

# Investigating the prospects of observing asteroseismic effects in M and K dwarfs with PLATO

M. Tomes      V. Arena      L. Booth      S. Khan-Bennett      D. Lanning  
S. Burton-Weiss      D. Hart      M. Osborne      C. Gnanavel      J. Chow  
P. Ramsden      E. Penny      K. Hinds      S. Shrestha.

March 2020



School of Physics and Astronomy  
University of Birmingham  
United Kingdom

## Abstract

Attempts were made to place a quantitative limit on the detection of acoustic oscillation modes in the late main-sequence of K and M dwarf stars to be observed by the European Space Agency's PLATO mission. An extensive catalogue of frequency-power spectra was successfully modelled for K and M dwarf stars using simulated data from TRILEGAL. Detection techniques were tested on the modelled spectra; these were a combination of manual and automatic Gaussian fitting and power-integration methods. From the 40 brightest of each modelled K and M dwarfs, oscillation modes were only successfully detected for the brightest K dwarf with magnitude 7.319. It was concluded that PLATO would only be able to observe oscillation modes in stars as bright or brighter than this. The detection methods could be refined to potentially improve this limit.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Introduction: TSM Group (Maddie O)</b>	<b>7</b>
<b>3</b>	<b>Target Selection (Chamini)</b>	<b>7</b>
3.1	TRILEGAL (Chamini) . . . . .	8
3.2	K and M dwarfs (Chamini) . . . . .	8
3.3	Plotting HR and CM Diagrams (Chamini) . . . . .	10
3.3.1	Binary Systems (Jeff) . . . . .	11
<b>4</b>	<b>Oscillations</b>	<b>11</b>
4.1	Convection (Sharif) . . . . .	11
4.2	Granulation (Sharif) . . . . .	13
4.3	P-Modes (Sharif and Victoire) . . . . .	15
<b>5</b>	<b>Components of the spectrum</b>	<b>16</b>
5.1	Introduction to scaling relations (Sharif and Maddie O) . . . . .	16
5.2	The Frequency of Maximum Power (Victoire) . . . . .	17
5.3	The Large Frequency Separation (Victoire) . . . . .	18
5.4	The Small Frequency Separation, D (Jeff and Sam) . . . . .	18
5.5	The Gaussian Envelope (Maddie O and Jeff) . . . . .	19
5.5.1	Amplitude (Maddie O) . . . . .	19
5.5.2	Width (Jeff) . . . . .	21
5.6	The Lorentzian peaks . . . . .	22
5.6.1	Definition (Maddie T and Victoire) . . . . .	22
5.6.2	The width of the peaks (Victoire) . . . . .	23
5.6.3	The height of the peaks (Maddie T) . . . . .	26
5.7	D (Maddie O and Sam) . . . . .	28
5.8	Epsilon (Sam) . . . . .	29
5.9	Granulation (Sharif) . . . . .	31
5.10	Cadence, Radial Order (Sharif) . . . . .	33
5.11	Shot noise (Jeff) . . . . .	33
<b>6</b>	<b>Models/Results</b>	<b>35</b>
6.1	The Sun Spectrum (Maddie T) . . . . .	35
6.2	Models of K dwarfs (Maddie T and Victoire) . . . . .	36
6.3	Models of M dwarfs (Chamini) . . . . .	40
6.4	Motivation for the targets selected (Chamini) . . . . .	43
6.5	Check of the spectrum (Maddie O) . . . . .	45
<b>7</b>	<b>Discussion</b>	<b>46</b>
7.1	Approximations Made (Maddie T) . . . . .	46
7.2	Satellite Jitter (Sam) . . . . .	47

<b>8</b>	<b>Intro DA (Luke)</b>	<b>48</b>
<b>9</b>	<b>Theory (Luke)</b>	<b>48</b>
9.1	How can detections be made . . . . .	48
9.2	Parseval's Theorem . . . . .	50
9.3	Importance of preserving and extracting measurables from successful detections . . . . .	51
<b>10</b>	<b>Noise (Dan H)</b>	<b>51</b>
10.1	Statistical Noise . . . . .	51
10.2	Effects of Shot Noise . . . . .	53
<b>11</b>	<b>Binning (Paige and Dan L)</b>	<b>54</b>
11.1	Binning Basics (Dan L) . . . . .	54
11.2	Binning Exponential Statistics (Dan L) . . . . .	54
11.3	Illustration of Binning (Paige) . . . . .	55
<b>12</b>	<b>Gaussian Fitting (Edward and Dan H)</b>	<b>56</b>
12.1	Curve Fitting Processes (Dan H) . . . . .	56
12.2	Gaussian Fitting Methods . . . . .	57
12.2.1	Multiple Generic Gaussian Fitting (Dan H) . . . . .	57
12.2.2	Single Generic Gaussian Fitting with Cut-off (Dan H) . . . . .	58
12.3	Background fitting and Gaussian fitting (Edward) . . . . .	60
12.4	Finding $\nu_{\max}$ (Edward) . . . . .	61
12.5	Temperature method (Edward) . . . . .	64
<b>13</b>	<b>Power Integration Fitting (Paige and Dan L)</b>	<b>64</b>
13.1	Fitting to the Background and P-modes (Dan L) . . . . .	64
13.2	Initial Investigations of Power Ratios (Dan L) . . . . .	65
13.3	Optimising the fit using Sliders (Paige) . . . . .	66
13.4	Increasing the Shot Noise Level (Paige) . . . . .	69
13.5	Using Scaling Relations as an Estimate to Check for Congruent Results (Dan L) . . . . .	70
13.6	Power Ratio Evaluation (Paige) . . . . .	72
<b>14</b>	<b>Detection Thresholds (San)</b>	<b>73</b>
14.1	$\sigma$ test . . . . .	73
14.2	Varying the Gaussian: 'Plateau' test and Grid Stars comparison . . . . .	74
14.3	Power Integration Ratio . . . . .	75
<b>15</b>	<b>Measuring Global Parameters (K-Ryan)</b>	<b>75</b>
15.1	Measuring $\nu_{\max}$ and $\Delta\nu$ and from Gaussian Fitting . . . . .	76
15.2	Measuring $\Delta\nu$ for Solar-Like Spectra . . . . .	77
<b>16</b>	<b>Results and Analysis (K-Ryan and San)</b>	<b>79</b>
16.1	Shot noise & Magnitude for the catalogue stars' range (San) . . . . .	79
16.2	$1\sigma$ test observations (San) . . . . .	79
16.2.1	K dwarfs (San) . . . . .	79

16.2.2	M dwarfs (San) . . . . .	83
16.3	Grid Stars (K-Ryan) . . . . .	83
16.4	Temperature from fitted $\nu_{\max}$ values (K-Ryan and San) . . . . .	84
16.5	Power Integration Ratio (San) . . . . .	86
16.6	$\Delta\nu$ Results (K-Ryan) . . . . .	86
16.7	Overall Findings (K-Ryan and San) . . . . .	87
<b>17</b>	<b>Discussion of Results (Luke)</b>	<b>88</b>
17.1	Implications for the ESA PLATO Mission . . . . .	90
17.2	Other Considerations . . . . .	90
<b>Appendix</b>		<b>93</b>
A	Derivation of convective instability (Sharif) . . . . .	93
B	Additional figures . . . . .	94

## Acknowledgements

We would like to give thanks to Bill Chaplin for successfully organising and running this group study, your enthusiasm and dedication to asteroseismology is inspiring. We would also like to give thanks to Warrick Ball and Martin Nielsen for providing constant support and for putting up with all of our questions.

# 1 Introduction

**SBW** - The upcoming ‘PLanetary Transits and Oscillations of Stars’ (PLATO) mission is the third medium-class mission in the European Space Agency’s Cosmic Vision programme. It is due to be launched in 2026 with two primary objectives.

The first of these is the detection of seismic stellar oscillations which can be used to effectively probe the stellar interior and gain insight into stars’ internal structures.

Gauging PLATO’s ability to observe seismic oscillations in M and K dwarfs motivates this project, which will use stellar population synthesis-based modelling derived from asteroseismic Kepler data to place a visual magnitude detection limit on PLATO for observing seismic stellar oscillations. If successful, such a prediction will allow the PLATO mission to target regions of the sky containing many stars of the lowest possible mass for which oscillations will be observable. This will maximise the probability of detecting transiting Earth-like exoplanets during observation of a given patch of the sky whilst simultaneously gathering useful asteroseismic data from the largest number of M and K dwarfs possible. PLATO will build upon previous missions’ accomplishments in this regard, and thus it is important to first establish the successes and limitations of these missions.

**DL** - The earliest dedicated mission for observations of stellar oscillations using a space-based telescope was the ‘Convection, Rotation and planetary Transits’ (CoRoT) mission, launched in late 2006. The mission, equipped with a single telescope with a field of view of  $2.7^\circ \times 3.05^\circ$  ([Auvergne et al. \(2009\)](#)), observed each of its targets for 150 days in an attempt to detect seismic stellar oscillations which would facilitate accurate estimates of these stars’ fundamental properties through analysis of their frequency-power spectra. CoRoT achieved

some success here, in that the size of the convective cores could be deduced for target stars as cool as those in the F-class. However, collected data was not of the required signal-to-noise ratio to observe these oscillations in any dimmer stars. This is because observational noise scales with the visual magnitude of a target (see section 5.11), and thus detection of seismic oscillations in stars dimmer than F-type dwarfs requires a stronger observed signal than that detectable by CoRoT. The main factors limiting CoRoT in this respect were the insufficient collection area of its single telescope and the relatively short amount of time it spent observing each of its targets.

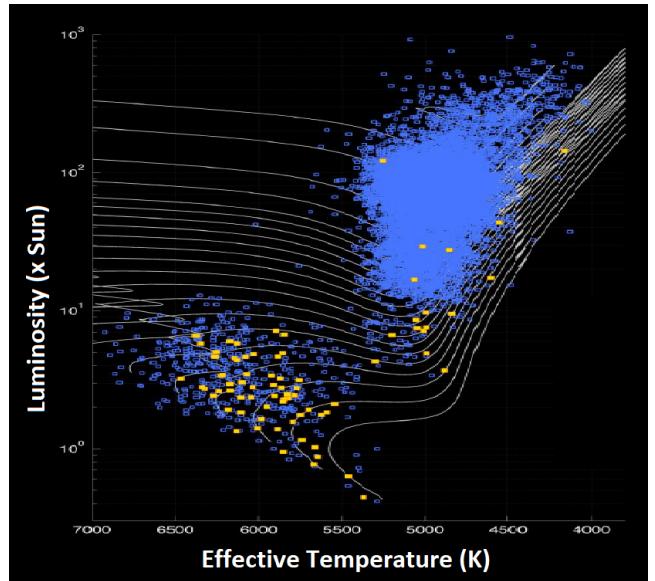


Figure 1.1: Hertzsprung-Russell diagram of Kepler stars in which seismic oscillations have been observed. Includes approximately 20000 red giant stars and 700 main-sequence stars.

Notably, a number of bright K-dwarfs are present ( $T_{\text{eff}} \leq 5000K$ ). Taken from [Chaplin and Miglio \(2013\)](#).

