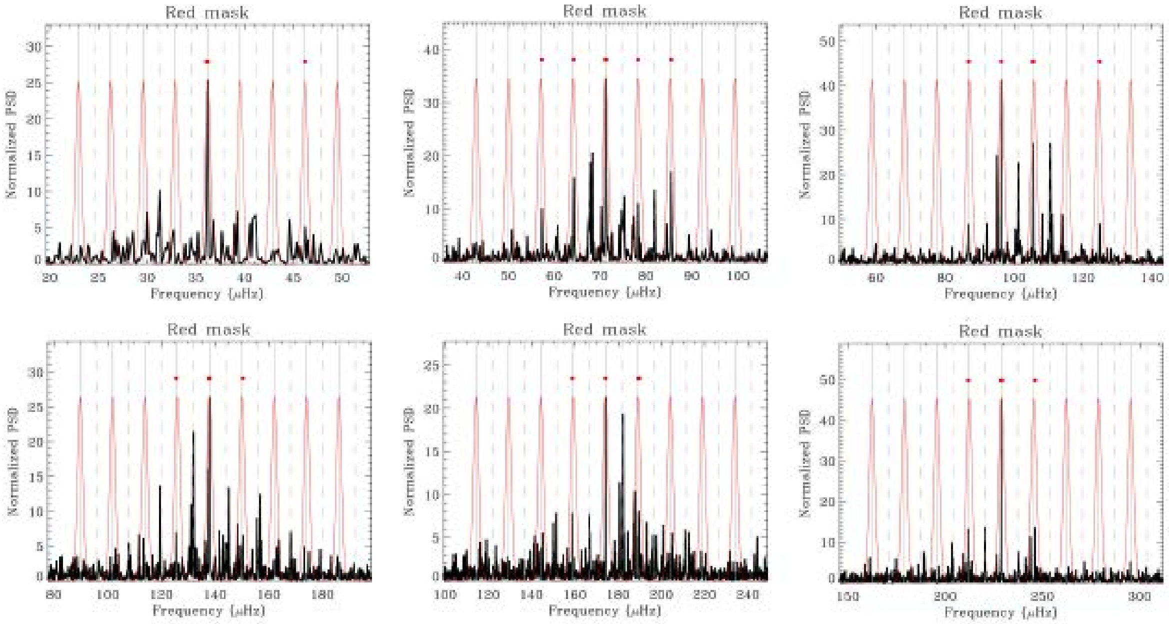


Figure 5.7: The first set of priors, centred on frequencies $\pm i\Delta\nu$ relative to the maximum peak. The black lines show these frequencies, and the priors themselves are shown in red. Peaks which pass the false alarm test are marked by red dots.

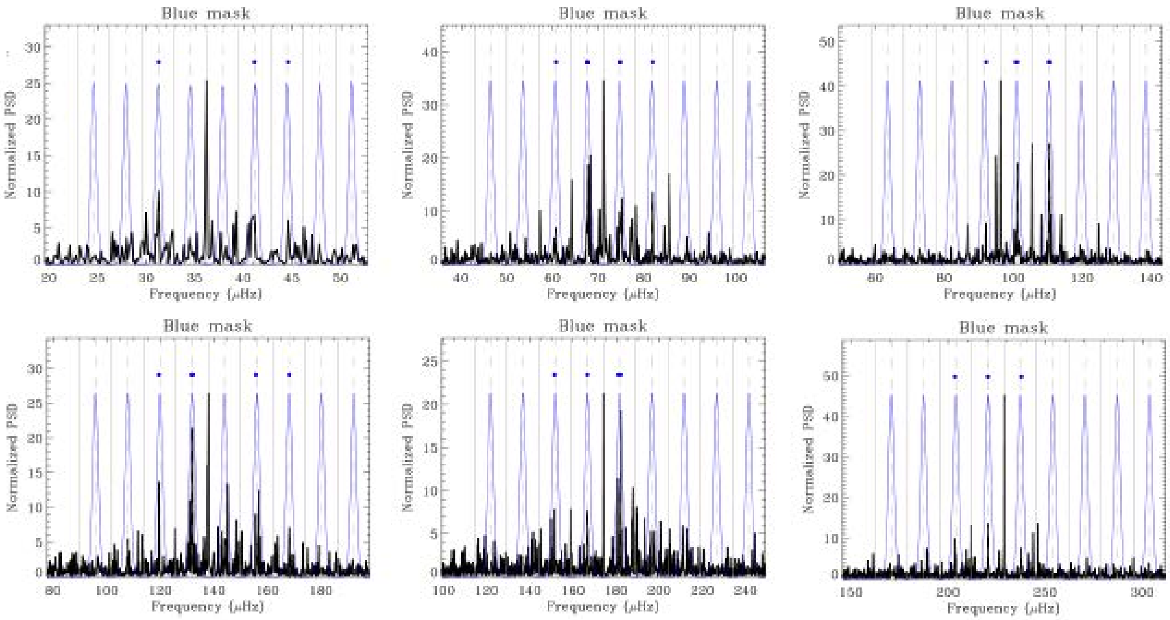


Figure 5.8: As in Figure 5.7, but with priors centred on the dashed lines lying at $\pm i\Delta\nu + 0.5\Delta\nu$ relative to the maximum peak.

To get our estimate of $\Delta\nu$ from this, we first need to mark the relative order n of these frequencies. Following Kallinger et al. (2012), we recall the asymptotic relation for the p-modes

$$\nu_{\rm nl} = \Delta\nu\left(n + \frac{l}{2} + \epsilon\right) + \delta\nu_{ll+2}(n). \tag{5.20}$$

and that for radial modes we have

$$\nu_{n0} = \Delta\nu(n + \epsilon) \tag{5.21}$$

i.e., modes are at frequencies modulo the large separation $\Delta\nu$, with exact locations fixed by the constant ϵ. It therefore follows that we may estimate ϵ using

$$\epsilon = \frac{\nu_{\rm nl}}{\Delta\nu} \text{mod } 1 \tag{5.22}$$

In practice, for consistency with solar-type stars we need to add an extra factor of unity to get the correct match to the order n if $\Delta\nu > 3\mu$Hz and $\epsilon < 0.5$. Having an estimated epsilon we can derive a relative order n

$$n = \frac{\nu_{\rm nl} - \epsilon\Delta\nu}{\Delta\nu}. \tag{5.23}$$

This is not the actual order n, here it is unimportant, we only need the order of modes relative to each other. Then our new estimate of $\Delta\nu$ comes from a simple linear fit between the relative order and frequency, where the gradient of this line is $\Delta\nu$. Figure 5.9 shows the result of this for our six K2 stars.

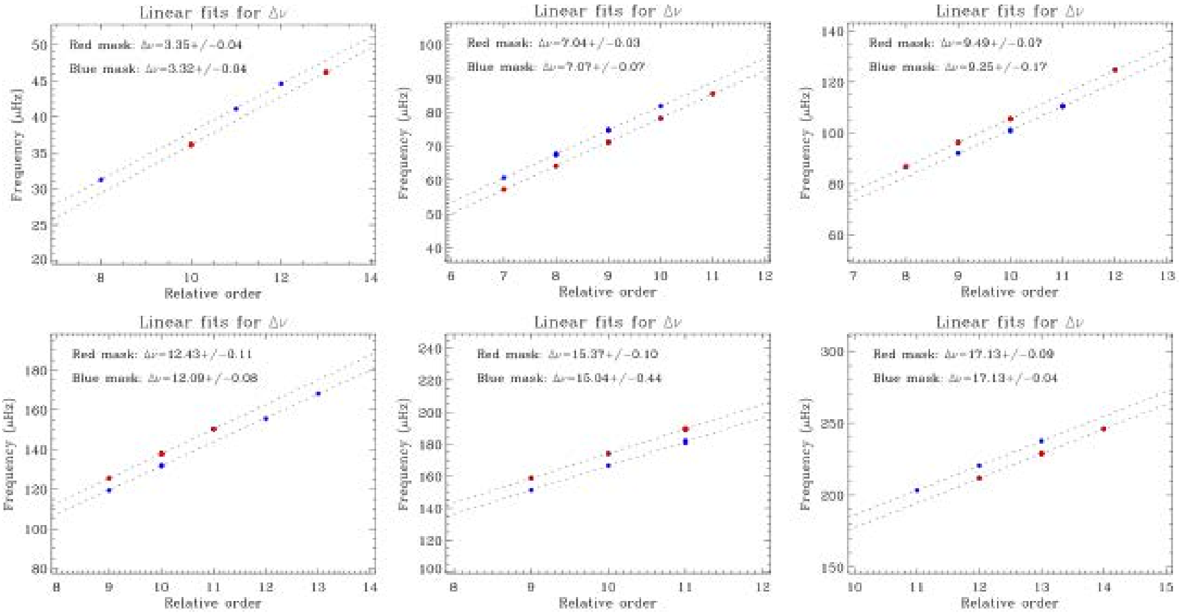


Figure 5.9: The result of the linear fit to the frequencies and relative order for the two sets of priors.

The uncertainty on the value of $\Delta\nu$ simply comes from the uncertainty on the fit.

5.4 $\nu_{\rm max}$ Estimation

We are by now familiar with the modes of oscillation being modulated by a Gaussian envelope, so it is fairly intuitive, when attempting to find the frequency that the modes are at their maximum power, to fit a Gaussian to the power excess in the relevant frequency range. We fit to the median-normalised power spectrum here, as, given our background approximation is good, it represents the power excess due to the modes well, and simplifies the fitting process owing to not having to consider a frequency dependent background component. We fit a function of the form

$$P(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} + k_{\text{offset}} \tag{5.24}$$

allowing for some flat background k_{offset}. The median-normalised power spectrum data are binned according to a few overlapping averages per $\Delta\nu$. This reduces the random noise while retaining the form of the peak we are trying to fit. The error bars on the fitted points shown in Figure 5.10 are simply the uncertainties derived from the binned data, and are included as weights for the fit.

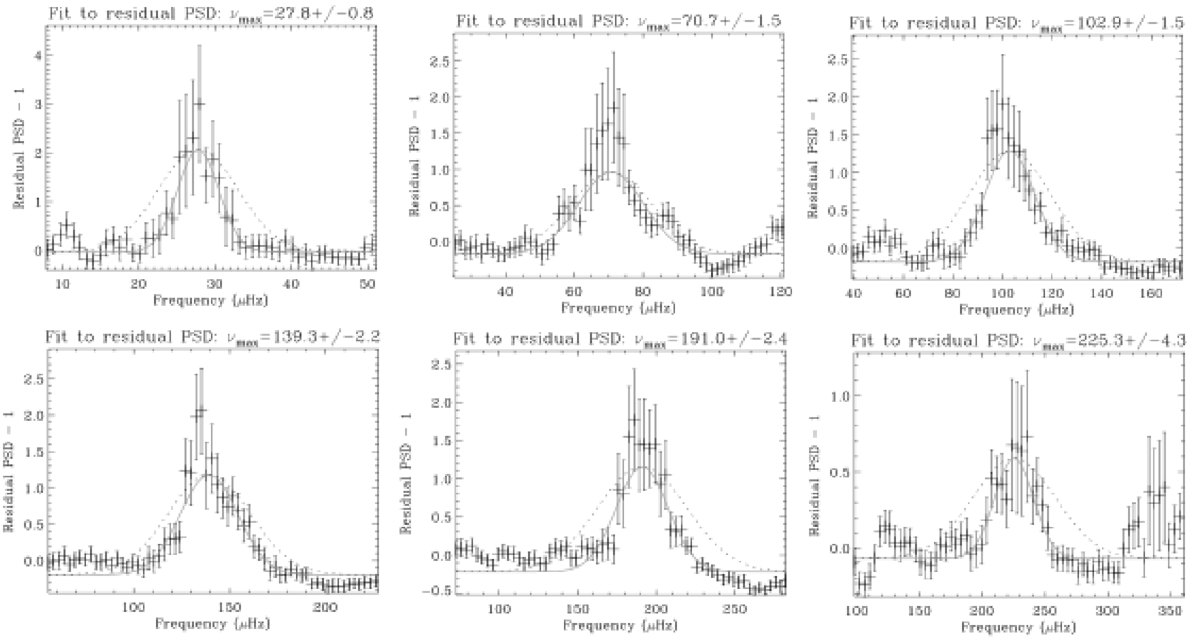


Figure 5.10: ν_{max} determined using the PSPS method for the six stars.