Following CoRoT, the next significant mission in asteroseismology was the Kepler mission, launched by NASA in March 2009. The primary aim of Kepler was to detect Earth-like exoplanets as they transit their host star. For-

tunately for asteroseismologists, however, these investigations rely on sustained observations of a single target for a period of time - much like the observations required for detecting stellar oscillations. Indeed, the telescope used in the Kepler mission was well suited for making successful observations of dim stars; it was pointed at the same patch of sky for roughly 3.5 years using a telescope with a collecting area of approximately  $0.7\text{m}^2$ , much larger than the  $0.2\text{m}^2$  diameter telescope used in the CoRoT mission [Johnson \(2017\)](#). Consequently, manipulation of the Kepler data successfully revealed seismic oscillations in over 20000 red giant stars and 700 main-sequence stars, including a small number of bright K-dwarfs (figure 1.1).

Despite this success, however, the Kepler mission had limitations. Firstly, the telescope was pointed at a small  $105\text{deg}^2$  section of the sky for the entirety of its lifetime [Barclay \(2020\)](#). This means that, while the collected data is of high quality, asteroseismologists are still keen to discover the properties of similar stars that are situated in the rest of the sky. Additionally, Kepler's ability to resolve oscillations in even dimmer stars was limited by a number of factors. Most notably, its 60s short cadence ([Gilliland et al. \(2010\)](#)) may have proved insufficient for detecting seismic oscillations in many dim K-type or M-type dwarfs (see section 6.3). In light of this, stellar targets were selected such that the limited number of these short cadence slots available for observing seismic oscillations were used most effectively. This meant targeting stars of bright visual magnitude which are, in general, quite luminous.

The most recent mission which aids in the study of asteroseismology is NASA's Transiting Exoplanet Survey Satellite (TESS), launched in 2018. Similar to CoRoT, TESS observes a large portion of the sky in its attempts to iden-

tify oscillations within stars. It will, however, eventually exceed CoRoT significantly in this regard, with an expected total sky coverage of approximately 85%. This will be achieved by using four separate wide-angle cameras operating simultaneously, each with a field-of-view of  $24\text{deg}^2$  ([Chrisp et al. \(2015\)](#)). These will observe the sky in 26 sectors, with each sector being observed for approximately 27 days, as per [Barclay \(2020\)](#). While this short observing time increases sky coverage, however, it significantly diminishes TESS's ability to observe seismic oscillations in dim stars. Indeed, Kepler was able to identify oscillations in stars that are much too dim to be observed by the TESS mission due to the larger collecting power on the telescope used and the longer observing time for a single patch of sky. Asteroseismologists are currently without a comprehensive study of the sky that shows oscillations in dim stars.

**SBW** - The PLATO mission is expected to provide such a study. The 24(+2) cameras<sup>1</sup> that will be mounted to the craft will provide a total field-of-view of  $2232\text{deg}^2$  ([ESA \(2017\)](#)) which, combined with the long time period over which it will collect data, will allow for coverage of around 50% of the sky ([Rauer et al. \(2010\)](#)). Furthermore, with these cameras exhibiting a smaller cadence of 2.5s/25s, it is also possible that PLATO will observe seismic stellar oscillations in dimmer stars than previously possible. If this is the case, the ability to detect seismic oscillations in M and K dwarf stars would be significant for several reasons. Firstly, such stars are extremely common, and it is useful, therefore, to understand their internal structure (see section 3.2). More relevant to the PLATO mission, however, is that these red dwarfs are the most suitable targets for fulfilling the mission's second primary objective: the identification of Earth-like exoplanets as they transit their parent stars.

---

<sup>1</sup>PLATO will be equipped with 24 cameras with a 25s cadence, and 2 extra cameras with a 2.5s cadence that will be used for studying bright targets with  $m_v \leq 8$  ([Catala \(2011\)](#)).

Identifying Earth-like exoplanets is a task motivated by the search for extra-terrestrial life in the Universe, the emergence of which, for PLATO’s purposes, can be assumed to only occur under conditions similar to those on Earth. An exoplanet can only be considered ‘Earth-like’ if its orbit is radially positioned within the habitable zone of its host star. The planet must also be sufficiently large that it can retain an atmosphere and exhibit continuous geological activity, while being sufficiently small such that it is terrestrial in nature.

To understand why such planets are more likely to be detected in orbit around red dwarf stars, it is important to explain the factors upon which PLATO’s ability to detect Earth-like exoplanets depends. These are the probability of this planet existing in a suitable orbital plane such that its transit can be observed, and the distinguishability of the corresponding transit signal above observational noise. The former of these is somewhat random in nature, requiring the orbital plane to be almost parallel to a near-Earth observer’s line-of-sight<sup>2</sup>. The latter, however, can be optimised during the target selection process as it is dependent upon the fractional size of the dip in an observed star’s light curve as it is transited by an orbiting exoplanet, relative to the amount of noise present in the observations. It is obvious that this transit signal-to-noise ratio increases with the relative size of the exoplanet to its host star, as a larger transiting planet will cause a more significant dip in the star’s observed intensity; this, therefore, motivates the search for such planets around low-mass main-sequence stars. Furthermore, the transit of an exoplanet with a smaller orbital radius, and thus a shorter orbital period, will result in a higher signal-to-noise ratio than the transit of a more remote exoplanet. This is because shorter orbital periods allow for more

transits to be observed within a given time frame, meaning that phase-folding can be more effectively employed to smooth over the noise. As Earth-like exoplanets exist, by definition, within the habitable zone of their solar system, and the orbital radius of this habitable zone decreases with the mass of the system’s host star, it is clear that this also points to red dwarfs as the prime candidates around which orbiting Earth-like exoplanets can be detected.

As PLATO has a finite lifespan over which it can collect data, the patches of the sky which the mission will cover must be selected to most effectively fulfil both of its mission objectives: the detection of seismic stellar oscillations and the identification of Earth-like exoplanets. As established, the likelihood of detecting a transiting Earth-like exoplanet of a given size increases as the size of its observed host star decreases. However, whilst it is known that PLATO could potentially be more successful at observing seismic oscillations in red dwarfs than its predecessors, the degree to which this is the case remains to be seen.

---

<sup>2</sup>The probability of this being the case increases somewhat with the exoplanet’s orbital proximity to the host star and this host star’s size, but this is usually not significant.

## 2 Introduction: TSM Group (Maddie O)

The role of the target selection and modelling group was to compile a list of simulated targets and produce fake, noise-free data using the properties of these simulated stars to represent the real data that will be obtained. The catalogue of targets produced by this group needed to be comprised of K and M dwarf stars, the foci of this project.

The selected stars were taken from a simulated galactic population, generated by the code TRILEGAL. By using both PLATO's observational capabilities and the intrinsic properties of M and K dwarfs, the catalogues were narrowed down to a suitable selection of stars that could test the limits of PLATO.

The second task was to generate fake, noise-free data. Achieving this required the computation and simulation of the multiple components that constitute the real, frequency-power oscillation spectrum, an example of which is shown in figure 2.1. The three main components of the spectrum are the oscillations themselves, the granulation and sources of noise that may impact the visibility of the oscillations such as the shot noise. Other aspects of the spectrum that would ideally be included in a perfect model are the artefact and activity and instrumental noise shown in figure 2.1. However, these components should not affect the detectability of the oscillations, making the modelling of these features less of a priority in this project.

Once the spectra were produced, they were passed over to the data analysis (DA) group for the next stage of the project. To ensure that the spectra being produced would match real-life observations and to test initial detection capabilities, the spectrum for the Sun was created before moving on to K and M dwarfs. This would allow for the TSM group to ensure

all the components of the spectra work correctly as the spectrum for the Sun is already well-defined by previous missions.

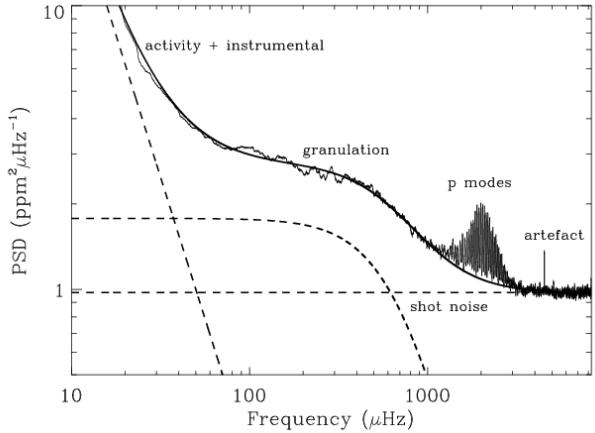


Figure 2.1: Full range of the frequency-power spectrum of Kepler-410, on a log-log scale. This shows all the components contributing to the data. ([Basu and Chaplin \(2017\)](#))

## 3 Target Selection (Chamini)

According to [ESA \(2017\)](#), PLATO will have a FoV of  $2232 \text{ deg}^2$ , which will be focused on 12 predetermined patches of the sky centred around  $30^\circ$  above and below the galactic equator ( $|b| = 30^\circ$ ), covering  $\sim 40\%$  of the sky; 2 Long-Duration (LOP) fields (with data being collected for 2-3 years each) and 10 STEP-and-stare fields (observed for a few months each). For this study, simulations of the STEP01, STEP03, and STEP08 fields were used ( $(l, b) \approx (313, -30), (13, -30), (305, +30)$ ) (figure 3.1) but analysed to the maximum duration of 2 years to yield the most number of detections, and the highest precision in the subsequent measurements ([Miglio et al. \(2017\)](#)). STEP01 and STEP08 are predicted to be comprised of  $\sim 90\%$  thin disc stars, and a small trace of bulge stars, which can be more difficult to accurately measure the brightness of, whereas STEP03 contains  $\sim 3\%$  of stars from

the galactic bulge. The range of stars is useful to study the evolution of the Milky Way and the formation of the regions they are found in. The samples were kept separately to distinguish between the different seeing conditions in those regions. For this study, no binaries were included in the simulations.

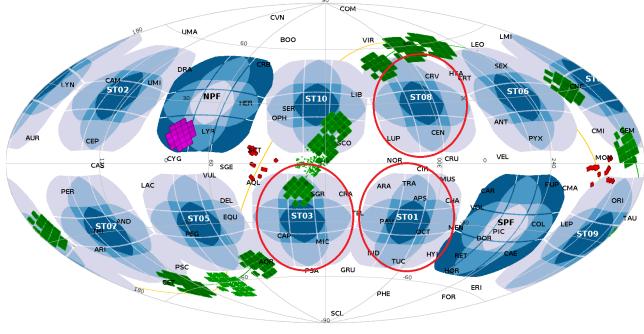


Figure 3.1: STEP01, STEP03 and STEP08

(circled in red), along with the 7 other proposed STEP-and-stare fields and 2 LOP (NPF and SPF), all located at  $|b| = 30^\circ$  about the galactic equator as shown in the galactic reference frame. Fields already observed by Kepler, K2 and CoRoT are shown in magenta, green and red respectively. Image adapted from [Miglio et al. \(2017\)](#) and [Maxted \(2018\)](#)

### 3.1 TRILEGAL (Chamini)

TRIdimensional modeL of thE GALaxy, TRILEGAL ([Girardi et al. \(2005\)](#)) was used to simulate the stars in the 3 fields. Known parameters such as galactic coordinates, age, metallicity, distance moduli, extinction coefficients, mass, luminosity and surface gravity as well as apparent magnitudes in several filters were provided in catalogues. The simulation was created to obey various key stellar relations and visibilities that are apparent in the Milky Way from PLATO's point of view. A faint magnitude limit of 15 was used in the simulations to model reality, as can be seen in figure 3.2.

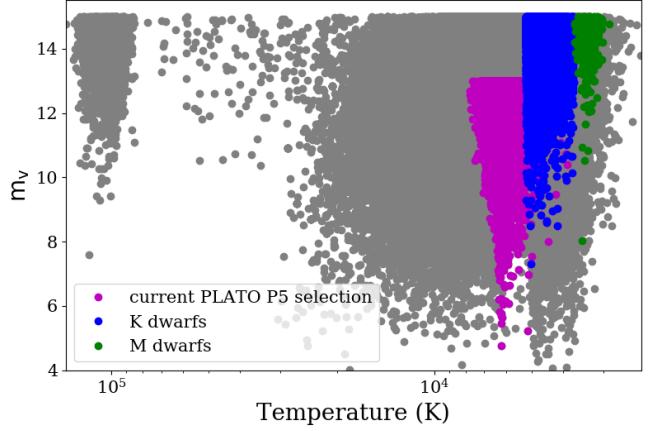


Figure 3.2: Effective Temperature of all the objects in STEP01 against their apparent visual magnitudes. Potential K dwarfs are in blue and M dwarfs in green. The cuts used to constrain the K dwarfs and M dwarfs are found in Table 1

### 3.2 K and M dwarfs (Chamini)

This report will focus on the detectability and spectra of cool K and M dwarfs. K and M dwarfs have the lowest temperatures, masses and luminosities on the main-sequence, yet have ages greater than 500 Myrs (see Table 1), distinguishing them from their pre-main-sequence counterparts. These stellar objects are particularly promising in the hunt for exoplanets being the most abundant stars in the galaxy (making up  $\sim 12\%$  and  $\sim 76\%$  of the solar neighbourhood ([Ledrew \(2001\)](#))) due to their low initial masses  $0.1 M_\odot \leq M_{\text{initial}} \leq 0.8 M_\odot$  placing them in the optimal range for the peak in Initial Mass Function (theorised distribution of stellar masses at the formation of the galaxy which is predicted to favour those with masses  $\sim 10^{-1} M_\odot$ , with the number density falling off at higher and lower masses.) Despite their abundance, due to their characteristic low luminosities and the limiting apparent magnitude,  $m_v \sim 15$ , of the sample, these dwarfs are less abundant on the current HR diagram as they are less easily observed.

According to [Cuntz and Guinan \(2016\)](#), early K dwarfs are interesting host-stars for advanced life-hosting planets; their Habitable-Zones are predicted to be close enough to receive enough radiation from the star to sustain liquid water, but not close enough to be fatally affected by solar flares. They are even predicted to be more likely to host life than Sun-like stars, due to their long main sequence lifetimes ( $\sim 10^2$  Gyrs) and stable masses, allowing for more time for planets and possible life to form in their steady habitable zones.

To characterise the properties of exoplanets, accurate estimates for the radius and mass of the host stars are vital ([Mulet-Marquis et al. \(2009\)](#)), hence the interest in the seismic analysis of K and M dwarfs. So far, early to mid K dwarfs and their exoplanets have been detected using Kepler, however, the probability of observing oscillations in M dwarfs remains uncertain.

In [ESA \(2017\)](#), 4 groups of targets P1,2,4 & 5, have been identified as categories of interest for PLATO. P5 contains dwarfs and subgiants up to  $m_v=13$ , in stellar classes between G and K7 (early K dwarfs), but neglect mid- to late K dwarfs and M dwarfs. Although M dwarfs are the subject of P4, PLATO will focus on their transits through transmitting back imagettes (small images created by just the pixels the star affects). The purpose of this study will be to produce and analyse possible spectra of K and M dwarfs to determine whether the P5 target list could be expanded on, or narrowed down, depending on the detectability of the oscillations in the spectra. Limits in telemetry (the sending and downloading of data back to Earth) imposes the need to carefully pre-select targets for the PLATO Input Catalogue (PIC) to maximise the yield of successful seismic and planetary observations.

K dwarfs	M dwarfs
$3700 \text{ K} \leq T_{\text{eff}} \leq 5200 \text{ K}$	$2400 \text{ K} \leq T_{\text{eff}} \leq 3700 \text{ K}$
$0.5M_{\odot} \leq M \leq 0.8M_{\odot}$	$M \leq 0.6M_{\odot}$
$L \leq 0.6L_{\odot}$	$L \leq 0.08L_{\odot}$
Age $\geq 500$ Myrs	Age $\geq 500$ Myrs
$4 \leq \log g \leq 7$	$4 \leq \log g \leq 7$

Table 1: Limits in the temperature, mass, luminosity, age and surface gravity used to constrain K and M dwarfs in the TRILEGAL samples, adapted from [Baraffe and Chabrier \(1996\)](#), [Habets and Heintze \(1981\)](#), [Masana \(2019\)](#) and [Montalto et al. \(2019\)](#)

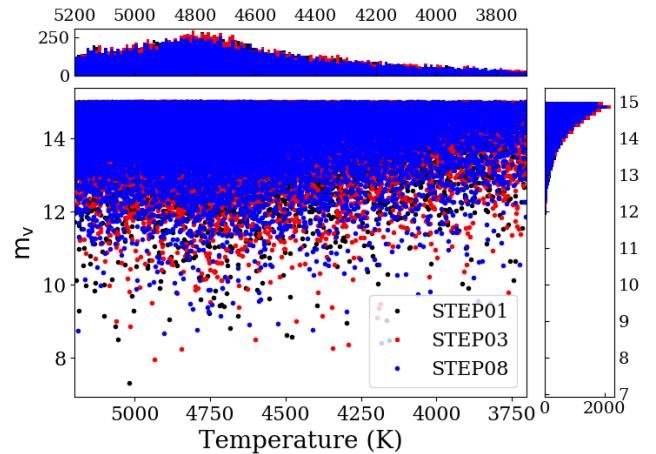


Figure 3.3: Effective Temperature of the K dwarfs (determined via Table 1) in STEP01, 03, & 08 against their apparent visual magnitudes. Histograms showing the distribution of apparent magnitudes and Temperatures show a sharp exponential decrease in the number of targets at decreasing magnitudes and a modal effective temperature of  $\sim 4800$  K.

Table 1 outlines the cuts used to define the K dwarfs and M dwarfs in the simulated TRILEGAL catalogues. Figures 3.3 and A2 show the distribution of apparent magnitudes and temperatures within K and M dwarfs.

### 3.3 Plotting HR and CM Diagrams (Chamini)

The current P5 PLATO selection of F5 to K7 spectral type sub-giants and dwarfs with  $m_v < 13$  was determined using the following cuts.

$$\begin{aligned} \text{P5 PLATO dwarf selection} = & 0.42 \leq \mathbf{B-V} \leq \\ & 1.38 \quad M_v \leq 5(\mathbf{B-V}) + 0.4 \quad \& \quad M_v \leq 5(\mathbf{B-V}) + 3.5 \\ & \& m_v \leq 13 \end{aligned}$$

(These cuts were created from RAVE DR5 to adapt the region occupied by  $\log g > 4$ ,  $4050 \text{ K} < T_{\text{eff}} < 6510 \text{ K}$ ,  $m_v \leq 13$  (Montalto et al. (2019))).

The PLATO Definition Study Report (2017) highlights the need for bright targets for asteroseismic analysis and sets a limit of  $m_v < 11$  for its most successful detections due to the lower NSR. The limits used to define the K dwarfs (in blue) and M dwarfs (in green) in figures 3.4 and 3.5 in the catalogues are outlined in Table 1. The resulting stellar populations are shown in Table 2.

	STEP01	STEP03	STEP08
Total	1007115	1170996	788081
P5 cuts	54445	57334	50064
K dwarfs	19417	20262	18579
M dwarfs	1051	985	935

Table 2: Stellar populations of the entire simulated catalogue, current P5 PLATO dwarf cuts and K and M dwarf cuts outlined in Table 1

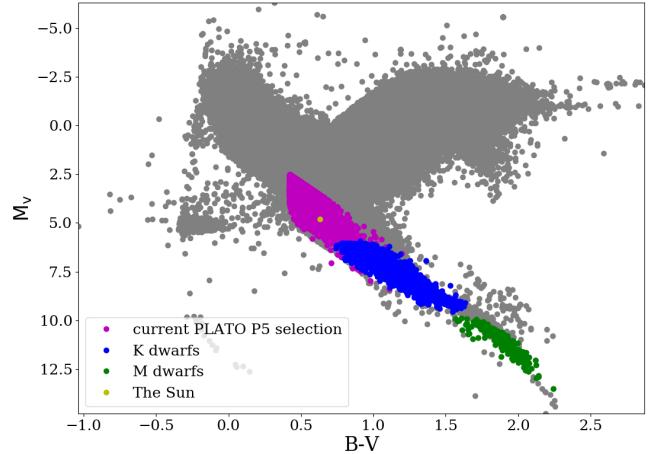


Figure 3.4: Colour-Magnitude Diagram of the TRILEGAL simulation for STEP01, with the absolute magnitudes being calculated from the extinction co-efficient,  $A_v$ , the distance modulus,  $\mu$ , and  $m_v$

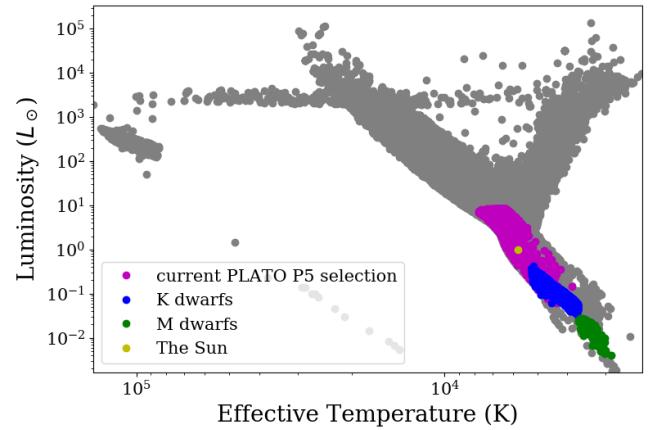


Figure 3.5: Hertzsprung-Russell diagram of the TRILEGAL simulation for STEP01, with objects satisfying the P5 PLATO cuts in magenta, criteria for K dwarfs in blue; M dwarfs in green. The approximate position of the Sun is in yellow.

As can be seen in figure 3.5, the Sun (a class G star) currently fits within the outlined P5 limits, as well as some luminous, lower temperature K dwarfs. The current selection does not contain any M dwarfs.

ESA (2017) predicts an M dwarf population  $\sim 5200 \pm 20\%$  per FoV for a limiting magnitude of  $m_v = 15$ . Although the populations in Table 2 are much lower than predicted, the more selective cuts decrease the number of anomalies and falsely identified targets.