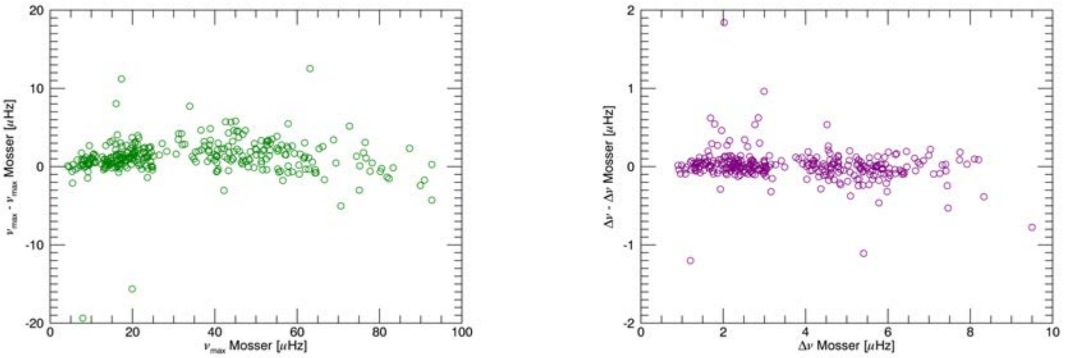
The location of the peak of the best fit Gaussian provides our estimate of ν_{max}. Figure 5.10 shows the result of this for our six representative K2 stars.

5.5 Results

Here, results on simulated spectra and data from CoRoT and K2 are presented. The CoRoT and K2 results are compared to values obtained by other pipelines, while for the simulations the correct values exist. Discussion of the uncertainties and the success in recovering the parameters accurately follow.

5.5.1 *CoRoT*

The CoRoT data used here are from the long cadence observations of the LRc01 exofield. All are red giants, and this list contains 345 stars. In the case of CoRoT, only the frequency range $< 100\mu$Hz is tested for signatures of oscillation, due to the presence of strong instrumental artefacts at higher frequencies due to the spacecraft passing through the South Atlantic Anomaly (see Chapter 2.) The results are are compared to values obtained by Mosser (private communication), see Mosser and Appourchaux (2009) for discussion of this pipeline.

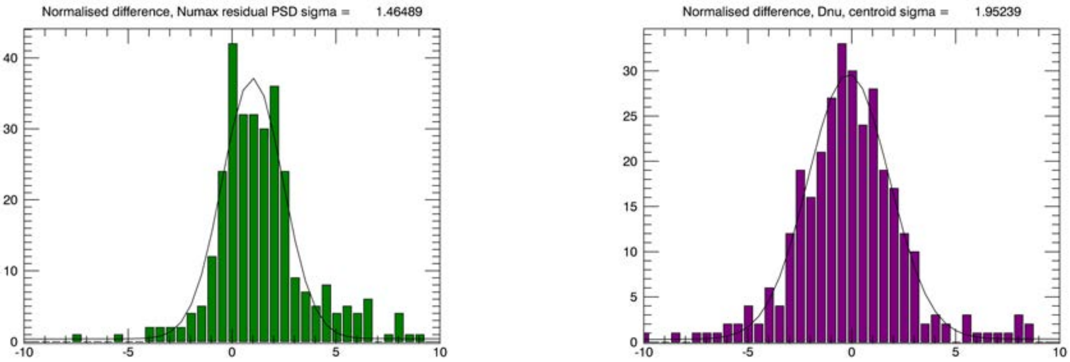


(a) ν_{max}

(b) $\Delta\nu$

Figure 5.11: Results on 319 CoRoT LRc01 stars returned to the CoRoT ExoGiants collaboration. These are shown compared to results obtained by Mosser. The black line shows the 1-1 relation,and serves to guide the eye.

The pipeline returns results for 319 of the 345 stars. The $\Delta\nu$ results shown in Figure 5.11 are taken from the PSPS method, and agree well with the values from the other pipeline. The ν_{max} values are generally high relative to the other pipeline, but did not fall outside of the range of values obtained by other pipelines in the collaboration that also returned results on these stars (Montalban, private communication).



(a) Normalised difference $\nu_{\max}$.

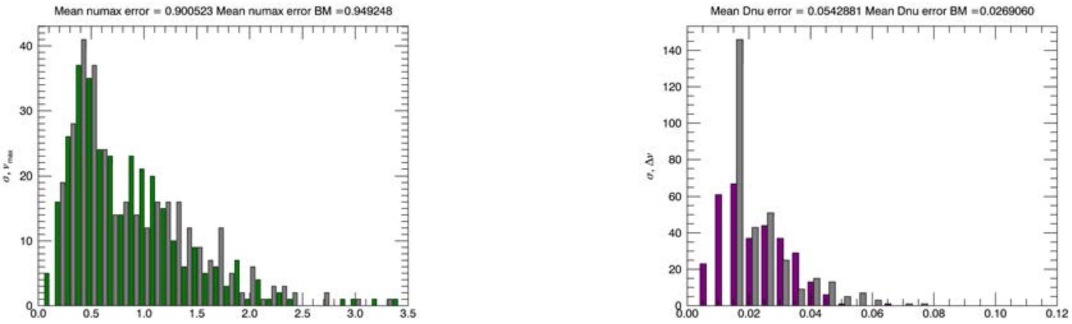
(b) Normalised difference $\Delta\nu$.

Figure 5.12: The error-bar normalised difference between the $\nu_{\max}$ and $\Delta\nu$ values obtained by this pipeline and that of Mosser and Appourchaux (2009).

In determining whether the errors obtained by the pipeline are representative of the true uncertainty, it is helpful to look at the difference between the “true” value and the computed value normalised by the error bar;

$$\frac{\nu_{\text{max, pipeline}} - \nu_{\text{max, true}}}{\sigma_{\nu_{\text{max, pipeline}}}} \tag{5.25}$$

The standard deviation of this quantity ought to be close to unity if the error bar is representative of the scatter in the data. Unlike the simulated stars, the real stars do not have a true (error free) value to compare to, and so we would expect the width of this in the real stars to be somewhat wider. Figure 5.12 shows this for both $\nu_{\max}$ and $\Delta\nu$ and this is indeed what we see. Figure 5.13 shows a direct comparison of the 1-σ uncertainties of the two pipelines and they are broadly consistent.

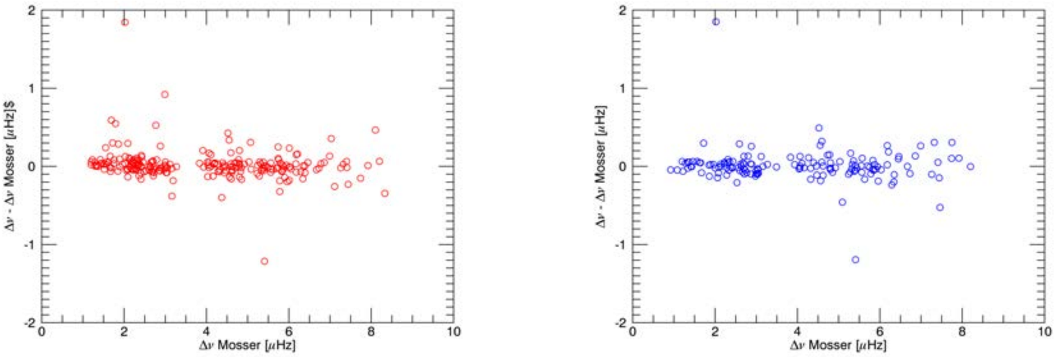


(a) Uncertainties on $\nu_{\max}$. Green bars are the values from this pipeline, grey are those from the work by Mosser.

(b) Uncertainties on $\Delta\nu$. Purple bars are the values from this pipeline, grey are those from the work by Mosser.

Figure 5.13: Uncertainties on $\nu_{\max}$ and $\Delta\nu$.

The results (shown in Figure 5.14) of the red-blue mask method also show good agreement with values from the Mosser pipeline, albeit with fewer stars returning results.



(a) 233 $\Delta\nu$ values returned by the red mask.

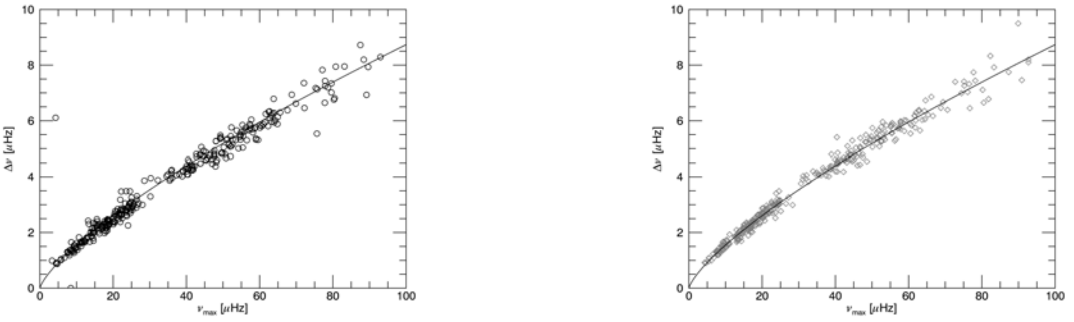
(b) 149 $\Delta\nu$ values returned by the blue mask.

Figure 5.14: $\Delta\nu$ values returned by the linear fits to frequencies and relative order.

A useful visual for the results is to plot the relationship between $\nu_{\max}$ and $\Delta\nu$. The theoretical relation is given by

$$\Delta\nu = 0.276\nu_{\max}^{0.751}. \tag{5.26}$$

Figure 5.15 shows this relation for the pipeline described in this chapter, alongside that of the Mosser et al. pipeline. They both visually are very similar, and the χ^2 values of the fits are comparable.



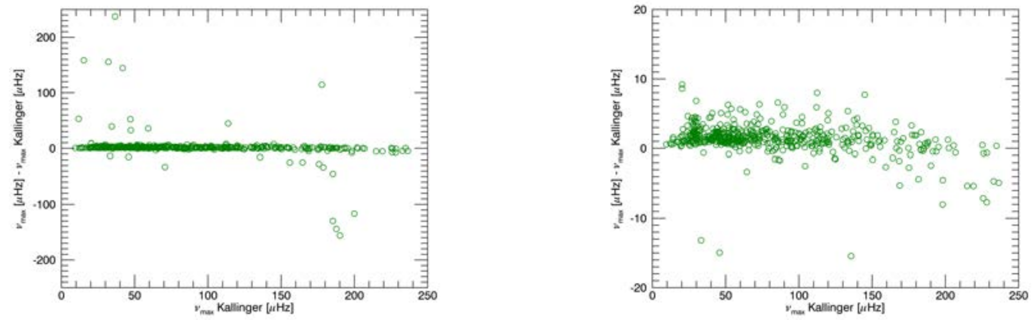
(a) $\nu_{\max} - \Delta\nu$ relationship for the pipeline described here. $\chi^2 = 11.23$

(b) $\nu_{\max} - \Delta\nu$ relationship for the Mosser et al. pipeline. $\chi^2 = 9.73$

Figure 5.15: $\nu_{\max} - \Delta\nu$ relationship for our pipeline, and that of Mosser et al. The solid line shows the theoretical relationship given by Equation 5.26.

5.5.2 *K2*

The following are part of the work towards results contributed to the K2 Galactic Archeology project, and published in Stello et al. (2016). Similarly to the CoRoT results, they are compared to results from another pipeline described in Kallinger et al. (2016) and Stello et al. (2017) (private communication).

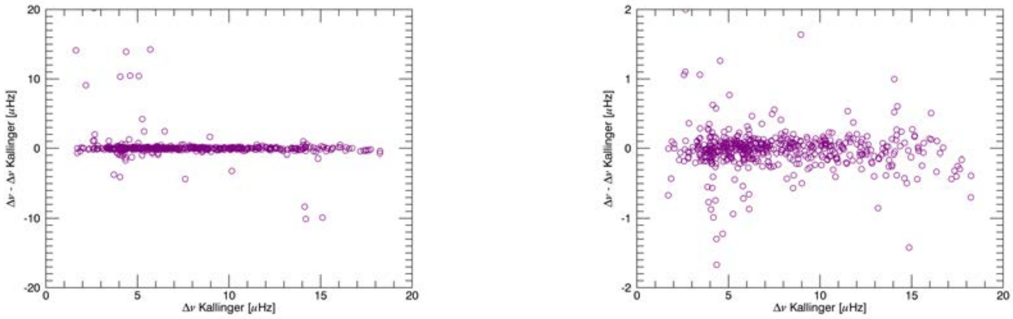


(a) $\nu_{\max}$

(b) A zoomed in version of panel a, showing only values with the best agreement.

Figure 5.16: $nu_{\max}$ results on 1051 K2 C1 stars. The y-axes show the differences between the results obtained by our pipeline and those reported by Kallinger et al.

The agreement between the results for the two pipelines for the K2 data (as shown in Figures 5.16 amd 5.17) is again generally good. For the relatively few stars where the pipelines disagree dramatically, the pipeline, despite accurately identifying the region of

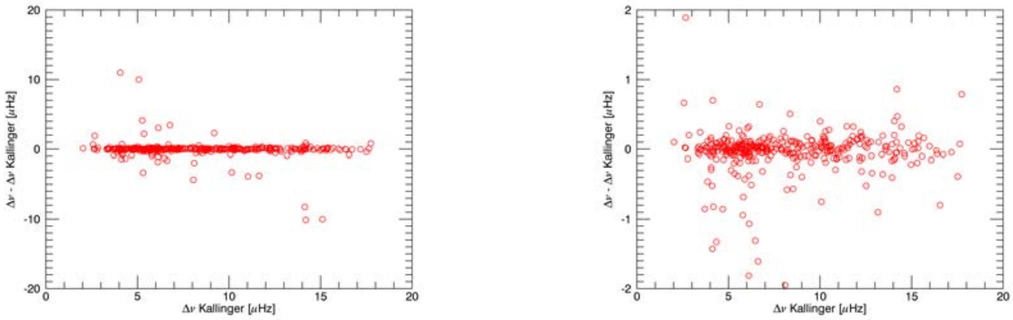


(a) $\Delta\nu$

(b) A zoomed in version of panel c, showing only values with the best agreement.

Figure 5.17: $\Delta\nu$ results on 1051 K2 C1 stars. The y-axes show the differences between the results obtained by the PSPS method in our pipeline and those reported by Kallinger et al.

frequency of the oscillations failed to fit one or other of the parameters correctly in cases of low SNR. In practice these results would be removed from any sample returned to a collaboration. Figures 5.18 and 5.19 show the results for the red-blue mask method. As expected, this method returns fewer values but has in general a good agreement with the Kallinger pipeline. Since the $\Delta\nu_{\mathrm{guess}}$ used to lay down the priors on the red and blue masks is based on the PSPS value, it is not surprising that similar incorrect values are returned by this method.

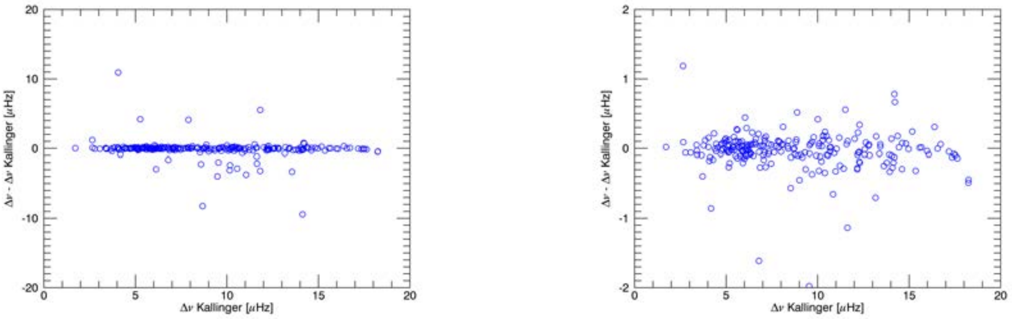


(a) 374 $\Delta\nu$ values returned by the red mask.

(b) A zoom of panel a, excluding the values which depart the most.

Figure 5.18: The difference between $\Delta\nu$ values returned by the linear fits to frequencies and relative order using the red mask, and those returned by Kallinger et al.

Figure 5.20 shows the relationship between ν_{max} and $\Delta\nu$ output by both pipelines

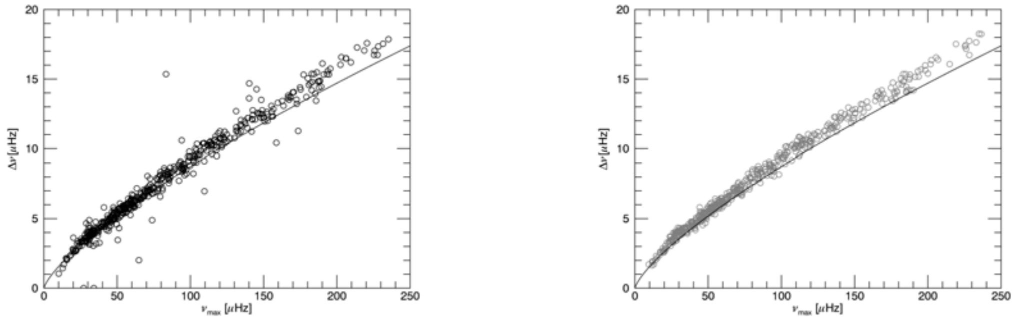


(a) 241 $\Delta\nu$ values returned by the blue mask.

(b) A zoom of panel c, excluding the values which depart the most.

Figure 5.19: As Figure 5.18, but using the blue mask.

depart in a very similar fashion from the theoretical relationship, as is expected for stars observed by K2. This is due to the underlying population of stars being slightly different to the *Kepler* field, on which this relation is based.



(a) $\nu_{\rm max} - \Delta\nu$ relationship for our pipeline.

(b) The same, for the Kallinger pipeline.

Figure 5.20: $\nu_{\rm max} - \Delta\nu$ relationship for our pipeline, and that of Kallinger et al. The solid line shows the theoretical relationship given by Equation 5.26.

5.5.3 Simulated Spectra

Below are results for 420 simulated spectra at the frequency resolution of K2. The results for $\nu_{\rm max}$ shown in Figure 5.21 do seem to display a slight overestimate of the parameter at low frequency and an underestimate at higher frequencies. For the very largest departures this discrepancy is of the order of 15%. This was reported when returning results to collaborations, and it did not seem that this systematic error caused the pipeline to underperform relative to others, however, the reason for this is a cause for some concern and would benefit from further work in the future.

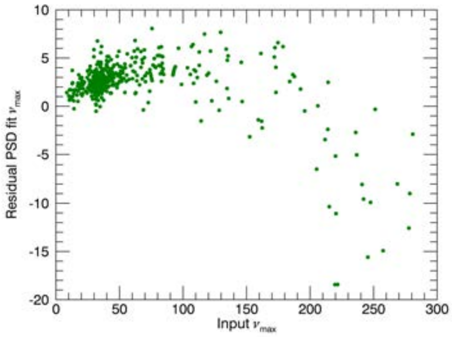
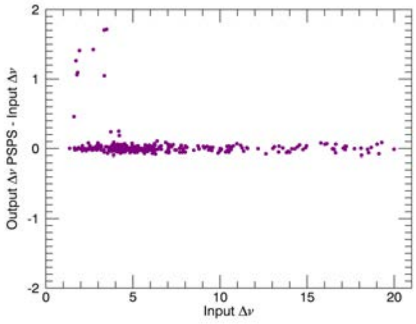
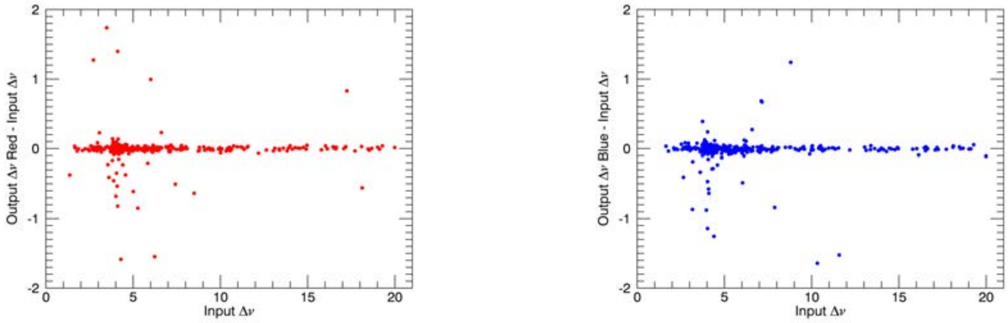


Figure 5.21: The difference between input ν_{max} output ν_{max} for simulated K2 stars.

The picture for $\Delta\nu$ is much better, however. As well as displaying good agreement with other pipelines for real data, the recovery of the correct simulated $\Delta\nu$ is excellent for all methods, as shown in Figure 5.22.



(a) $\Delta\nu$ from PSPS for simulated K2 stars.



(b) $\Delta\nu$ from red mask for simulated K2 stars. (c) $\Delta\nu$ from blue mask for simulated K2 stars.

Figure 5.22: The difference between input $\Delta\nu$ and output $\Delta\nu$ for simulated K2 stars.

For these $\Delta\nu$ values, we of course do have the correct result. This provides a good test of the uncertainties. The normalised difference

$$\frac{\Delta\nu_{\mathrm{pipeline}} - \Delta\nu_{\mathrm{actual}}}{\sigma_{\Delta\nu_{\mathrm{actual}}}} \tag{5.27}$$

is informative on how well the output uncertainties represent the true underlying scatter. Figure 5.23 shows histograms of this quantity for the PSPS and red/blue mask results. This lends confidence to the obtained uncertainties, since the standard deviation of this quantity is indeed close to unity for all methods.

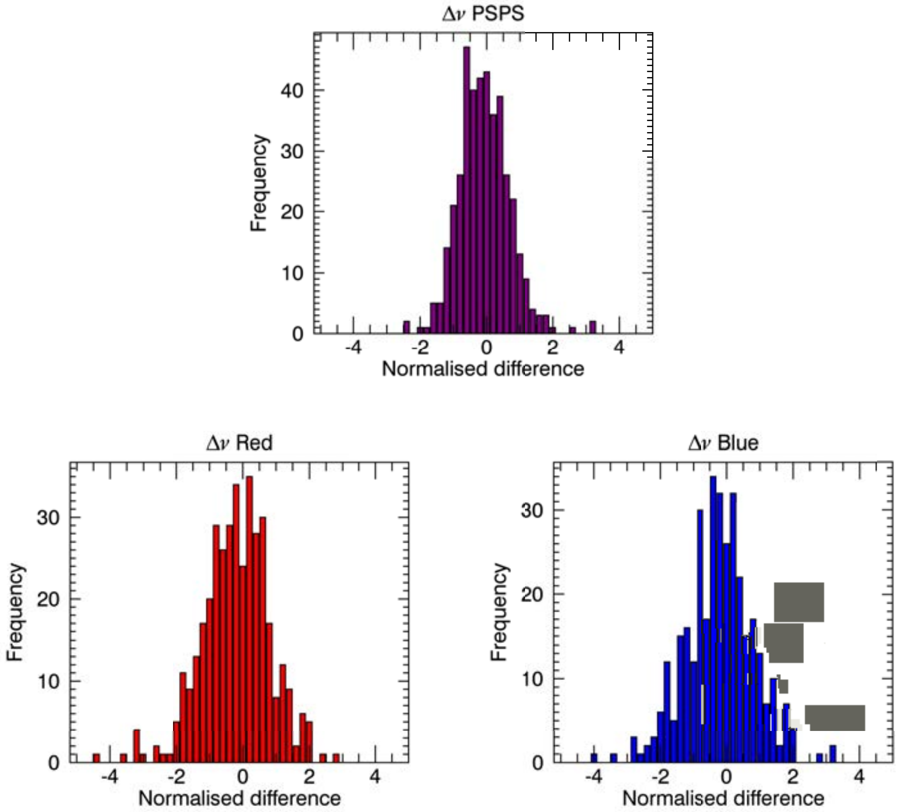


Figure 5.23: The normalised difference for simulated K2 stars. Recall that if the error bars output by the pipeline are representative of the uncertainty on the measurement, we would expect the standard deviation of the normalised difference, and hence the width of these histograms to be unity.

Chapter 6

Variance in the Red Clump

In Chapter 3 we see that the variance of red giant branch stars has a different relationship to ν_{max} than red clump stars. This chapter describes attempts to simulate these stars in the same way as the RGB stars, and a discussion of what properties of the star lead to this.

Recall the discussion of stellar evolution in Chapter 1. Red giant stars are in a late phase of their evolution, having exhausted the supply of hydrogen in their cores, and begun fusing hydrogen in a shell around their core. Low mass red giants go through the helium flash due to runaway nuclear fusion in their degenerate cores becoming hot enough to begin fusing helium. Slightly higher mass stars have core masses large enough to begin helium fusion without the core having to become degenerate. Core helium burning stars are classified into two “clumps”, depending on which of these regimes they fall into, the lower mass population forming what is generally referred to as the red clump, and the higher mass, the secondary clump.

The different internal structure of these outwardly very similar stars obviously impacts on the oscillations observed, as their propagation through the interior is effected. The échelle diagram Figure 6.1 shows this in a visual way. This very clear difference in the structure of the dipole modes is induced by the period spacing of the g-modes and hence the coupling with the $\ell = 1$ p-modes. The physical reason for this is that during the first core helium burning phase (generally referred to as “the red clump”), the core expands and the outer envelope expands, increasing both $\Delta\nu$ and $\Delta\Pi$.