### 3.3.1 Binary Systems (Jeff)

Binary stellar systems and how they may affect the appearance of the frequency-power spectra observed by PLATO have not been considered. Previous studies such as that conducted by Miglio et al. (2014) have attempted to shed light on the difference in binary power spectra and that of single star systems. It was found that there were two possible effects. Firstly, The two power spectra of the separate stars may overlap, with the individual oscillation peaks of the two stars being interspersed. For binary systems with a difference in the mass fraction of the two stars of  $\pm 0.5$ , the peak of the Gaussian curves for each star would lie within the FWHM of the other. It turns out that the case where the two oscillation spectra are largely overlapping (ie. the difference in their value of  $\nu_{\max}$ ) is more common. Figure 3.6 shows the probability of the two spectra of the binary system having different separations,  $\delta\nu_{\max}$ . The probability spikes at lower separations implying that overlapping spectra for binaries are more likely to be observed.

However, within Kepler's short-cadence (SC) data for 58.85 second periods, the probability that each star belongs to a binary system is only 0.3. It is predicted that the number of binaries with detectable seismic oscillations is roughly 200 which are mainly helium-burning stars. main-sequence binaries with detectable seismic oscillations are considered rare as the signal-to-noise ratios of these stars are lower than for red giants.

So in general, although binary star systems

would have altered frequency power spectra compared to single stars, the fact that oscillations are rarely detected in main-sequence binary systems means that this should not affect results when considering late main-sequence K and M dwarf stars for PLATO.

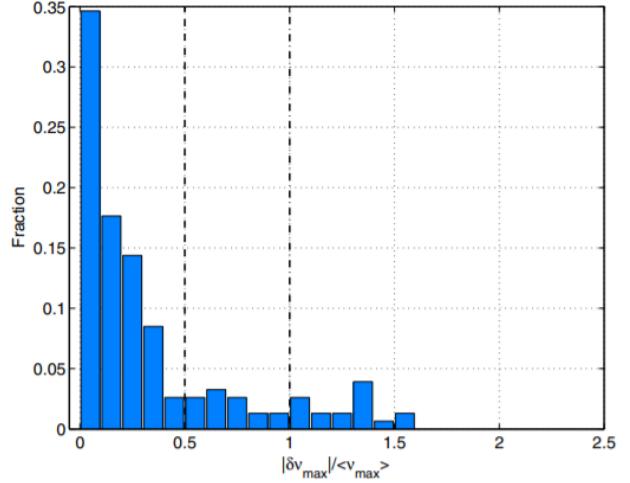


Figure 3.6: The probability of components of binary spectra having different  $\nu_{\max}$  separations.

## 4 Oscillations

### 4.1 Convection (Sharif)

The stratification of a star depends on hydrostatic equilibrium and any perturbation from this balance will be damped and restored. With this condition in place, instabilities arise as a result of the delicate balance between temperature and pressure. For example, by treating the stellar material as an ideal gas, its equation of state,

$$\rho = \frac{\mu m_u P}{k_b T} \quad (4.1)$$

shows that if the temperature drops too quickly with increasing radius,  $r$ , the density may have to increase to compensate for this and maintain the correct pressure gradient for hydrostatic equilibrium. However, a layer of higher density material sitting above a layer of lower

density material is unstable and leads to what is called convective instability. A region of the star is determined to be convectively stable or not, by applying the Schwarzschild criterion in which the motion of a perturbed bubble of gas is analysed by considering the buoyant forces acting on it. (For full deviation, see Appendix A). This reduces to the simple inequality, for which any section of the stellar radius must be convective when satisfied

$$\nabla > \nabla_{\text{ad}}. \quad (4.2)$$

This equation describes a region in which a rising bubble of gas loses energy and cools as a slower rate (less negative gradient) than the cooling of the surrounding plasma. Consequently, it maintains a greater temperature and smaller density than its surroundings, it continues to rise unhindered until new instabilities set in, such as turbulent flow causing the bubble to dissolve. This model omits a term which describes the changing molecular composition of the stellar plasma as a function of  $r$ ,  $\nabla_\mu$ . The data in TRILEGAL does not allow for the calculation of this term as it does not provide data of quantities which scale with  $r$ , nevertheless, this term is generally very small.

To calculate  $\nabla$ , it will assumed that energy is being transported completely by radiation, therefore,  $\nabla = \nabla_R$  which is given by

$$\nabla_R = \frac{3k_b}{16\pi c_s G m_u} \frac{\kappa}{\mu} \frac{l}{m} \frac{\rho}{T^3} \quad (4.3)$$

where  $c_s$  is the speed of sound in the star,  $G$  is the gravitational constant,  $l/m$  is the luminosity generated per unit mass,  $T$  is the temperature and  $\kappa$  is the opacity as given by Kramers law,  $\kappa = \kappa_0 \rho T^{-7/2}$ . Convection takes place when  $\nabla_R > \nabla_{\text{ad}}$ . It can be seen from the powers of  $T$  in equation 4.3, and in Kramer's law, that the criteria for convection is highly sensitive to the temperature. The effective temperature must be low to achieve surface convection

which is satisfied K dwarfs and M dwarfs. Stars that have a effective temperature less than 7500 K exhibit surface convection, decreasing the effective temperature further causes this envelope to deepen where some stars have convective envelopes which extend all the way to the core (Chaplin (2019)). Generally  $l/m$  is small in low mass stars (Pols (2011)) causing  $\nabla_{\text{rad}}$  to be small in the core (Pols (2011)). Consequently, their cores remain radiative.

The assumption that the stellar material can be modelled as an ideal gas allows for the adiabatic temperature gradient to be found fairly easily.  $\Gamma_1 = c_p/c_v = 5/3 = \Gamma_2$  for a fully ionised gas giving  $\nabla_{\text{ad}} = 0.4$ .  $\nabla_R$  was computed using data accessed online (Christensen-Dalsgaard (1996)) allowing for a model to be plotted for the Sun (see figure 4.1), showing where the convectively unstable region lies.

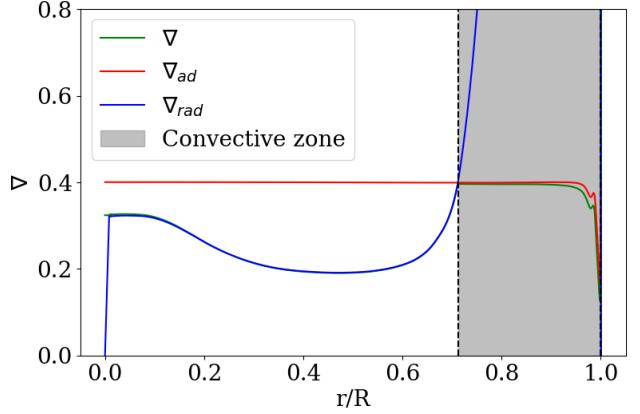


Figure 4.1: Model of the Sun showing three temperature gradients,  $\nabla$ ,  $\nabla_{\text{ad}}$  and  $\nabla_{\text{rad}}$ . It can be seen in the outer  $\simeq 30\%$  of the star satisfies  $\nabla_{\text{rad}} > \nabla_{\text{ad}}$  resulting in the convective envelope.

The stars that are considered here, however, have a lower effective temperature than the Sun. This increases the radiative temperature gradient resulting in the convective envelope extending much further towards the core of the

star as per equation 4.3. The effective temperature is dependent on the mass of the star and stars with masses less than  $0.35M_{\odot}$  are completely convective (Pols (2011)). The depth of the convective envelope as a function of the total mass is represented in figure 4.2.

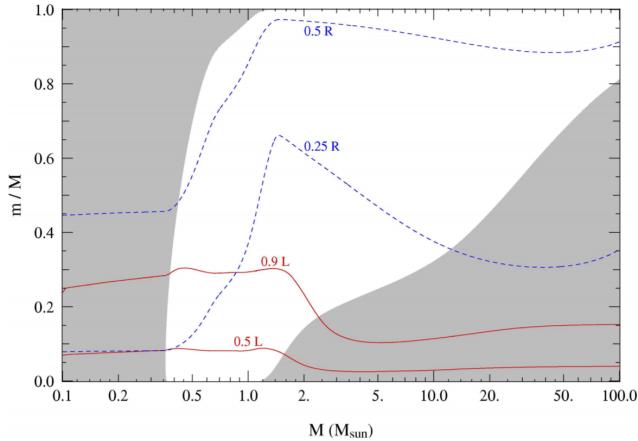


Figure 4.2: Diagram showing the depth of the convective envelope in grey as a function of the star’s total mass. Depth is represented in this plot by the proportion of stellar mass from the surface. The solid red line represents the shells in which 50% and 90% of the total luminosity are produced and the dashed blue lines show where the shells which lie at 25% and 50% of the total radius. (Pols (2011))

The stars involved in this project have masses between  $0.80 M_{\odot}$  and  $0.50 M_{\odot}$  and can, therefore, be expected to have convective zones where the outer  $\simeq 40\%$  of their mass lies.

By scaling  $\kappa$  in accordance with Kramer’s law, models for the stars in TRILEGAL can be found by scaling the model for the Sun.

$$\nabla_{\text{rad}} = \nabla_{\text{rad},\odot} \frac{T_{\text{eff}}}{T_{\text{eff},\odot}}^{-\frac{7}{2}}. \quad (4.4)$$

The model shown in figure 4.3 is obtained for a K dwarf of effective temperature 4682.7 K.

It can be seen that the convective envelope extends much deeper into the star, as expected. This scaling assumed that the only variable which changes between the Sun and the K dwarf is the opacity. More accurate models would use real data for  $l/m$ ,  $\nabla_{\mu}$ , and other quantities that scale with stellar radius, therefore, figure 4.3 should be treated as a rough picture of what is occurring. In reality, data would show that the convective envelope extends deeper than represented in figure 4.3 partially due to  $l/m$  being smaller at lower radii than for the Sun, causing the intersection to happen at smaller values of  $r/R$ .

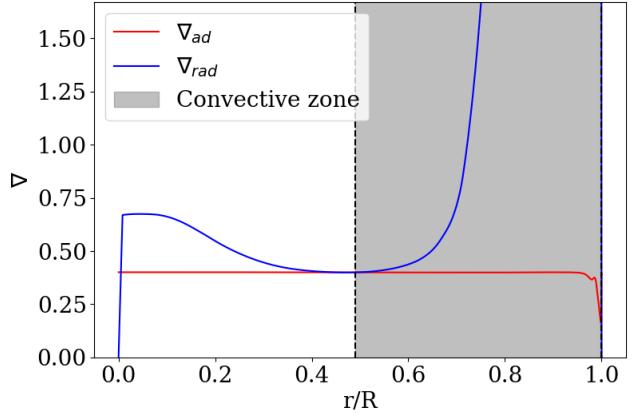


Figure 4.3: Stellar model for a K dwarf with  $T_{\text{eff}} = 4682.7$  K. It can be seen by comparing to figure 4.1 that the depth of the convective zone has increased as a result of the increase in opacity.