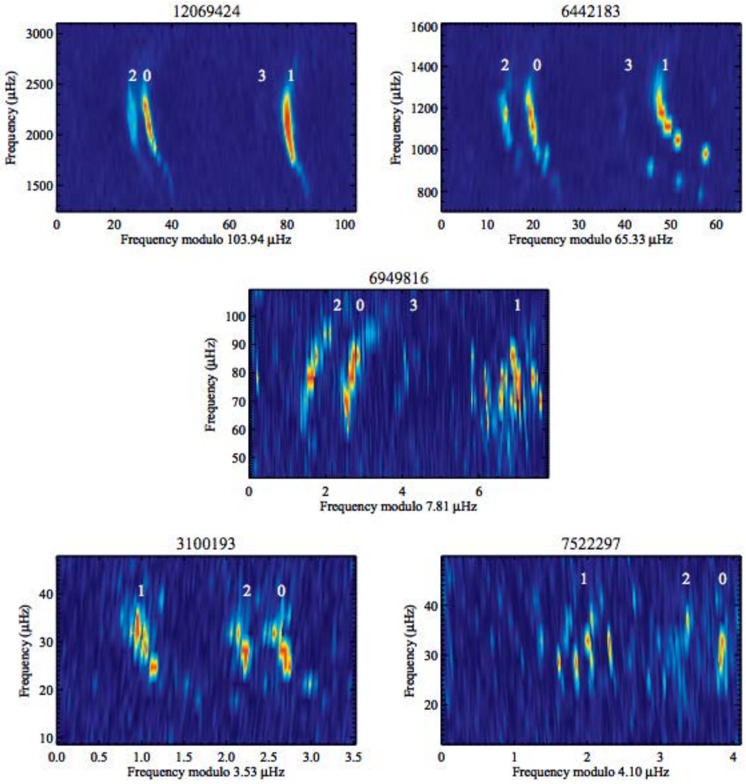


Figure 6.1: Example of échelle diagrams for stars at different evolutionary stages; from the top left, a main-sequence star 16 Cyg A, a subgiant, two RGB stars and an RC star. KIC 3100193 and KC 7522297 have very similar surface properties but have very different $\ell = 1$ mode structure. Image from Chaplin et al. (2013).

Chapter 3 described in detail how the frequencies and heights of these mixed modes are dependent on the coupling between the p- and g-modes in the star, and it is the considerably higher period spacing for the core helium burning stars that induces the differences in the $\ell = 1$ structure that we observe. Simulating modes with these different properties is key to attempting to reproduce the different relationship between variance and ν_{max} we see for red giant branch and red clump stars, and understanding what process it arises from, since the granulation and mode power both contribute.

6.1 Simulating modes in the Red Clump

Chapter 3 described in detail how the mixed modes present in red giant branch stars are simulated, including their frequencies, degree of coupling between the p- and g-modes and how the heights and linewidths are modified. All of these are dependent on their g-mode period spacing $\Delta\Pi$. Recall that the red giant branch period spacings are assumed to follow a linear relationship with $\Delta\nu$, described by Equation 3.34. The situation for red clump and secondary clump stars is very different, as shown in Figure 6.2, displaying period spacing data from Mosser et al. (2014). Red Clump stars have much higher period spacings than RGB stars for a given $\nu_{\rm max}$.

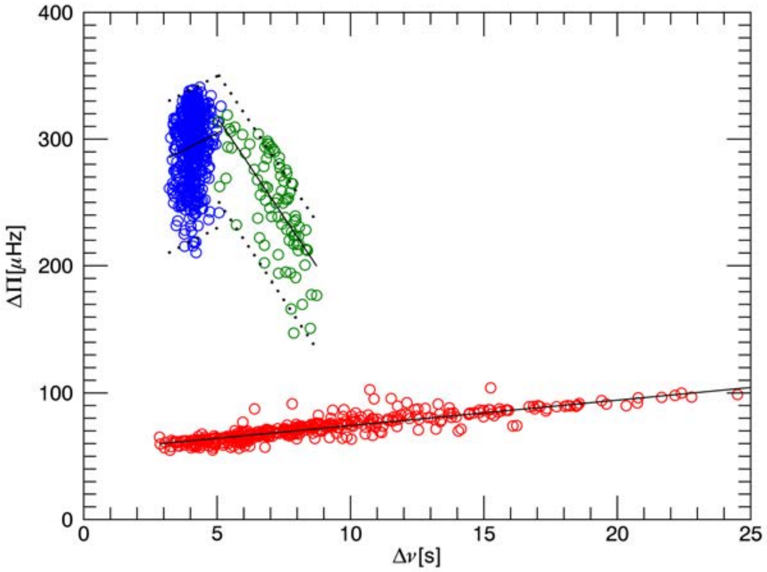


Figure 6.2: $\Delta\Pi$ vs $\Delta\nu$ data for *Kepler* red giants from Mosser et al. (2014). Red Giant Branch stars are shown in red, red clump stars in blue and secondary clump stars in green. Solid black lines show best linear fits for each evolutionary state, and the dotted lines show the upper and lower limits used in generating synthetic $\Delta\Pi$ values, as described by Equations 6.1 - 6.4.

To reproduce this observed behaviour in the artificial data, a linear fit is performed for each evolutionary state separately, then the $\Delta\Pi$ value is distributed uniformly between an upper and lower limit defined by this fit line, described by the equations;

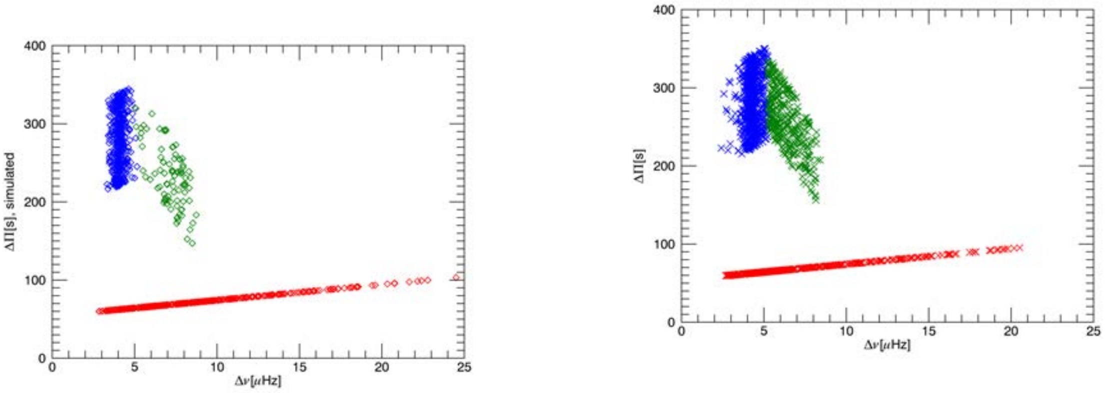
$$RC_{\text{low}} = 10.8\Delta\nu + 250.7 - 70 \tag{6.1}$$

$$RC_{\text{high}} = 10.8\Delta\nu + 250.7 + 45 \tag{6.2}$$

$$2RC_{\text{low}} = -31.5\Delta\nu + 474.8 - 50 \tag{6.3}$$

$$2RC_{\text{high}} = -31.5\Delta\nu + 474.8 + 40 \tag{6.4}$$

The result of calculating red clump period spacings by randomly distributing them between these limits using the observed $\Delta\nu$ values of the real stars from Mosser et al. (2014) is shown in Figure fig. 6.3a. Given the small number of stars, this approximates the real values reasonably well. Figure fig. 6.3b shows the same plot for a population of simulated stars from TRILEGAL (see Chapter 2) which acts as our synthetic *Kepler* field. The higher density of points here is entirely due to the TRILEGAL field having more stars than the Mosser et al. (2014) data. The distribution for both panels of Figure 6.3 is the same.



(a) Period spacing calculated from observed $\Delta\nu$ for the same real stars.

(b) The same plot for simulated stars from TRILEGAL.

Figure 6.3: Simulated period spacings following Equations 6.1 to 6.4.

We go on to simulate power spectra for the TRILEGAL stars using the values of $\Delta\Pi$ obtained from these relations, and the formulation as described in Chapter 3. The properties of the $\ell = 1$ mixed modes are governed by the period spacing, and as such we expect the simulated power spectra to be appreciably different for outwardly similar stars, as for the real stars in the échelle diagram in Figure 6.1. It is helpful at this point to look at the diagnostic plot of the bulk properties of the mixed mode parameters as we did in Chapter 3. Figure 6.4 shows the coupling parameter Q_1, the number of $\ell = 1$ mixed modes per order, and the total power present in both the $\ell = 1$ modes and modes from $\ell = 0, 3$ in units of total $\ell = 0$ power, ensuring that the power levels are consistent with a mixed mode free realisation.

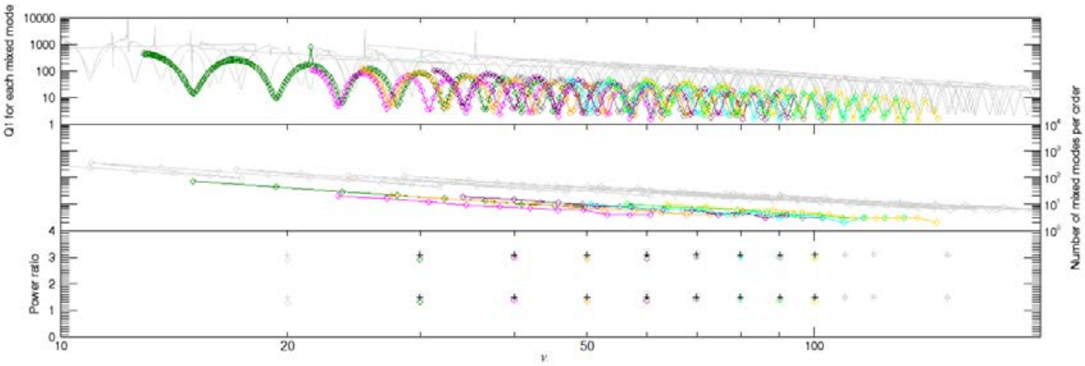
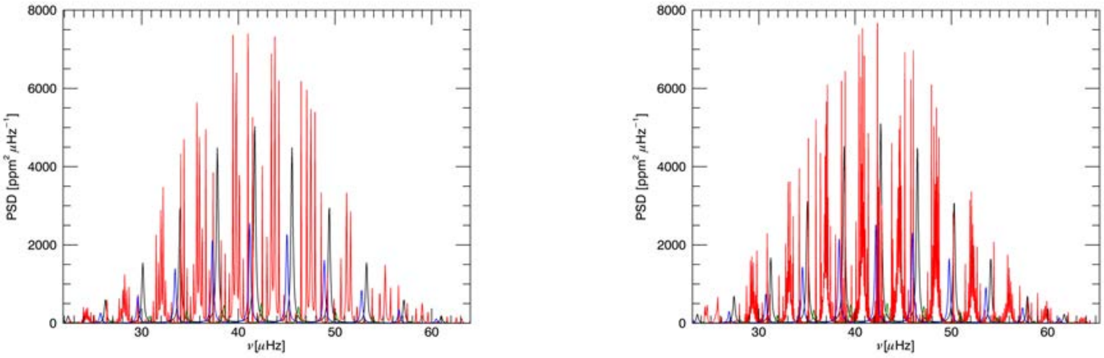


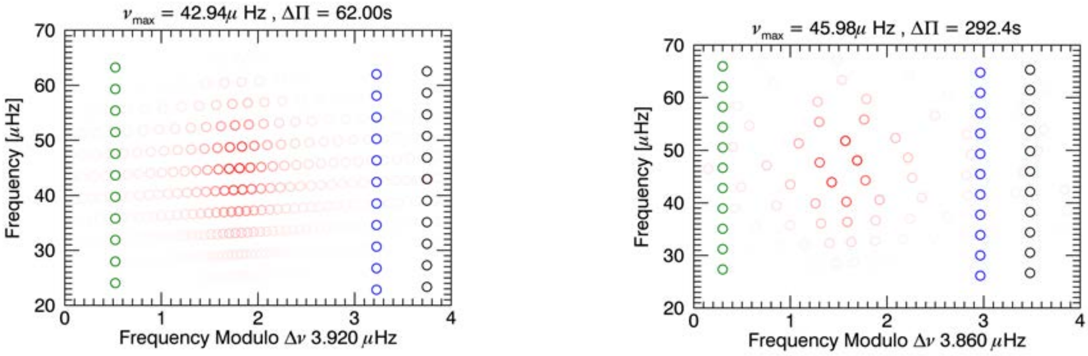
Figure 6.4: $\ell = 1$ mixed mode diagnostics for red clump and secondary clump stars. The grey plot shows these same quantities for RGB stars of the same $\nu_{\max}$. Top panel: the inertia ratio Q_1 for each mixed mode simulated. The values for modes of 16 different simulated stars are shown in different colours. Middle panel: The number of mixed modes per order, plotted against nominal $\ell = 1$ frequency. Bottom panel: A sanity check for power levels. Coloured stars show the total power in units of $\ell = 0$ power for the mixed mode simulations, black circles for the pure p-mode case. Coloured diamonds and black crosses follow the same convention for the ratio of $\ell = 1$ power to $\ell = 0$ power.

Notice that Q_1 in the red clump regime, while displaying similar behaviour to the that of an RGB star, is appreciably lower. This in turn results in a lower number of modes per order for stars of the same frequency. Q_1 also dictates the frequencies, heights and linewidths of the mixed modes. Figure 6.5 shows the oscillation spectra and échelle diagrams of 2 simulated stars with the same $\nu_{\max}$ but one has a typical RGB period spacing, while the other has a red clump period spacing.



(a) Simulated mode spectrum of an RGB star.

(b) Simulated mode spectrum of an RC star.



(c) Échelle diagram of an RGB star.

(d) Échelle diagram of an RGB star.

Figure 6.5: The appearance of mixed modes in the spectra and Échelle diagrams of two simulated red giants of the same $\nu_{\rm max}$ but different evolutionary states.

Again, it is the échelle diagram which illustrates most clearly the difference in the $\ell = 1$ modes under the two different period spacing regimes. There are far fewer modes in the red clump star, and they are much less densely distributed in frequency. Besides the period spacing governing the properties of the $\ell = 1$ mixed modes, all other mode parameters (such as p-mode visibilites, nominal frequencies, linewidths) remain the same between RGB and red clump stars. Since RGB and clump stars do have different distributions of luminosity, temperature and mass, all parameters that are in the scaling relations for $A_{\rm max}$, $\Delta\nu$, and $\nu_{\rm max}$, there ought to be slightly different distributions for these values, but the dependencies are not changed in these simulations. In any case, it is an interesting exercise to study the effect of only changing the period spacing on the relationship between variance and $\nu_{\rm max}$. Of course, both modes and the granulation background contribute to the variance metric we have worked with throughout this thesis.

6.2 The Granulation Background Contribution

The granulation parameters of red giant branch stars are spread over a much larger range than those of red clump stars. This is intuitive, since, as discussed in Mathur et al. (2011), red clump stars occupy a much more narrow range of temperature, luminosity and hence, ν_{max}. The granulation model used for our simulations is “model H” from Kallinger et al. (2014), and does include a mass dependency, and so we would expect the granulation contribution to the variance for the red clump stars to be impacted accordingly. Kallinger et al. (2014) does show a deviation from the power law relationship for an amplitude parameter in their model which approximately falls in the ν_{max} range occupied by the clump and secondary clump.

In the absence of separate relations for the granulation parameters of clump stars, we leave any differences in the model parameters down to the dependencies on the physical parameters of the star. Figure 6.6 shows the mass distributions for the simulated TRILEGAL stars used, colour coded for evolutionary stage. This clearly shows the higher mass regime of the secondary clump: recall these are the evolved red giants who have a core mass high enough to ignite helium burning without the core becoming degenerate.

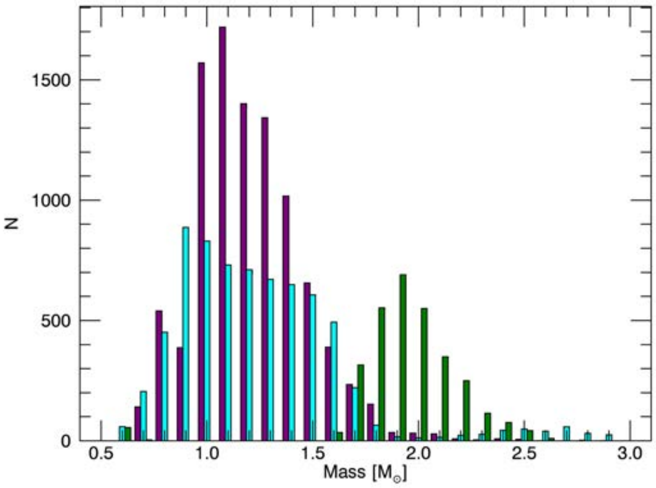
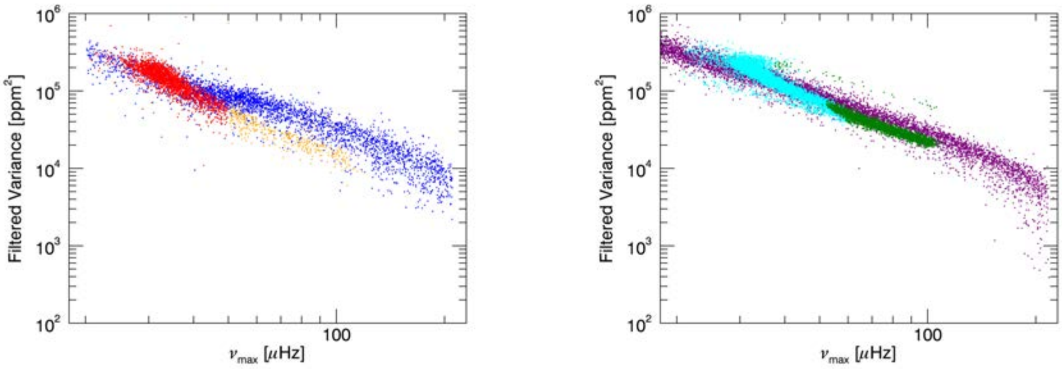


Figure 6.6: Histogram of masses of simulated stars. Purple shows RGB stars, cyan shows the first red clump and green, the secondary clump.

The most clear feature in this is the generally higher mass for the secondary clump stars, while the clump and RGB stars have reasonably similar mass distributions. Since the granulation background model does have a mass dependent term, the evolutionary stage will have an effect on the contribution of the granulation to the variance, particularly for secondary clump stars. We now go on to look at the results of simulating the variance, mode power and granulation power for these three evolutionary states, and comparisons with those for real stars.

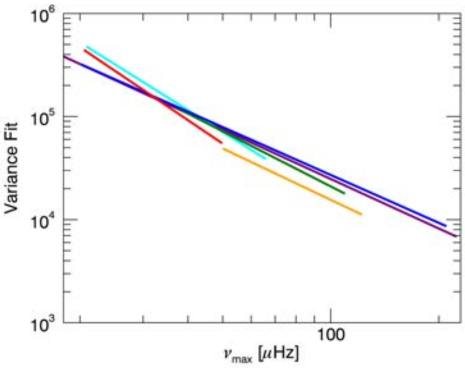
6.3 Results

The relationship between a bandpass filtered estimate of the stellar variance and ν_{max} has been a recurring theme in the work contained in this thesis. We noted in Chapter 3 that the slope of the approximate power law relationship between these quantities is different for the hydrogen shell burning red giant branch stars and the core helium burning red clump and secondary clump stars. Exactly the same method for measuring the filtered variance used in all previous chapters was also used here; integrating the power density spectrum over the frequency region centred on ν_{max} which contains the oscillation power, including the contribution from the granulation background. Figure 6.7 shows variance against $\nu_{\max}$ coloured by evolutionary state from the power spectra of both real and simulated stars, and comparison of power law fits to these.



(a) Variance vs $\nu_{\max}$ for real *Kepler* giants. Blue denotes RGB, red, clump stars and orange, the secondary clump.

(b) Variance vs $\nu_{\max}$ for simulated giants. Purple denotes RGB, cyan, clump stars and green, the secondary clump.



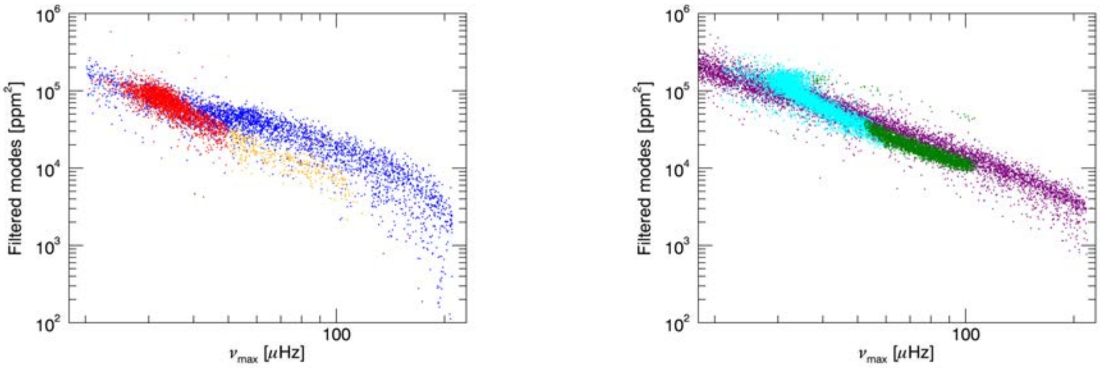
(c) Power law fits to the colour coded data above.

Figure 6.7: Variance-$\nu_{\max}$ relationship for real and simulated populations.

Recall that the only different formulation in the simulations is the period spacing in the clump stars. The slopes of the relationship for clump stars and RGB stars in the simulation are noticeably different, in fitting the relationships with power laws, we find that the the variance of RGB stars goes as $\nu_{\max}^{-1.59}$ while for the clump, this relation is $\nu_{\max}^{-2.18}$. The secondary clump relation goes as $\nu_{\max}^{-1.77}$, closer to the value for the RGB than the secondary clump, as is the case for the real stars, though the real secondary clump variance falls off more quickly with $\nu_{\max}$, and the real secondary clump stars have a lower coefficient, as is clear in the plot, these stars sit firmly below the RGB relation, where this is not the case for the simulations.

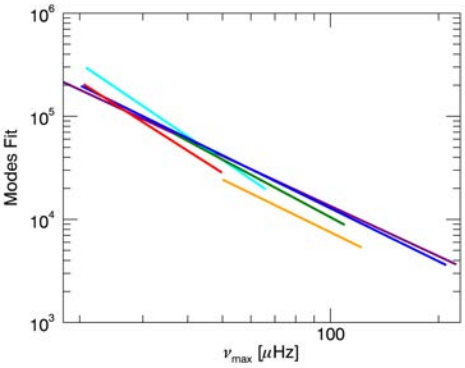
The parameters of all the power law fits in this section are contained in Table 6.1, and also power laws drawn of the fits for the real and simulated quantities are shown, allowing for more clarity in visually comparing the slopes than simply showing the data on the same plot. The exponents of the simulated and real red giant branch stars are broadly consistent with each other for the variance, as well as the modes and background contribution, lending confidence to the assumption that any discrepancies between the exponents of the simulated and real clump stars are the result of departures from the RGB scaling that are not accounted for in the simulations. We may conclude that the different behaviour observed in the variance can be at least partly explained by the different period spacings observed at the same $\nu_{\max}$.

It is helpful, alongside the variance scaling, to also quantify the slopes of the integrated background and mode contributions separately, as the only change we have made is in the simulation of the modes. The background profile is of course known for the simulated stars, while for the real stars, we rely on a median filter as described in Chapter 5. This can sometimes remove a small amount of power from the modes, but as Figures 6.7c, 6.8c, 6.9c and Table 6.1 show that the simulated and real RGB profiles are all consistent with one another, and there is no reason for any issue with the background filter to effect the clump stars differently, we can assume that the background estimation is not responsible for any difference in the slopes.



(a) Mode power vs $\nu_{\max}$ for *Kepler* giants. Colours as in Figure 6.7.

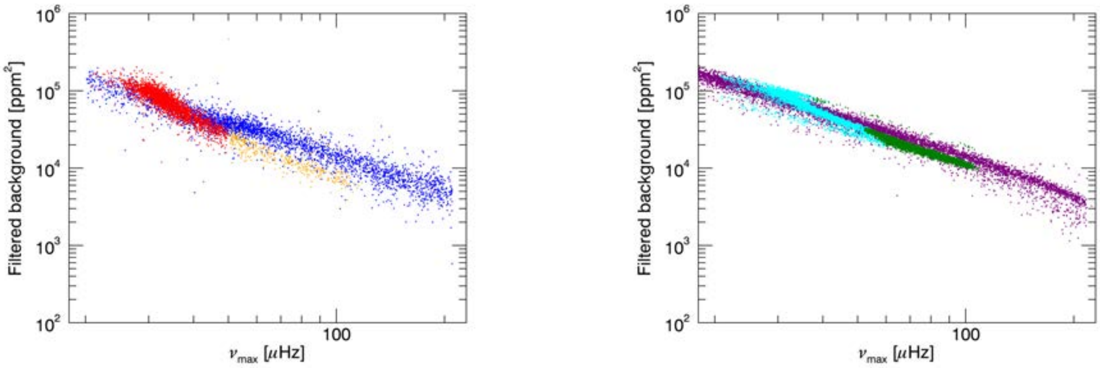
(b) Mode power vs $\nu_{\max}$ for TRILEGAL giants. Colours as in Figure 6.7.



(c) Power law fits to the colour coded data above.

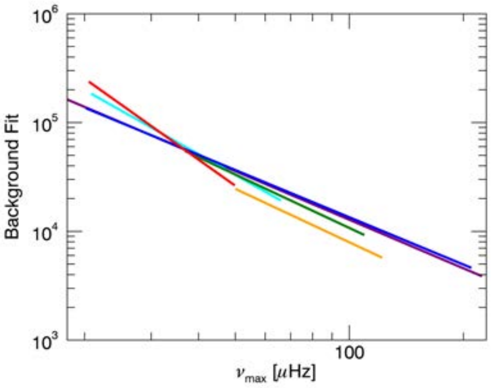
Figure 6.8: Integrated mode power-$\nu_{\max}$ relationship for real and simulated populations.