## 4.2 Granulation (Sharif)

At the upper boundary of the convective zone,  $\nabla_{\text{rad}} = \nabla_{\text{ad}}$  (Pols (2011)). This results in the buoyant force acting on a bubble vanishing. However, the inertia acquired by the bubble of gas causes it to overshoot the convective boundary where it is strongly braked, especially if it is entering an area of low density such as the photosphere. Here, the temperature is low enough that hydrogen recombines

causing the outer layers to become transparent (Ramírez (2009)). Consequently, a rising bubble experiences a sharp drop in opacity allowing it to radiate away its energy excess causing it to cool enough such that it is now denser than the surrounding material. The buoyant force acting on the bubble changes sign and the bubble falls back into the depths of the star. This occurs across many adjacent convective filaments where plasma is hot and bright as it rises, and cool and dark when it sinks. This process is known as granulation or surface convection. The overshooting and subsequent sinking results in a damped oscillatory motion at the surface of the star causing a granulation pattern (see figure 4.4). This brightness fluctuation results in a continuum background signal in the frequency-power spectrum of the star and due to its damped nature, it is represented as a Lorentzian in the frequency domain. In the time domain, a granule's equation of motion can be found from a stochastically damped oscillator.

$$\frac{d^2}{dt^2}y(t) + 2\eta \frac{d}{dt}y(t) + \nu_0^2 y(t) = f(t) \quad (4.5)$$

where  $y(t)$  is the amplitude of the oscillating granule,  $\eta$  is the linear damping rate,  $\nu_0$  is the frequency of the undamped oscillator and  $f(t)$  is the stochastic driving function. To evaluate this in the frequency domain, the Fourier transforms of  $y(t)$  and  $f(t)$  are introduced as:

$$Y(\nu) = \int y(t)e^{i\nu t} dt \quad (4.6a)$$

$$F(\nu) = \int f(t)e^{i\nu t} dt. \quad (4.6b)$$

The Fourier transform of equation 4.5 would be a Dirac delta function centred about  $\nu_0$ . Since the star will be observed for some finite amount of time,  $T$ , this has the effect of multiplying  $y(t)$  by a step function of width  $T$ . Integrating by parts and substituting in results in equation 4.5

being re-expressed as:

$$-\nu^2 Y(\nu) - 2i\eta\nu Y(\nu) + \nu_0^2 Y(\nu) = F. \quad (4.7)$$

Solving for  $Y(\nu)$  and squaring for the power gives:

$$P(\nu) = |Y(\nu)|^2 = \frac{|F(\nu)|^2}{(\nu_0^2 - \nu^2)^2 + 4\eta^2\nu^2}. \quad (4.8)$$

Assuming that  $\eta \ll \nu \rightarrow |\nu - \nu_0| \ll \nu_0$ , the power can be written as

$$\langle P(\nu) \rangle \simeq \frac{1}{4\nu_0^2} \frac{\langle P_f(\nu) \rangle}{(\nu - \nu_0)^2 + \eta^2} \quad (4.9)$$

where  $\langle P_f(\nu) \rangle$  is the average power of the driving force. This equation describes a Lorentzian profile for damped oscillators and is, therefore, fundamental for modelling the granulation as well as the p-modes described in the next section.

In the case of the granulation, the oscillation of granules are not sinusoidal, as assumed in the previous derivation, rather, that they are exponentially excited and then decay exponentially leading to an altered form of equation 4.9, referred to as a pseudo-Lorentzian.

$$\langle P(\nu) \rangle = \frac{4\sqrt{2}\sigma_c^2}{1 + (2\pi\nu\tau_c)^4} \quad (4.10)$$

where  $\sigma_c$  is the r.m.s intensity of a convective granule and  $\tau_c$  is the average mode lifetime or relaxation time. The granulation is modelled as a zero-centred Lorentzian (Jones (2018)), hence the lack of a  $\nu_0$  term. The choice of the exponent "4" in the denominator of equation 4.10 represents the fact that in the time domain, the granules are exponentially excited before decaying. Some modes, however, have solutions in which this exponent is fixed to "2" representing instantaneous excitation followed by an exponential decay. "4" was chosen, in this case, because it was believed to be more physically realistic, however, it makes little difference to the spectrum at the position of the p-modes.

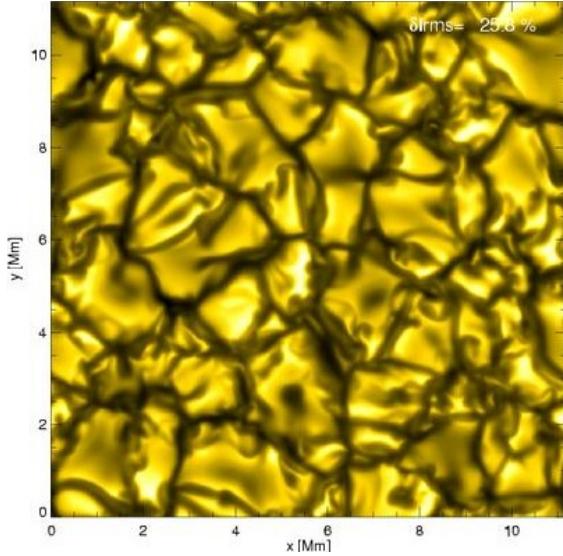


Figure 4.4: Example of granulation on the surface of a star, in this case, the Sun. ([Freytag \(2004\)](#)). The axis scale is Mm,  $10^6$  m

Since granulation is a direct effect of convection, any star which has a convective envelope must have some degree of granulation.

### 4.3 P-Modes (Sharif and Victoire)

**SKB** - The granulation at the surface of the star creates a highly turbulent region which acts as the excitation process for sound waves. A good analogy would be to think about a bell in a sandstorm where some grains of sand excite oscillations, whereas some damp, both in a random manner. This results in resonant frequencies being heard; in a star, there are thousands of resonant frequencies ([Jones \(2018\)](#)). These waves are mediated by pressure forces within the star, hence the name p-mode. The star acts as a spherical cavity for these waves to resonate in. A wave is projected inward from the surface of the star, as it propagates, it enters shells of increasing density causing the sound speed to increase resulting in the wave being refracted back towards the surface until it rebounds back off the inner surface of the star. This process

repeats until the wave traverses the entire star and coalesces with its starting point, forming a standing wave.

The stellar cavity can be approximated as a pipe (when considering radial modes) with one closed end and one end where the "openness" is determined by the acoustic cut off frequency. The acoustic cut-off frequency depends on the sound speed  $c_s$  and the scale height  $H$ :

VA -

$$\omega_{\text{ac}} \equiv 2\pi\nu_{\text{ac}} = \frac{c_s}{2H} \quad (4.11)$$

where the sound speed  $c_s$  in a medium is defined as:

$$c_s = \left( \frac{\gamma \mathfrak{R} T^{\frac{1}{2}}}{\mu} \right) \text{ i.e. } c_s \propto T^{1/2} \quad (4.12)$$

where  $\mathfrak{R}$  the gas constant,  $\gamma$  the adiabatic exponent and  $\mu$  the mean molecular weight. The scale height  $H$  for an isothermal atmosphere as presented in [Lamb \(1932\)](#) is:

$$H = \frac{\mathfrak{R} T_{\text{eff}}}{\mu g} \text{ i.e. } H \propto \frac{T}{g}. \quad (4.13)$$

Consequently,  $\nu_{\text{ac}} \propto g T^{-1/2}$ . Then, since this occurs at the surface, the surface gravity,  $g$ , scales as  $M/R^2$  and  $T_{\text{eff}}$  is the effective temperature, thus:

$$\nu_{\text{ac}} \propto M R^{-2} T_{\text{eff}}^{-1/2}. \quad (4.14)$$

**SKB** - Since the excitation method is stochastic in nature, the waves are projected across a range of angles which penetrate the star to varying depths which affects their visibility as described in later sections. These waves cause the star to oscillate in a way which can be analysed by solving the spherical wave equation. The solution is in the form of a spherical harmonic function which is described by the typical quantised numbers,  $n$ ,  $l$  and  $m$ .  $n$ , the radial order, describes the number of nodes along the radius of the star,  $l$ , the angular degree, describes the number of nodes along the surface of the star (therefore,  $l = 0$  refers

to a radial mode) and  $m$ , the azimuthal order, describes the number of node planes which bisect the equator. Since only non-rotating stars are considered, the  $m$  factor is ignored since these states are degenerate. These fluctuations in intensity were first observed in the Sun, and, therefore, p-modes are often referred to as stellar-like oscillations.

The frequencies of p-modes are found by equation 4.15:

$$\nu_{n,l} = \Delta\nu \left( n + \frac{l}{2} + \epsilon \right) - Dl(l+1) \quad (4.15)$$

where the terms in this equation will be discussed in section 5. This equation results in modes which resemble figure 4.5, which represents real data from previous work by [Chaplin and Miglio \(2013\)](#).

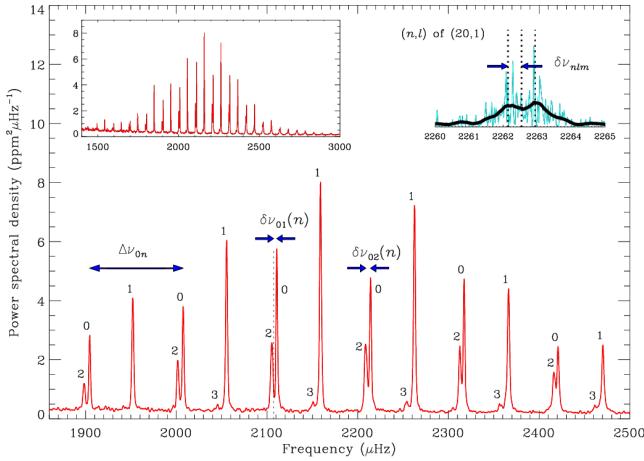


Figure 4.5: Power spectrum for a G-type main-sequence star. It can be seen that modes of equal angular degree (as numbered) are evenly spaced by  $\Delta\nu$  as a result of consecutive radial orders being multiples of the fundamental frequency of the star.

This section demonstrated how p-modes exists in a star ultimately as a result of surface convection occurring at the stellar surface. Section 4.1 quoted that stars of effective temperatures less

than 7500 K exhibit surface convection. Since it is expected that the K and M dwarfs that will be observed by PLATO will have effective temperatures less than this, it can be expected that they will all display granulation, and by extension, p-modes. This should be reflected in the modelled spectra from simulated stars from the TRILEGAL database.

## 5 Components of the spectrum

### 5.1 Introduction to scaling relations (Sharif and Maddie O)

**SKB** - A key method used in producing aspects of the spectrum was the use of scaling relations. The TRILEGAL database provided only a limited number of stellar properties for each star for which plotting the components directly would be impossible. Rather, extensive data from the Sun was utilised for which the models produced here would be based off. The components modelled here are each dependent on a set of stellar parameters. Therefore, dividing the function for the star in question by that of the Sun has the effect of removing any unknown constants in the function and leaving only the stellar parameters, which are often listed in TRILEGAL. This results in a direct scale of the function between the Sun and the star as a product of ratios of the scaling variables. For example, given some arbitrary stellar quantity  $f$ ,

$$f(x, y, z) = Ax \times y^2 \times z^{-\frac{1}{2}} \quad (5.1)$$

where  $A$  is the collection of constants and  $x, y, z$  represent quantities tabulated in TRILEGAL. Its scaling relation would be given by

$$f = f_{\odot} \times \frac{x}{x_{\odot}} \times \frac{y^2}{y_{\odot}} \times \frac{z^{-\frac{1}{2}}}{z_{\odot}}. \quad (5.2)$$

This allowed for complex calculations to be completed without the need to ever find  $A$ .

This approach, however, assumes that the stellar radius can scale all quantities including total mass, effective temperature and luminosity etc. This method completely negates any more complex changes between stars such as the differences in the convective envelopes, degree of degeneracy pressure, changing chemical composition, type of nuclear burning<sup>3</sup> and any other differences between stars which are a result of extreme conditions.