The fits to the mode power for the secondary clump and the RGB in the real stars have the same exponent of $\nu_{\max}^{-1.70}$, though the red clump stars have a considerably lower coefficient against the branch stars. This is not replicated in the simulations, however the first clump exponents for the real and simulated stars are comparable, going as $\nu_{\max}^{-2.22}$ and $\nu_{\max}^{-2.53}$ respectively. This suggests that there are properties of the power spectrum of the secondary clump stars which are different from both the RGB and the first clump. Figure 6.8b shows the slopes of all the quantities are reasonably consistent between the real and simulated stars, suggesting it is the different coefficients which are the main contributor to the mismatch between the real and simulated data.



(a) Background power vs $\nu_{\rm max}$ for *Kepler* giants. Colours as in Figure 6.7.

(b) Background power vs $\nu_{\rm max}$ for TRILEGAL giants. Colours as in Figure 6.7.



(c) Power law fits to the colour coded data above.

Figure 6.9: Integrated granulation power-$\nu_{\rm max}$ relationship for real and simulated populations.

The slopes of the background contribution to the variance are consistent between the real and simulated stars for the RGB and secondary clump, though the simulation seems to overestimate the absolute value of the integrated power at a given frequency. The red clump relationship, however, is not captured well by the simulations. This suggests that the mass dependency of the granulation amplitude ($M^{-0.26}$) described in the Kallinger et al. (2014) model does not necessarily capture the difference between RGB and clump stars adequately purely based on their different mass distributions. A measure of the background value at $\nu_{\rm max}$ by Mosser et al. (2012a) has a much stronger mass dependency of $M^{-3.0}$, though this is not something we can translate into the background profile across

	$10^a\nu_{\max}^{-b}$		
	Variance	Modes	Background
Real RGB	7.51;1.54	7.52;1.70	7.01;1.44
Simulated RGB	7.58;1.59	7.35;1.61	7.07;1.48
Real Clump	8.73;2.35	8.22;2.22	8.62;2.47
Simulated Clump	8.56;2.18	8.58;2.35	7.85;1.96
Real Second Clump	7.50;1.65	7.27;1.70	7.16;1.63
Simulated Second Clump	7.87;1.77	7.74;1.86	7.29;1.63

Table 6.1: Parameters of power law fits to the listed properties of real and simulated stars. Values are shown as $\log_{10}$ coefficient;exponent.

the whole frequency range.

While this different relationship between variance and $\nu_{\max}$ for red giants has been noted in previous work (Hekker et al. (2012), among others) and is often attributed to the mass distribution of the giants, it is likely that simple power law relationships for many quantities used in the simulations here (envelope height and width, for example) do not adequately capture the properties of the power spectrum where red clump stars are concerned. Mosser et al. (2011) frequently makes mention of subtle departures from these power laws for red clump stars. It is however interesting that in the absence of any such changes in the dependency, the differing g-mode period spacing and therefore $\ell = 1$ profiles alone do induce a different scaling for the mode power of clump giants.

In summary, in this short investigation, we can see that making one subtle alteration to the input physics we use in simulations has an impact on the resulting power spectrum outside of simply altering the appearance of the $\ell = 1$ region. The only parameter we change here is the $\ell = 1$ g-mode period spacing, which reflects the difference in structure in the deep interiors of red giant branch and red clump stars. The fact that in simulations, the distribution of power among the modes is sufficiently impacted such that our filtered variance metric has a measurably different relationship to $\nu_{\max}$ for the three evolutionary states is interesting. This shows that the physics of the deep interior makes a significant contribution to the relationship of mode power to variance, something which has hitherto only been discussed in terms of clump and branch stars having subtly different amplitude scaling and granulation profiles, for example. While these reflect the physics of the convective region, the g-mode spacing is a signature of the deep interior. The fact that the slopes of the relationship to nu_{max} of both mode power and variance are also broadly consistent between the simulations and the real data, suggest in fact that the physics of the deep interior is a candidate for the dominant reason the variance-ν_{max} relationships are so markedly different for these three evolutionary stages, since they are reproduced to a good extent without including any subtle differences in the physics of the outer regions. This result certainly merits further, more detailed work, and represents an opportunity to update current understanding of red giant evolution.

Chapter 7

Conclusion and the Future.

The common theme throughout this thesis has been the use of a simple but powerful global asteroseismic metric of filtered stellar variance to study red giant stars both on a star-by-star basis and as a ensemble population which reflects the formation and evolution of the Galaxy as we know it today.

Chapter 3 described the considerations and testing of relations and assumptions in creating realistic red giant power spectra. The power spectra components due to granulation, oscillations and shot noise are included. Extensive testing of amplitude scaling relations and the inclusion of mode spacings and mixed mode properties informed by theory and observation led to the creation of simulated power spectra that not only were consistent with real stars in terms of the variance metric, but realistic enough to test methods of asteroseismic parameter estimation.

Chapter 4 detailed work into inferring the fraction of stars in binary systems among *Kepler* red giant branch stars. The presence of an unseen companion star "washes out" the oscillation and granulation signal that makes up the variance metric, causing it to be lower than would be predicted by theory and observations of single stars. As such the distribution of the residuals of the variance metric about some relation with $\nu_{\rm max}$ reflects the number of stars which suffer from this "washout". To probe the underlying distribution of residuals, a Bayesian inference technique was used, allowing the most probable model of the distribution as a combination of binary and single stars to be determined by choosing a suitable model and constraining the parameters of the fit informed by the simulated populations. A strength of this technique is the ability to update the prior assumptions included in response to any future constraints placed on parameters of the stellar population. The result of this investigation is an inferred binary fraction of $57.4 \pm 2.5\%$, consistent with older studies of main sequence stars. Future possibilities for this work include using the method on stars of other evolutionary states, further investigating the effect of the mass ratio distribution on the inferred fraction, and extending the analysis to different populations in the Galaxy, for example stars observed by K2, and in the near future, TESS. TESS will provide all-sky observations of the brightest stars, allowing for investigation of stars in the local solar neighbourhood.

Chapter 5 describes the use of predicted and observed signal-to-noise and variance to automate the detection of oscillating stars in some of the newest K2 data as well as in the older CoRoT data, covering different regions in space and across the red giant lifetime of these stars. The detection test successfully constrains the frequency region containing the oscillation modes well. ν_{max} is found by fitting the Gaussian peak modulating the modes, and $\Delta\nu$ is estimated using a fit to the second power spectrum, as well as a method based on searching for modes directly in the power spectrum. Results from this pipeline were submitted to two studies, and were in generally good agreement with results obtained by other pipelines. Future work ought to focus on correcting a systematic error in the ν_{max} determination.

Chapter 6 shows how the variance metric captures to some degree the changing properties of the oscillations as red giants begin fusing helium during the red clump phase. In the attempt to replicate the relationship between variance and ν_{max} for the red clump stars, we found that a subtle change in the mixed mode properties is enough to produce a measurable differences in this metric as the star evolves, suggesting that the different mass distribution of red clump and red giant branch stars is not the only contributor to the different variance - ν_{max} relationship, as has been suggested. There is evidence to suggest that the granulation signal of red clump stars differs from red giant branch stars, however in the absence of an empirical relation for these stars it was not possible to create these simulations at this time.

The binary star fraction inference study has the scope to be extended. More investigation of the effect of the binary mass ratio distribution is needed since the result as it stands is based on priors and simulations which assume a binary mass close to unity. More simulation and testing of the precise effects of the mass ratio distribution would lend more weight to the conclusion of this investigation by allowing some relaxation of the reasonable (but not necessarily correct) assumption we make here on the mass ratio distribution. There is also nothing to stop the method in future being applied to other populations of other solar-like oscillators. It would be interesting to determine binary fractions for each of the many different K2 fields to see if there are any significant dependence of binary fraction on Galactic location.

The main future work on the pipeline ought to focus on finding the cause of the systematic error on $\nu_{\max}$. I believe that the shape of the background used to fit for $\nu_{\max}$ may be the cause of the bias that we see. Even when fitting to the residual power spectrum, there will always be some residual background power remaining, since the variable width moving median filter provides a good estimate of the background but does not attempt to fit the actual form of the granulation. This is of course for the sake of efficiency and time, and works very well for the detection test, but it may be the case that the form of the function we fit to the residual power may need more care when determining accurate parameters. While the red-blue mask method works very effectively in it's current form, there is also the possibility of improving it by including tests for multiple spikes over a given window, as described by Basu and Chaplin (in the press). Initial tests done on this were promising, however more careful tuning of search window sizes and spike widths are needed.

The short study of the effect of different mixed mode properties on red giant branch and red clump stars described by Chapter 6 was not expected to show such intriguing results. This was largely included for completeness, since when we introduced the variance metric we mentioned that it's relationship to ν_{max} differs between these three evolutionary stages, and it seemed odd to not revisit it. However what we found; that the properties of the $\ell = 1$ mixed modes, and therefore the deep interior of the star have a profound influence on the distribution of oscillation power, and therefore the variance metric. In my view, more advanced simulation of red clump power spectra is warranted, since the idea that the deep interior has such an effect is in opposition to some current thinking. Detailed thought on the avenue further work should take is a responsibility I will leave for others, however deriving robust scaling laws for the granulation background power of these stars, as well as the relationship between $\Delta\nu$ and $\Delta\Pi$ would be an excellent starting point. Once this is combined with our new understanding of the impact that the very subtle change in the $\ell = 1$ mixed mode profiles has, it may be possible to determine what the dominant physical reason for the departure of red clump stars from the red giant branch relation is.

The work contained within this thesis of course represents a very small corner of what is possible to investigate by studying global asteroseismic properties of evolved solar-like oscillators. With the continuing operation of the K2 mission (at the time of writing!), and with much more data to come from the TESS and PLATO missions, the future of asteroseismology certainly looks bright and busy. I will watch it from the outside with keen interest, and fond memories of my time in the field.

Bibliography

Edwin R. Anderson, Jr. Duvall, Thomas L., and Stuart M. Jefferies. Modeling of solar oscillation power spectra. *The Astrophysical Journal*, 364:699, dec 1990. ISSN 0004-637X. doi: 10.1086/169452. URL `http://adsabs.harvard.edu/abs/1990ApJ...364..699A`.

A. Baglin, E. Michel, M. Auvergne, and COROT Team. The seismology programme of the CoRoT space mission. In *Proceedings of SOHO 18/GONG 2006/HELAS I, Beyond the spherical Sun*, volume 624 of *ESA Special Publication*, page 34.1, October 2006.

J Ballot, C Barban, and C Van 't Veer-Menneret. Astronomy & Astrophysics Visibilities and bolometric corrections for stellar oscillation modes observed by Kepler (Research Note). *A&A*, 531, 2011. doi: 10.1051/0004-6361/201016230.

R.J. Barlow. *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences.* Manchester Physics Series. Wiley, 1989. ISBN 9780471922940. URL `https://books.google.co.uk/books?id=IhmoQgAACAAJ`.

S Basu and W J Chaplin. *Asteroseismic Data Analysis: Foundations and Techniques.* Princeton University Press, in the press.

T R Bedding, D Huber, D Stello, Y P Elsworth, S Hekker, T Kallinger, S Mathur, B Mosser, H L Preston, J Ballot, C Barban, A M Broomhall, D L Buzasi, W J Chaplin, R A García, M Gruberbauer, S J Hale, J De Ridder, I W Roxburgh, V Sangaralingam, I R Stevens, M D Suran, N J Tarrant, and A Weiss. SOLAR-LIKE OSCILLATIONS IN LOW-LUMINOSITY RED GIANTS: FIRST RESULTS FROM KEPLER. *The Astrophysical Journal Letters*, 713:176–181, 2010. doi: 10.1088/2041-8205/713/2/L176.

Timothy R Bedding, Benoit Mosser, Daniel Huber, Josefina Montalb, Paul Beck, Yvonne P Elsworth, Rafael A Garc, Andrea Miglio, Timothy R White, Joris De Ridder, Saskia Hekker, Conny Aerts, and Caroline Barban. Gravity modes as away to distinguish between hydrogen- and helium-burning red giant stars. *Nature*, 471:608–611, 2011.

T. M. Brown, R. L. Gilliland, R. W. Noyes, and L. W. Ramsey. Detection of possible p-mode oscillations on procyon. *The Astrophysical Journal*, 368:599–609, feb 1991. doi: 10.1086/169725.

T. M. Brown, D. W. Latham, M. E. Everett, and G. A. Esquerdo. Kepler input catalog: Photometric calibration and stellar classification. *The Astrophysical Journal*, 142:112, oct 2011. doi: 10.1088/0004-6256/142/4/112.

W J Chaplin, G Houdek, C Karoff, Y Elsworth, and R New. Mode lifetimes of stellar oscillations Implications for asteroseismology. *A&A*, 500:21–24, 2009. doi: 10.1051/0004-6361/200911952.