**MO** - The majority of the scaling relations used throughout the production of the oscillation spectra were derived from observations of solar-type stars. This meant it was important to ensure they still held for M and K dwarfs by testing the extremes of the available simulated data and that the spectra being produced still appeared realistic. This involved generating multiple spectra for stars in TRILEGAL and seeing if the plots still appeared to give sensible results and did not change too much from the expected patterns. There are many ways in which the spectra should vary from star to star with some of the key features being the position of  $\nu_{\max}$ , the width of the Gaussian envelope that defines the modelled oscillations and the power of the oscillations.

## 5.2 The Frequency of Maximum Power (Victoire)

A power-frequency spectrum is modulated by a Gaussian-like envelope, the centre of which is defined by the frequency of maximum oscillation power where the observed power is strongest.

This frequency of maximum oscillation power  $\nu_{\max}$  scales with the acoustic cut-off frequency  $\nu_{\text{ac}}$  which depends on the surface conditions (Brown et al. (1991)). The latter defines the

<sup>3</sup>PP-chain or CNO cycles.

frequency above which no waves escape the surface of the star; effectively a boundary for reflection (see equation 4.11). Thus,  $\nu_{\max}$  should also depend on surface conditions.

The trapped modes must have frequencies which are below  $\nu_{\text{ac}}$ . Then, the quantities  $\nu_{\max}$  and  $\nu_{\text{ac}}$  are considered as proportional to each other. Therefore, from equation 4.14:

$$\nu_{\max} \propto \nu_{\text{ac}} \propto MR^{-2}T_{\text{eff}}^{-1/2}. \quad (5.3)$$

Thereby, the frequency of maximum power can be found from the following scaling relation:

$$\frac{\nu_{\max}}{\nu_{\max,\odot}} \simeq \left( \frac{g}{g_{\odot}} \right) \left( \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} \right)^{-1/2}. \quad (5.4)$$

Typical values for the Sun have been imported from the Astropy constants library (Robitaille et al. (2013)):  $R_{\odot} = 695700\text{km}$ ,  $T_{\text{eff},\odot} = 5777\text{ K}$ ,  $g_{\odot} = 274\text{ m.s}^{-2}$ ,  $\nu_{\max,\odot} = 3100\text{ }\mu\text{Hz}$ .

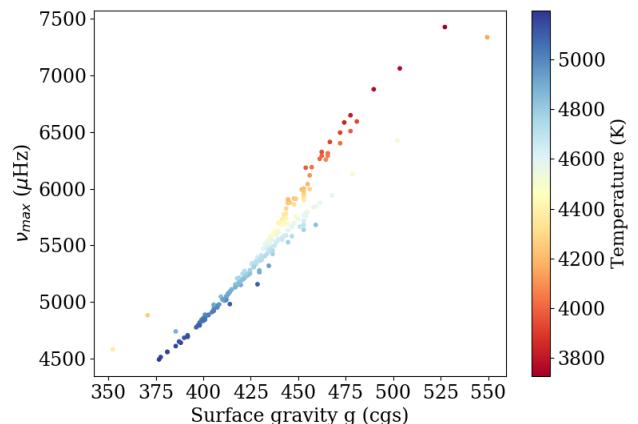


Figure 5.1: The frequency of maximum oscillation power as a function of surface gravity and temperature. This is a simulated sample of K dwarfs from the TRILEGAL synthesis code. This relation applies for a diverse stellar population (see Appendix A3).

Figure 5.1 shows that there is an inverse relationship between surface gravity and tempera-

ture. Additionally, there is a monotonically increasing relationship between the surface gravity and the frequency of the most visible oscillations, following equation 5.4. The relation seems to hold for K dwarfs since figure 5.1 is based on a K dwarfs population. Nevertheless, observations for the hottest F dwarfs have shown flatter maxima than yielded by this scaling relation (Chaplin and Miglio (2013)). Therefore, this relation could be imperfect for red dwarfs; this can only be confirmed once observations of the stellar oscillations for these spectral types are conducted.

### 5.3 The Large Frequency Separation (Victoire)

The oscillation spectrum presents a near-regular frequency structure, reflecting a similar-sized cavity for the trapped oscillations. This is defined by the large frequency separation  $\Delta\nu_{\text{nl}}$ , the average spacing between successive modes of similar spherical angular degrees  $l$  and consecutive radial orders  $n$ . Thus, it follows:

$$\Delta\nu_{\text{nl}} = \nu_{n+1,l} - \nu_{nl}. \quad (5.5)$$

The large frequency separation also relates to the sound travel-time across the star  $c(R)$ :

$$\Delta\nu = \left( 2 \int_0^R \frac{dr}{c(r)} \right)^{-1}. \quad (5.6)$$

The observed frequency spacing scales as the square root of the mean density of the star (Christensen-Dalsgaard et al. (1995)):

$$\Delta\nu_{\text{nl}} \propto <\rho>^{1/2}. \quad (5.7)$$

Consequently, the large frequency separation also provides a measurement of the acoustic radius of a star. Then, the large frequency separation can be found from the following scaling relation:

$$\frac{\Delta\nu}{\Delta\nu_{\odot}} \simeq \left( \frac{M}{M_{\odot}} \right)^{1/2} \left( \frac{R}{R_{\odot}} \right)^{-3/2} \quad (5.8)$$

with  $\Delta\nu_{\odot} = 134.9 \pm 0.02 \mu\text{Hz}$  (Toutain and Fröhlich (1992)).

As the star ages, its radius increases and one can observe an increasing acoustic travel-time, a decreasing stellar density and consequently, a decreasing large frequency separation (see figure 5.2).

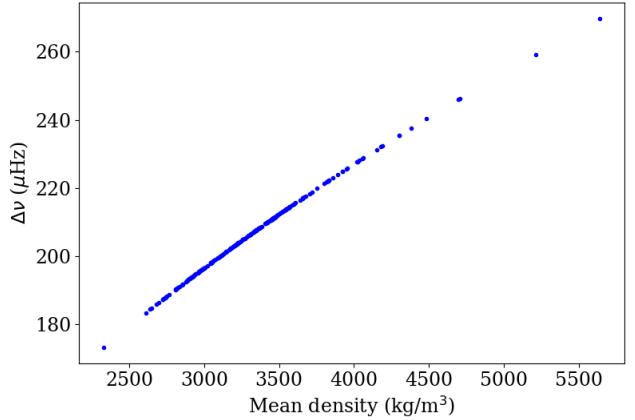


Figure 5.2: The large frequency separation as a function of mean density. This is a simulated sample of K dwarfs from the TRILEGAL synthesis code

### 5.4 The Small Frequency Separation, D (Jeff and Sam)

The small frequency separation is the difference in the observed frequencies of two distinct angular degree modes within each overtone  $n$  on a star's frequency-power spectrum. It is labelled as:

$$\delta\nu_{\text{nl}} = \nu_{nl} - \nu_{n-1,l+2}. \quad (5.9)$$

Irrespective of the particular angular degree modes in question, the small separation is always a multiple of the parameter  $D$ . This is encoded within the  $Dl(l+1)$  term in equation 4.15 which, for example, implies  $\delta\nu_{02} = 6D$  and  $\delta\nu_{13} = 10D$ .

For stars of a given stellar mass, this parameter  $D$  (and therefore,  $\delta\nu_{\text{nl}}$ ) is sensitive pre-

dominantly to the sound-speed gradient in the stellar core. This is formalised in [Christensen-Dalsgaard \(1997\)](#) as:

$$\delta\nu_{\text{nl}} \simeq -(4l + 6) \frac{\Delta\nu}{4\pi^2\nu_{\text{nl}}} \int_0^R \frac{dc_s}{dr} \frac{dr}{r}. \quad (5.10)$$

[Equation 4.12](#) demonstrates that the small separation for stars of a given stellar mass is largely dependent upon the composition profile of their core.

Taking Kepler main-sequence star data from [Aguirre et al. \(2017\)](#) and plotting  $\delta\nu_{02}$  against stellar mass and age visualises this dependency (figure 5.3).

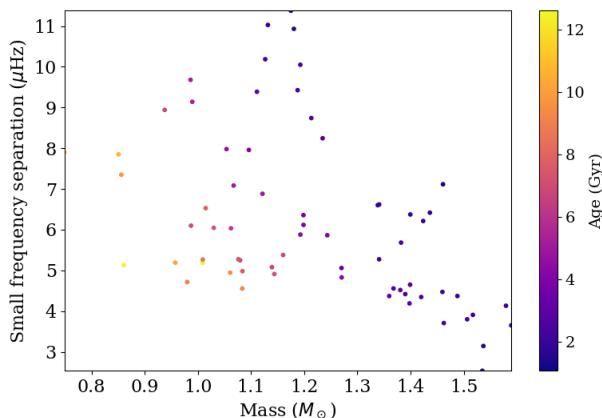


Figure 5.3: The small frequency separation of main-sequence stars plotted against their mass and age. A clear variation in  $\delta\nu_{02}$  with stellar age is exhibited.

As demonstrated in figure 5.3, the composition profile of the stellar core is related to its age. This is because the mean molecular weight,  $\mu$ , of the stellar core effectively doubles over a star's main-sequence lifetime as the initial pure hydrogen core is progressively fused into helium. Longer lived stars of a given mass have higher concentrations of helium in their cores, and thus a larger sound-speed gradient. Figure 5.3 demonstrates that this holds proportionally for all stars in the sample, irrespective of mass.

Figure 5.3 reflects the negative correlation predicted by equation 5.10: longer lived stars of a given stellar mass exhibit lower  $\delta\nu_{02}$  values.

## 5.5 The Gaussian Envelope (Maddie O and Jeff)

**MO** - The spectrum at the location of the p-modes can usually be modelled using a Gaussian-shaped envelope ([Chaplin and Miglio \(2013\)](#)). The Gaussian envelope was defined as

$$H_{\text{gau}}(\nu) = H_{\text{env}} \exp\left(-\frac{(\nu - \nu_{\max})^2}{2\sigma^2}\right), \quad (5.11)$$

where  $H_{\text{env}}$  is the maximum height of the envelope at  $\nu_{\max}$  and  $\sigma$  is the width of the envelope.

### 5.5.1 Amplitude (Maddie O)

The amplitudes of the Lorentzian peaks that represent the p-mode oscillations are determined by the balance between damping and excitation of the modes. As described by [Chaplin and Miglio \(2013\)](#) the peaks should not be perfectly symmetrical due to the "very localised excitation source" causing small changes between modes that can be detected. This asymmetry has been observed in solar p-mode oscillations and is predicted to be similar in other stars. Therefore, the amplitude of the Gaussian curve used to model the oscillations of the p-modes could not be that of the standard Gaussian distribution form. Instead, the maximum root mean square amplitude, or the amplitude at  $\nu_{\max}$ , was determined for the radial modes using scaling relations based from solar values, as outlined by [Ball et al. \(2018\)](#) and this value was used to scale the Gaussian. The amplitudes of the radial mode oscillations were given as:

$$A_{\max} = A_{\max,\odot} \beta \left( \frac{L}{L_{\odot}} \right) \left( \frac{M}{M_{\odot}} \right)^{-1} \left( \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} \right)^{-2} \quad (5.12)$$

where  $A_{\max,\odot} = 2.1$  ppm, as determined in the TESS bandpass ([Ball et al. \(2018\)](#)). The beta

factor in equation 5.12 takes into account that the amplitudes of the hottest dwarf stars appear to decrease. This value can be calculated using:

$$\beta = 1 - \exp\left(\frac{T_{\text{eff}} - T_{\text{red}}}{\Delta T}\right), \quad (5.13)$$

where  $\Delta T$  is taken to be 1250 K (Ball et al. (2018)) and  $T_{\text{red}}$  was calculated with

$$T_{\text{red}} = 8907 \text{K} \left(\frac{L}{L_{\odot}}\right)^{-0.093}. \quad (5.14)$$

This  $T_{\text{red}}$  value is the “temperature of the red edge of the  $\delta$ -Scuti instability strip at the star’s luminosity” (Ball et al. (2018)). The instability strip is a region found on the Hertzsprung-Russell diagram containing multiple different types of variable stars as demonstrated in figure 5.4.

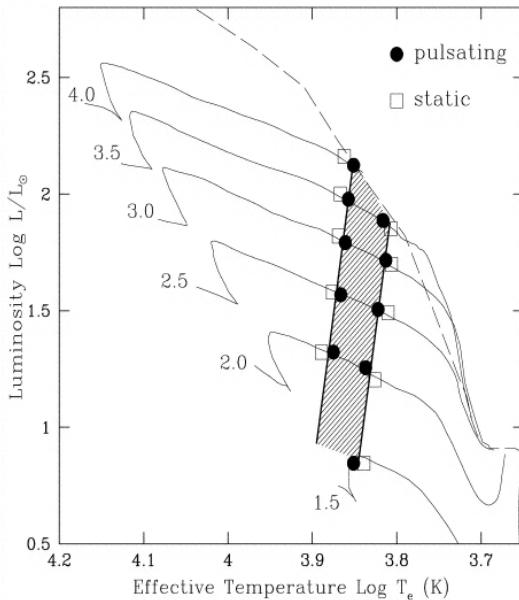


Figure 5.4: The instability strip extends from the main-sequence to the red giant phase of a star’s evolution, with multiple types of variable stars found within this strip, (Marconi and Palla (1998)).

Stars that are more massive than the Sun become variable at least once in their evolution.

Once they enter into the instability strip, instabilities cause these stars to pulsate in size and therefore, they have variable luminosity.  $\delta$ -Scuti stars are a type of small amplitude variable star that are found in the instability strip, either on the main-sequence or evolving from the main-sequence to the Giant phase of their life. The value of  $A_{\text{max}}$  was then used to calculate the prefactor  $H_{\text{gau}}$  that determines the maximum power of the Gaussian amplitude, using

$$H_{\text{env}} = \frac{A_{\text{max}}^2 \zeta}{2\Delta\nu}, \quad (5.15)$$

where  $\zeta$  is the sum of the relative visibilities (defined in section 6.2.3) squared,

$$\zeta = \sum_l \left(\frac{V_l}{V_0}\right)^2. \quad (5.16)$$

It is necessary to sum over the mode visibilities because every  $\Delta\nu$  segment has a power that depends on a mode for each degree and it is sufficient to sum up to  $l = 3$  (spe). The relations used to determine the height of the radial modes were based from observations of solar-like stars but there were a couple of simple checks that could be done to ensure the calculated amplitudes were scaling as they should be, even in dwarf stars. The radial mode amplitude should be inversely proportional to  $\log(g)$  (Chaplin and Miglio (2013)) and proportional to  $T_{\text{eff}}$ . The amplitudes for the radial mode were plotted against these properties of the stars, for 100 stars from the K dwarf catalogue and M dwarf catalogue. The relations for K dwarfs are shown in figures 5.5 and 5.6 and the correct trends can be seen.

The same correlations between the maximum radial mode amplitude and  $T_{\text{eff}}$  and  $\log(g)$  can still be seen in the plots produced for the M dwarf stars. However, to be sure that these relations are accurate for usage, the best value to look at is the residual.

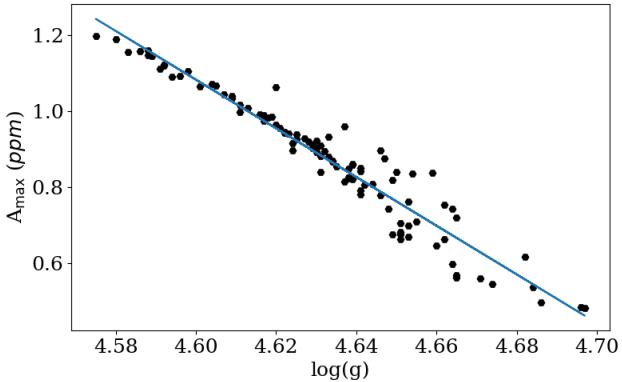


Figure 5.5: The amplitude of the radial mode oscillations scales inversely with  $\log(g)$  in K dwarf stars, in the same way it has been shown to scale in larger main-sequence stars. The residual for the best fit in this figure is 0.2056.

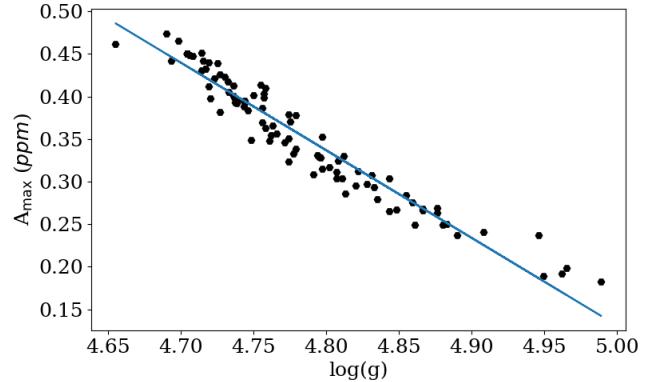


Figure 5.7: The amplitude of the radial mode oscillations scales inversely with  $\log(g)$  in M dwarf stars, as anticipated. The residual for the best fit in this figure is 0.0346.

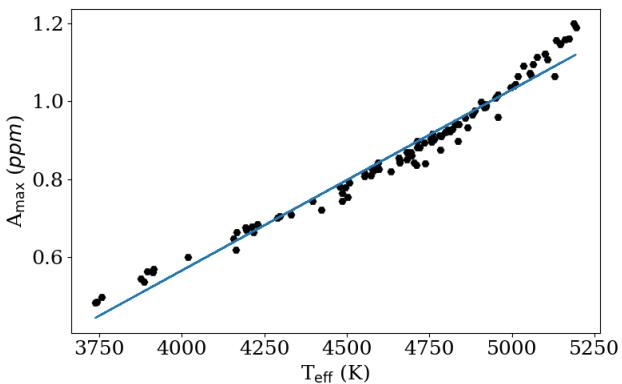


Figure 5.6: The amplitude of the radial mode oscillations scales proportionally to  $T_{\text{eff}}$  in K dwarf stars, as expected. The residual for the best fit in this figure is 0.0927.

The residual from the best fit plot of figure 5.6 is 0.0927 and that of figure 5.5 is 0.2056. In comparison to the M dwarf stars where the residual of figure 5.8 is 0.0952 and that of figure 5.7, 0.0346. All of these residual values demonstrate that the relations used produce a good representation for the amplitudes of the radial modes in relation to the intrinsic properties of the stars.

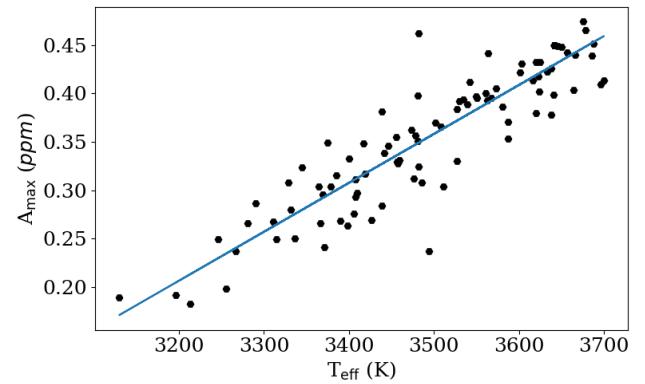


Figure 5.8: The amplitude of the radial mode oscillations scales proportionally to  $T_{\text{eff}}$  in M dwarf stars. The residual for the best fit in this figure is 0.0952.

### 5.5.2 Width (Jeff)

In equation 5.11,  $\sigma$  is defined as the width of the Gaussian distribution. In statistics, this corresponds to the fluctuation of the mean, here, this defines where the peak of the curve lies. In probability theory, the width of the Gaussian distribution is the standard deviation. It is used to measure the amount of variation in a set of measurements, to see how much the data fluctuates from the mean.

The standard form of a Gaussian is as follows ([Weisstein](#)):

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}. \quad (5.17)$$

In this project,  $x$  was associated to  $\nu$ , the frequency and  $\mu$  to  $\nu_{\max}$ , the frequency of maximum power, as seen in equation [5.11](#). The FWHM of a Gaussian can be found by finding the half-maximum points ( $x_0$ ). This yields  $x_0 \pm \sigma\sqrt{2\ln 2} + \mu$  with the consideration that  $f(x_{\max})$  is obtained at  $\mu = x_{\max}$ . Using this, the FWHM is thus given by:

$$\text{FWHM} = 2\sqrt{2\ln 2}\sigma \approx 2.355\sigma. \quad (5.18)$$

Then, [Christensen-Dalsgaard et al. \(2010\)](#) analysed smoothed power spectra by confining each spectrum to a Gaussian-shaped band with a Full Width Half Maximum (FWHM) as defined in equation [5.19](#). Since this worked throughout his investigation on Kepler data, it was deemed reasonable to use this condition for this project too.

$$\text{FWHM} = 5\Delta\nu. \quad (5.19)$$

Hence, combining equations [5.18](#) and [5.19](#), it is found that:

$$5\Delta\nu = 2.355\sigma. \quad (5.20)$$

This leads to the width of the Gaussian distribution being defined as:

$$\sigma = 2.123\Delta\nu. \quad (5.21)$$

From equation [5.21](#), it seems that the only parameter on which the width of the Gaussian envelope depends is the large frequency separation. Furthermore, the standard deviation  $\sigma$  shows how the amplitude of  $\Delta\nu$  is distributed. If it is large, then the values are more distributed, if it is low, then the amplitude will be more central.

## 5.6 The Lorentzian peaks

### 5.6.1 Definition (Maddie T and Victoire)

**VA** - Similar to granulation, the p-modes which are stochastically-excited by turbulent convection and intrinsically-damped are analogous to a damped harmonic oscillator. Thus they are well described by a Lorentzian function ([Batchelor \(1953\)](#), [Kumar et al. \(1988\)](#)),

$$L_{nlm}(\nu) = \frac{\varepsilon_{lm}(i_*)\tilde{v}_l^2 S_{n0}}{1 + \frac{4}{\Gamma_{nl}^2}(\nu - \nu_{nl} + m\nu_s)^2}. \quad (5.22)$$

The numerator corresponds to a height, dependent on the stellar inclination  $i_*$ ,  $S_{n0}$  is the height of the radial mode of order  $n$  and the squared visibilities,  $V_{nl}$ .  $\Gamma_{nl}$  is the FWHM mode width and  $\nu_{nl}$  is the frequency of the  $m = 0$  modes. However, here a non-rotating case is considered, thus rotational splitting is not included ( $m\nu_s$ ) and  $i_*$  is set to 0. Then, the equation reduces to:

$$L_{nlm}(\nu) = \frac{\tilde{v}_l^2 S_{n0}}{1 + \frac{4}{\Gamma_{nl}^2}(\nu - \nu_{nl})^2}. \quad (5.23)$$

This is the form that was considered to model the Lorentzian peaks.