W J Chaplin, H Kjeldsen, T R Bedding, J Christensen-Dalsgaard, R L Gilliland, S D Kawaler, T Appourchaux, Y Elsworth, R A García, G Houdek, C Karoff, T S Metcalfe, J Molenda, Zakowicz, M J P F G Monteiro, M J Thompson, G A Verner, S J Hale, R Handberg, S Hekker, D Huber, S Mathur, B Mosser, D Stello, and M D Suran. PREDICTING THE DETECTABILITY OF OSCILLATIONS IN SOLAR-TYPE STARS OBSERVED BY KEPLER. *The Astrophysical Journal*, 732:54–9, 2011. doi: 10.1088/0004-637X/732/1/54.

W J Chaplin, H Kjeldsen, J Christensen-Dalsgaard, R L Gilliland, S D Kawaler, S Basu, J De Ridder, D Huber, T Arentoft, J Schou, R A García, T S Metcalfe, K Brogaard, T L Campante, Y Elsworth, A Miglio, T Appourchaux, T R Bedding, S Hekker, G Houdek, C Karoff, J Molenda, Zakowicz, M J P F G Monteiro, V Silva Aguirre, D Stello, W Ball, P G Beck, A C Birch, and D L Buzasi. Asteroseismology of Solar-Like Oscillators in a 2-Wheel Mission. *ArXiv E-prints*, 2013.

William J. Chaplin, Yvonne Elsworh, Rachel Howe, George R. Isaak, Clive P. McLeod, Brek A. Miller, H. B. Van Der Raay, Sarah J. Wheeler, and Roger New. BiSON performance. *Solar Physics*, 168(1):1–18, sep 1996. doi: 10.1007/bf00145821. URL `https://doi.org/10.1007/bf00145821`.

J. Christensen-Dalsgaard. Asteroseismology. *Astrophysics and Space Science*, 261:1–12, 1998. doi: 10.1023/A:1002067301259.

J. Christensen-Dalsgaard, W. Dappen, S. V. Ajukov, E. R. Anderson, H. M. Antia, S. Basu, V. A. Baturin, G. Berthomieu, B. Chaboyer, S. M. Chitre, A. N. Cox, P. Demarque, J. Donatowicz, W. A. Dziembowski, M. Gabriel, D. O. Gough, D. B. Guenther, J. A. Guzik, J. W. Harvey, F. Hill, G. Houdek, C. A. Iglesias, A. G. Kosovichev, J. W. Leibacher, P. Morel, C. R. Proffitt, J. Provost, J. Reiter, E. J. Rhodes, Jr., F. J. Rogers, I. W. Roxburgh, M. J. Thompson, and R. K. Ulrich. The Current State of Solar Modeling. *Science*, 272:1286–1292, May 1996. doi: 10.1126/science.272.5266.1286.

Jørgen Christensen-Dalsgaard. *Lecture Notes on Stellar Oscillations*. Aarhus, 2003. URL `http://astro.phys.au.dk/jcd/oscilnotes/`.

Enrico Corsaro, Dennis Stello, Daniel Huber, Timothy R. Bedding, Alfio Bonanno, Karsten Brogaard, Thomas Kallinger, Othman Benomar, Timothy R. White, Benoit Mosser, Sarbani Basu, William J. Chaplin, Jrgen Christensen-Dalsgaard, Yvonne P. Elsworth, Rafael A. Garca, Saskia Hekker, Hans Kjeldsen, Savita Mathur, Sren Meibom, Jennifer R. Hall, Khadeejah A. Ibrahim, and Todd C. Klaus. Asteroseismology of the open clusters ngc 6791, ngc 6811, and ngc 6819 from 19 months of kepler photometry. *The Astrophysical Journal*, 757(2):190, 2012. URL `http://stacks.iop.org/0004-637X/757/i=2/a=190`.

D. S. P. Dearborn, J. C. Lattanzio, and P. P. Eggleton. Three-dimensional Numerical Experimentation on the Core Helium Flash of Low-Mass Red Giants. *The Astrophysical Journal*, 639:405–415, March 2006. doi: 10.1086/499263.

S. Deheuvels, J. Ballot, P. G. Beck, B. Mosser, R. Østensen, R. A. García, and M. J. Goupil. Seismic evidence for a weak radial differential rotation in intermediate-mass core helium burning stars. *Astronomy and Astrophysics*, 580:A96, August 2015. doi: 10.1051/0004-6361/201526449.

A. Duquennoy and M. Mayor. Multiplicity among solar-type stars in the solar neighbourhood. II - Distribution of the orbital elements in an unbiased sample. *Astronomy and Astrophysics*, 248:485–524, August 1991.

Arthur S. Eddington. *The Internal Constitution of the Stars*. Cambridge Science Classics. Cambridge University Press, 1926. doi: 10.1017/CBO9780511600005.

Y. Elsworth, S. Hekker, S. Basu, and G. R. Davies. A new method for the asteroseismic determination of the evolutionary state of red-giant stars. *MNRAS*, 466:3344–3352, April 2017. doi: 10.1093/mnras/stw3288.

L. Eyer, P. Dubath, N. Mowlavi, P. North, A. Triaud, F. Barblan, C. Siopis, L. Guy, B. Tingley, S. Zucker, D. W. Evans, L. Wyrzykowski, M. Süveges, and Z. Ivezic. The Impact of Gaia and LSST on Binaries and Exoplanets. In M. T. Richards and I. Hubeny, editors, *From Interacting Binaries to Exoplanets: Essential Modeling Tools*, volume 282 of *IAU Symposium*, pages 33–40, April 2012. doi: 10.1017/S1743921311026822.

D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125:306, March 2013. doi: 10.1086/670067.

J. H. J. de Bruijne A. G. A. Brown A. Vallenari C. Babusiaux C. A. L. Bailer-Jones U. Bastian M. Biermann D. W.Evans et al. Gaia Collaboration, T. Prusti. The gaia mission. *Astronomy and Astrophysics*, 595, 2016.

R A García, S Hekker, D Stello, J Gutiérrez-Soto, R Handberg, D Huber, C Karoff, K Uytterhoeven, T Appourchaux, W J Chaplin, Y Elsworth, S Mathur, J Ballot, J Christensen-Dalsgaard, R L Gilliland, G Houdek, J M Jenkins, H Kjeldsen, S Mccauliff, T Metcalfe, C K Middour, J Molenda-Zakowicz, M J P F G Monteiro, J C Smith, and M J Thompson. Preparation of Kepler light curves for asteroseismic analyses. *Mon. Not. R. Astron. Soc*, 414:6–10, 2011. doi: 10.1111/j.1745-3933.2011.01042.x.

R. A. García, F. Pérez Hernández, O. Benomar, V. Silva Aguirre, J. Ballot, G. R. Davies, G. Doğan, D. Stello, J. Christensen-Dalsgaard, G. Houdek, F. Lignières, S. Mathur, M. Takata, T. Ceillier, W. J. Chaplin, S. Mathis, B. Mosser, R. M. Ouazzani, M. H. Pinsonneault, D. R. Reese, C. Régulo, D. Salabert, M. J. Thompson, J. L. van Saders, C. Neiner, and J. De Ridder. Study of KIC 8561221 observed by Kepler: an early red giant showing depressed dipolar modes. *Astronomy and Astrophysics*, 563:A84, March 2014. doi: 10.1051/0004-6361/201322823.

R. L. Gilliland, J. M. Jenkins, W. J. Borucki, S. T. Bryson, D. A. Caldwell, B. D. Clarke, J. L. Dotson, M. R. Haas, J. Hall, T. Klaus, D. Koch, S. McCauliff, E. V. Quintana, J. D. Twicken, and J. E. van Cleve. Initial Characteristics of Kepler Short Cadence Data. *The Astrophysical Journal*, 713:L160–L163, apr 2010. doi: 10.1088/2041-8205/713/2/L160.

L. Girardi. A secondary clump of red giant stars: why and where. *Mon. Not. R. Astron. Soc*, 308:818–832, sep 1999. doi: 10.1046/j.1365-8711.1999.02746.x.

L Girardi, M A T Groenewegen, E Hatziminaoglou, and L Da Costa. Astronomy & Astrophysics Star counts in the Galaxy Simulating from very deep to very shallow photometric surveys with the TRILEGAL code. *Astronomy and Astrophysics*, 436: 895–915, 2005. doi: 10.1051/0004-6361:20042352.

J. Harvey. High-resolution helioseismology. In E. Rolfe and B. Battrick, editors, *Future Missions in Solar, Heliospheric & Space Plasma Physics*, volume 235 of *ESA Special Publication*, June 1985.

S. Hekker, R. L. Gilliland, Y. Elsworth, W. J. Chaplin, J. De Ridder, D. Stello, T. Kallinger, K. A. Ibrahim, T. C. Klaus, and J. Li. Characterization of red giant stars in the public kepler data. *MNRAS*, 414:2594–2601, July 2011. doi: 10.1111/j.1365-2966.2011.18574.x.

S Hekker, Y Elsworth, B Mosser, T Kallinger, W J Chaplin, J De Ridder, R A García, D Stello, B D Clarke, J R Hall, and K A Ibrahim. Astrophysics Solar-like oscillations in red giants observed with Kepler: influence of increased timespan on global oscillation parameters. *Astronomy and Astrophysics*, 544, 2012. doi: 10.1051/0004-6361/201219328.

G Houdek, N J Balmforth, J Christensen-Dalsgaard, and D O Gough. Amplitudes of stochastically excited oscillations in main-sequence stars. *Astronomy and Astrophysics*, 351:582–596, 1999.

S. B. Howell, C. Sobeck, M. Haas, M. Still, T. Barclay, F. Mullally, J. Troeltzsch, S. Aigrain, S. T. Bryson, D. Caldwell, W. J. Chaplin, W. D. Cochran, D. Huber, G. W. Marcy, A. Miglio, J. R. Najita, M. Smith, J. D. Twicken, and J. J. Fortney. The K2 Mission: Characterization and Early Results. *Publications of the Astronomical Society of the Pacific*, 126:398, April 2014. doi: 10.1086/676406.

D Huber, T R Bedding, D Stello, B Mosser, S Mathur, T Kallinger, S Hekker, Y P Elsworth, D L Buzasi, J De Ridder, R L Gilliland, H Kjeldsen, W J Chaplin, R A García, S J Hale, H L Preston, T R White, W J Borucki, J Christensen-Dalsgaard, B D Clarke, J M Jenkins, and D Koch. ASTEROSEISMOLOGY OF RED GIANTS FROM THE FIRST FOUR MONTHS OF KEPLER DATA: GLOBAL OSCILLATION PARAMETERS FOR 800 STARS. *The Asrophysical Journal*, 723:1607–1617, 2010.

D Huber, T R Bedding, D Stello, S Hekker, S Mathur, B Mosser, G A Verner, and A Bonanno. TESTING SCALING RELATIONS FOR SOLAR-LIKE OSCILLATIONS FROM THE MAIN SEQUENCE TO RED GIANTS USING KEPLER DATA. *The Astrophysical Journal*, 743(2):10, 2011.

Jon M. Jenkins, Douglas A. Caldwell, Hema Chandrasekaran, Joseph D. Twicken, Stephen T. Bryson, Elisa V. Quintana, Bruce D. Clarke, Jie Li, Christopher Allen, Peter Tenenbaum, Hayley Wu, Todd C. Klaus, Jeffrey Van Cleve, Jessie A. Dotson, Michael R. Haas, Ronald L. Gilliland, David G. Koch, and William J. Borucki. INITIAL CHARACTERISTICS OF ¡i¿KEPLER¡/i¿ LONG CADENCE DATA FOR DETECTING TRANSITING PLANETS. *The Astrophysical Journal*, 713(2), 2010. doi: 10.1088/2041-8205/713/2/L120.

T. Kallinger, S. Hekker, B. Mosser, J. De Ridder, T. R. Bedding, Y. P. Elsworth, M. Gruberbauer, D. B. Guenther, D. Stello, S. Basu, R. A. García, W. J. Chaplin, F. Mullally, M. Still, and S. E. Thompson. Evolutionary influences on the structure of red-giant acoustic oscillation spectra from 600d of Kepler observations. *Astronomy and Astrophysics*, 541:A51, May 2012. doi: 10.1051/0004-6361/201218854.

T Kallinger, J De Ridder, S Hekker, S Mathur, B Mosser, and M Gruberbauer. The connection between stellar granulation and oscillation as seen by the Kepler mission. *Astronomy and Astrophysics*, 2014.

T. Kallinger, S. Hekker, R. A. Garcia, D. Huber, and J. M. Matthews. Precise stellar surface gravities from the time scales of convectively driven brightness variations. *Science Advances*, 2:e1500654–e1500654, January 2016. doi: 10.1126/sciadv.1500654.

H Kjeldsen and T R Bedding. Amplitudes of solar-like oscillations: a new scaling relation. *Astronomy and Astrophysics*, 529:L8, 2011. doi: 10.1051/0004-6361/201116789.

Hans Kjeldsen and Timothy R. Bedding. Amplitudes of Stellar Oscillations: The implications for asteroseismology. *Astronomy and Astrophysics*, 293:87–106, 1995.

D.G. Koch, W.J. Borucki, G. Basri, N.M. Batalha, T.M. Brown, D. Caldwell, J. Christensen-Dalsgaard, W.D. Cochran, E. DeVore, E.W. Dunham, T.N. Gautier, III, J.C. Geary, R.L. Gilliland, A. Gould, J. Jenkins, Y. Kondo, D.W. Latham, J.J. Lissauer, G. Marcy, D. Monet, D. Sasselov, A. Boss, D. Brownlee, J. Caldwell, A.K. Dupree, S.B. Howell, H. Kjeldsen, S. Meibom, D. Morrison, T. Owen, H. Reitsema, J. Tarter, S.T. Bryson, J.L. Dotson, P. Gazis, M.R. Haas, J. Kolodziejczak, J.F. Rowe, J.E. Van Cleve, C. Allen, H. Chandrasekaran, B.D. Clarke, J. Li, E.V. Quintana, P. Tenenbaum, J.D. Twicken, and H. Wu. Kepler mission design, realized photometric performance, and early science. *The Astrophysical Journal*, 713:L79, apr 2010. doi: 10.1088/2041-8205/713/2/L79.

M.L. Kutner. *Astronomy: A Physical Perspective*. Cambridge University Press, 2003. ISBN 9780521529273. URL `https://books.google.co.uk/books?id=2QVmiMW0O0MC`.

Charles J. Lada. Stellar multiplicity and the initial mass function: Most stars are single. *The Astrophysical Journal Letters*, 640(1):L63, 2006. URL `http://stacks.iop.org/1538-4357/640/i=1/a=L63`.

Robert Leighton, Robert Noyes, and George Simon. Velocity Fields In The Solar Atmosphere I. Preliminary Report. *The Astrophysical Journal*, 135, 1961.

N. R. Lomb. Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39:447—-46, 1976. doi: 10.1007/BF00648343.

M. N. Lund, R. Handberg, G. R. Davies, W. J. Chaplin, and C. D. Jones. K2P2 A Photometry Pipeline for the K2 Mission. *The Astrophysical Journal*, 806:30, June 2015. doi: 10.1088/0004-637X/806/1/30.

S Mao. Lecture notes for physics of stellar structure, 2003. URL `http://www.jb.man.ac.uk/ smao/starHtml/stellarEquation.pdf`.

S Mathur, S Hekker, R Trampedach, J Ballot, T Kallinger, D Buzasi, R A García, D Huber, A Jiménez, B Mosser, T R Bedding, Y Elsworth, C Régulo, D Stello, W J Chaplin, J De Ridder, S J Hale, K Kinemuchi, H Kjeldsen, F Mullally, and S E Thompson. Granulation in Red Giants: Observations by the Kepler Mission and Three-dimensional Convection Simulations. *The Astrophysical Journal*, 741(2), 2011. doi: 10.1088/0004-637X/741/2/119.

E Michel, R Samadi, F Baudin, C Barban, T Appourchaux, and M Auvergne. Astronomy & Astrophysics Intrinsic photometric characterisation of stellar oscillations and granulation Solar reference values and CoRoT response functions. *Astronomy and Astrophysics*, 495:979–987, 2009. doi: 10.1051/0004-6361:200810353.

A. Miglio, W. J. Chaplin, R. Farmer, U. Kolb, L. Girardi, Y. Elsworth, T. Appourchaux, and R. Handberg. Prospects for detecting asteroseismic binaries in kepler data. *The Astrophysical Journal Letters*, 784(1):L3, 2014. URL `http://stacks.iop.org/2041-8205/784/i=1/a=L3`.

A. Miglio, W. J. Chaplin, K. Brogaard, M. N. Lund, B. Mosser, G. R. Davies, R. Handberg, A. P. Milone, A. F. Marino, D. Bossini, Y. P. Elsworth, F. Grundahl, T. Arentoft, L. R. Bedin, T. L. Campante, J. Jessen-Hansen, C. D. Jones, J. S. Kuszlewicz, L. Malavolta, V. Nascimbeni, and E. L. Sandquist. Detection of solar-like oscillations in relics of the Milky Way: asteroseismology of K giants in M4 using data from the NASA K2 mission. *MNRAS*, 461:760–765, September 2016. doi: 10.1093/mnras/stw1555.

B Mosser and T Appourchaux. On detecting the large separation in the autocorrelation of stellar oscillation times series. *Astronomy and Astrophysics*, 508:877–887, 2009. doi: 10.1051/0004-6361/200912944.

B Mosser, K Belkacem, M.-J Goupil, A Miglio, T Morel, C Barban, F Baudin, S Hekker, R Samadi, J De Ridder, W Weiss, M Auvergne, and A Baglin. Astronomy & Astrophysics Red-giant seismic properties analyzed with CoRoT. *Astronomy and Astrophysics*, 517, 2010. doi: 10.1051/0004-6361/201014036.

B. Mosser, K. Belkacem, M J. Goupil, E. Michel, Y. Elsworth, C. Barban, T. Kallinger, S. Hekker, J. De Ridder, R. Samadi, F. Baudin, F.J.G. Pinheiro, M. Auvergne, A. Baglin, and C. Catala. The universal red-giant oscillation pattern. an automated determination with corot data. *Astronomy and Astrophysics*, 525:L9, jan 2011. doi: 10.1051/0004-6361/201015440.

B Mosser, Y Elsworth, S Hekker, D Huber, T Kallinger, S Mathur, K Belkacem, M J Goupil, R Samadi, C Barban, T R Bedding, W J Chaplin, R A García, D Stello, J De Ridder, C K Middour, R L Morris, and E V Quintana. Characterization of the power excess of solar-like oscillations in red giants with Kepler. *Astronomy and Astrophysics*, 537, 2012a. doi: 10.1051/0004-6361/201117352.

B Mosser, M J Goupil, K Belkacem, E Michel, D Stello, J P Marques, Y Elsworth, C Barban, P G Beck, T R Bedding, J De Ridder, R A García, S Hekker, T Kallinger, R Samadi, M C Stumpe, T Barclay, and C J Burke. Astrophysics Probing the core structure and evolution of red giants using gravity-dominated mixed modes observed with Kepler. *Astronomy and Astrophysics*, 540, 2012b. doi: 10.1051/0004-6361/201118519.

B. Mosser, O. Benomar, K. Belkacem, M. J. Goupil, N. Lagarde, E. Michel, Y. Lebreton, D. Stello, M. Vrard, C. Barban, T. R. Bedding, S. Deheuvels, W. J. Chaplin, J. De Ridder, Y. Elsworth, J. Montalban, A. Noels, R. M. Ouazzani, R. Samadi, T. R. White, and H. Kjeldsen. Mixed modes in red giants: a window on stellar evolution. *Astronomy and Astrophysics*, 572:L5, December 2014. doi: 10.1051/0004-6361/201425039.

T. S. H. North, W. J. Chaplin, R. L. Gilliland, D. Huber, T. L. Campante, R. Handberg, M. N. Lund, D. Veras, J. S. Kuszlewicz, and W. M. Farr. A simple model to describe intrinsic stellar noise for exoplanet detection around red giants. *MNRAS*, 465:1308–1315, February 2017. doi: 10.1093/mnras/stw2782.

M. H. Pinsonneault, Y. Elsworth, C. Epstein, S. Hekker, S. Mészáros, W. J. Chaplin, J. A. Johnson, R. A. García, J. Holtzman, S. Mathur, A. García Pérez, V. Silva Aguirre, L. Girardi, S. Basu, M. Shetrone, D. Stello, C. Allende Prieto, D. An, P. Beck, T. C. Beers, D. Bizyaev, S. Bloemen, J. Bovy, K. Cunha, J. De Ridder, P. M. Frinchaboy, D. A. García-Hernández, R. Gilliland, P. Harding, F. R. Hearty, D. Huber, I. Ivans, T. Kallinger, S. R. Majewski, T. S. Metcalfe, A. Miglio, B. Mosser, D. Muna, D. L. Nidever, D. P. Schneider, A. Serenelli, V. V. Smith, J. Tayar, O. Zamora, and G. Zasowski. The APOKASC Catalog: An Asteroseismic and Spectroscopic Joint Survey of Targets in the Kepler Fields. *The Astrophysical Journal Supplement*, 215:19, December 2014. doi: 10.1088/0067-0049/215/2/19.

R Samadi, H.-G Ludwig, K Belkacem, M J Goupil, and M.-A Dupret. The CoRoT target HD 49933 I. Effect of the metal abundance on the mode excitation rates. *Astronomy and Astrophysics*, 509:A15, 2010. doi: 10.1051/0004-6361/200911867.

R. W. Slawson, A. Prša, W. F. Welsh, J. A. Orosz, M. Rucker, N. Batalha, L. R. Doyle, S. G. Engle, K. Conroy, J. Coughlin, T. A. Gregg, T. Fetherolf, D. R. Short, G. Windmiller, D. C. Fabrycky, S. B. Howell, J. M. Jenkins, K. Uddin, F. Mullally, S. E. Seader, S. E. Thompson, D. T. Sanderfer, W. Borucki, and D. Koch. Kepler Eclipsing Binary Stars. II. 2165 Eclipsing Binaries in the Second Data Release. *The Astronomical Journal*, 142:160, November 2011. doi: 10.1088/0004-6256/142/5/160.

D. Stello, M. Cantiello, J. Fuller, D. Huber, R. A. García, T. R. Bedding, L. Bildsten, and V. Silva Aguirre. A prevalence of dynamo-generated magnetic fields in the cores of intermediate-mass stars. *Nature*, 529:364–367, January 2016. doi: 10.1038/nature16171.

D. Stello, J. Zinn, Y. Elsworth, R. A. Garcia, T. Kallinger, S. Mathur, B. Mosser, S. Sharma, W. J. Chaplin, G. Davies, D. Huber, C. D. Jones, A. Miglio, and V. Silva Aguirre. The K2 Galactic Archaeology Program Data Release. I. Asteroseismic Results from Campaign 1. *The Astrophysical Journal*, 835:83, January 2017. doi: 10.3847/1538-4357/835/1/83.

Dennis Stello, Daniel Huber, Thomas Kallinger, Sarbani Basu, Benot Mosser, Saskia Hekker, Savita Mathur, Rafael A García, Timothy R Bedding, Hans Kjeldsen, Ronald L Gilliland, Graham A Verner, William J Chaplin, Othman Benomar, Søren Meibom, Frank Grundahl, Yvonne P Elsworth, Joanna Molenda, Zakowicz, Robert Sza O 15, Jørgen Christensen-Dalsgaard, Peter Tenenbaum, Joseph D Twicken, and Kamal Uddin. AMPLITUDES OF SOLAR-LIKE OSCILLATIONS: CONSTRAINTS FROM RED GIANTS IN OPEN CLUSTERS OBSERVED BY KEPLER. *The Astrophysical Journal Letters*, 737:L10, 2011. doi: 10.1088/2041-8205/737/1/L10.

Dennis Stello, Daniel Huber, Timothy R Bedding, Othman Benomar, Lars Bildsten, Yvonne P Elsworth, Ronald L Gilliland, Benot Mosser, Bill Paxton, and Timothy R White. ASTEROSEISMIC CLASSIFICATION OF STELLAR POPULATIONS AMONG 13,000 RED GIANTS OBSERVED BY KEPLER. *The Astrophysical Journal Letters*, 765:41–5, 2013. doi: 10.1088/2041-8205/765/2/L41.

Y Taroyan. Lecture notes for course ph38310: The solar interior, 2010.

R. K. Ulrich. The Five-Minute Oscillations on the Solar Surface. *The Astrophysical Journal*, 162:993, December 1970. doi: 10.1086/150731.

F. van Leeuwen. Validation of the new Hipparcos reduction. *Astronomy and Astrophysics*, 474:653–664, November 2007. doi: 10.1051/0004-6361:20078357.

G. A. Verner, Y. Elsworth, W. J. Chaplin, T. L. Campante, E. Corsaro, P. Gaulme, S. Hekker, D. Huber, C. Karoff, S. Mathur, B. Mosser, T. Appourchaux, J. Ballot, T. R. Bedding, A. Bonanno, A. M. Broomhall, R. A. García, R. Handberg, R. New, D. Stello, C. Regulo, I. W. Roxburgh, D. Salabert, T. R. White, D. A. Caldwell, J. L. Christiansen, and M. N. Fanelli. Global asteroseismic properties of solar-like oscillations observed by Kepler: A comparison of complementary analysis methods. *Monthly Notices of the Royal Astronomical Society*, 415(4):3539–3551, 2011. doi: 10.1111/j.1365-2966.2011.18968.x.