### Inclination

**MT** - The data provided for the simulated stellar population included no indication of the orientation of the stars and their rotation axis relative to the observer. It was consequently assumed that the simulated stars had an inclination angle of the stellar rotation axis,  $i_* = 0$  relative to the observer's line of sight. At an inclination of zero, the visibility of all modes with  $|m| > 0$  reduces to zero.

Using this simplified model, the  $\varepsilon_{lm}(i_*)$  term in equation [5.22](#), and defined in equation [5.24](#) was therefore set to 1. In doing this, the model

disregards any effects on the visibility of oscillation modes of different  $m$  due to the inclination of each star's rotation axis.

Nevertheless, including an estimation of this term would have been an improvement to the stellar oscillation models developed here and altered the appearance of the oscillation modes. The visibilities of modes depending on  $l$  and  $m$  according to [Ball et al. \(2018\)](#) in power are,

$$\varepsilon_{lm}(i_s) = \frac{(l - |m|)!}{(l + |m|)!} [P_l^{|m|}(\cos i_s)]^2 \quad (5.24)$$

which is a function of inclination angle  $i_s$  where  $P_l^{|m|}$  are the Legendre Polynomials. In the true selection of stars observed by PLATO, there will likely be distribution of inclination angles from 0-90°. This will have an impact on the visibilities of modes due to equation 5.24 and hence the height of the Lorentzian peaks in the modelled spectra ([Nielsen et al. \(2014\)](#)).

## Rotational Splitting

**MT** - The  $m\nu_s$  term in equation 5.22 accounts for the effects of rotational splitting of oscillation modes. This term was not considered when modelling the spectra of the simulated stars however, would improve the accuracy of the appearance of the modelled spectra; aiming to mimic those that will be seen by PLATO.

For a non-rotating star, each oscillation mode with an azimuthal order  $|m| > 0$  is degenerate with the modes of order  $m = 0$ . However, at small rotation velocities, the intensity of each p-mode is Doppler shifted in frequency by an amount linearly proportional to the rotation rate ([Nielsen et al. \(2014\)](#)). This gives rise to the frequency of each mode  $l$  splitting uniformly into  $2l + 1$  azimuthal modes as  $m$  can take values of  $-l < m < l$ .

$$\nu_{nlm} = \nu_{nl} + m\nu_s \quad (5.25)$$

where  $\nu_{nlm}$  are the frequencies of modes for different  $m$ ,  $\nu_{nl}$  is the frequency of the  $m = 0$

mode and  $\nu_s$  is the rotating frequency of the star. An observer in the non-rotating frame, therefore, sees the frequencies  $\nu_{nlm}$  of modes as split depending on their value of  $m$  ([Aerts et al. \(2010\)](#)).

This is a basic model assuming constant angular velocity which is inevitably not always the case. Nevertheless, the split modes would have significant implications for the power of the oscillation spectra. If the model were to consider rotationally split multiplets, the data analysis aiming to quantify the power above and below the smooth line of the spectrum would be somewhat altered; perhaps making detection of p-modes easier due to the increased power encompassed by the doubled modes.

The amplitudes of the oscillation modes would likely not change from the inclusion of rotational splitting however, so the basic, non-rotating model was deemed sufficient to test PLATO's detection limits.

### 5.6.2 The width of the peaks (Victoire)

To evaluate the width of the peaks, three approaches were considered.

#### The linewidth from [Lund et al. \(2017\)](#):

The relation defining the width of the Lorentzian peaks relied on the analysis provided by [Lund et al. \(2017\)](#). Accessing the data available from that paper allowed the linewidth of the p-modes to be plotted as a function of frequency for the radial mode as well as effective temperature. It was seen that for each star, as frequency increases, so do the linewidth, with an apparent dip centred on the  $\nu_{\max}$ . The relation between the two appears more linear following the dip. Similarly, as the temperature increases, so does the linewidth, as demonstrated in figure 5.9.

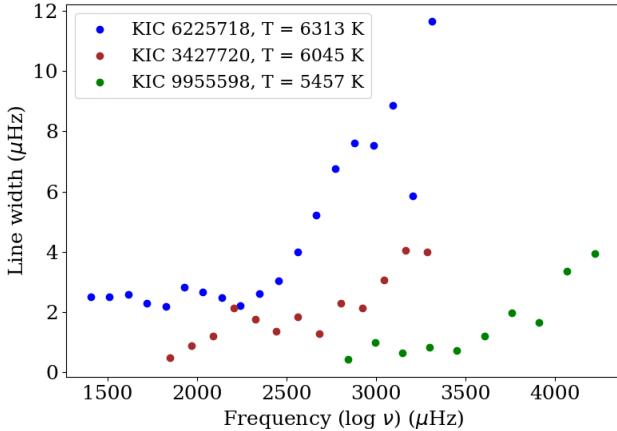


Figure 5.9: The radial linewidth with respect to frequency at different temperatures, based on the data from Lund et al. (2017).

Then, following Appourchaux et al. (2014), the linewidth may be characterised by:

$$\ln \Gamma = \alpha \ln \left( \frac{\nu}{\nu_{\max}} \right) + \ln \Gamma_\alpha + \frac{\ln \Delta \Gamma_{\text{dip}}}{1 + \left( \frac{2 \ln(\nu/\nu_{\text{dip}})}{\ln(W_{\text{dip}}/\nu_{\max})} \right)^2}$$

where  $\Gamma_\alpha$  is the linewidth value at  $\nu_{\max}$ . The first two terms correspond to a power law of exponent  $\alpha$ . Then, the final term accounts for a Lorentzian dip (as seen in figure 5.9) centred on  $\nu_{\text{dip}}$ . However, when extrapolating this relation for M and K dwarfs, it was decided to disregard the dip. Lower temperature values (than solar-like effective temperatures) yield a negative amplitude of  $\Delta \Gamma_{\text{dip}}$  which prevents the log from being evaluated. Therefore, the linewidth was estimated using:

$$\ln \Gamma = \alpha \ln(\nu/\nu_{\max}) + \ln \Gamma_\alpha. \quad (5.26)$$

Fortunately, Lund et al. (2017) fitted the variables ( $\Gamma_\alpha$  and  $\alpha$ ) against the effective temperature  $T_{\text{eff}}$ . The following functions were obtained:

$$P_\alpha = -25.5 T_{\text{eff}} + 29.1 \quad (5.27)$$

$$P_{\Gamma_\alpha} = 6.3 T_{\text{eff}} - 1.8. \quad (5.28)$$

These are used to define the linewidth  $\Gamma$  along with the scaling relation to obtain  $\nu_{\max}$ , the frequency of maximum oscillation power, following equation 5.26. As the stars get smaller and cooler, the linewidth should decrease. Therefore, the values for K and M dwarfs are expected to be smaller than those for solar-like stars. This was plotted in figure 5.10.

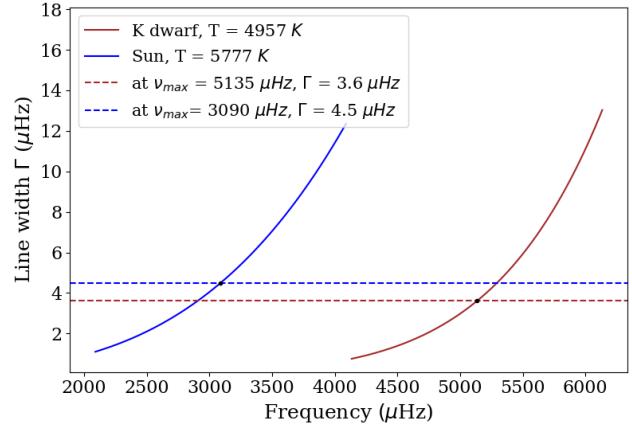


Figure 5.10: The radial linewidth as a function of frequency. This is using the relation as defined in equation 5.26. Therefore, it does not account for the Lorentzian dip that should appear in the fit for the Sun.

The Lorentzian dip can be taken into account for the Sun when plotting the radial linewidth (see figure 5.11). As can be seen in figure 5.11, the linewidth for the Sun, evaluated using the full equation, yields a smaller value than that of the K dwarf, evaluated using equation 5.26. Thus, there is a caveat here: disregarding the Lorentzian dip seems to lead to an overestimate of the width.

Table 1 in Chaplin et al. (1997) gives an estimate of the solar linewidths. It can be seen that at 3040.1  $\mu$ Hz the width is measured to be  $1.089 \pm 0.151 \mu$ Hz. This agrees with the fit in figure 5.11 which includes the Lorentzian dip in the analysis; at a frequency of 3040  $\mu$ Hz, it yields a value of  $\approx 0.89 \mu$ Hz. A common value of the linewidth for the Sun is  $\approx 1 \mu$ Hz. The fit

in figure 5.10 which disregarded the dip, provided a width of  $\approx 4.3 \mu\text{Hz}$  at  $\nu_{\max}$ . Therefore, there is an overestimate by a factor of 4.3, using this method.

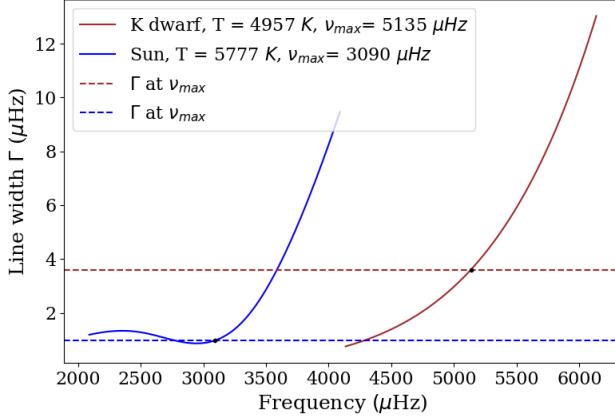


Figure 5.11: The radial linewidth as a function of frequency. This is using the relation as defined by Appourchaux et al. (2014). Thus it accounts for the Lorentzian dip for the Sun.

Consequently, this would yield a value of the amplitude that would be smaller than expected. The height is inversely proportional to the width, as explained later (equation 5.33). Due to the overestimate of the widths there would be an impact on detectability.

Other methods were then considered to estimate widths deemed more reasonable for red dwarfs.

### The damping rate:

The Lorentzian has an FWHM of the damping rate  $\eta$ . Observations show that this increases with increasing frequency (Christensen-Dalsgaard (1997)). The mode linewidth is then related to the mode damping rate through:

$$\Gamma_{\text{nl}} = \frac{\eta_{\text{nl}}}{\pi}. \quad (5.29)$$

From Chaplin et al. (2009), the FWHM linewidth,  $\Delta$  of a mode is defined as:

$$\Delta = \frac{1}{\tau\pi} = \frac{\Gamma_{\text{nl}}}{2\pi} \quad (5.30)$$

with  $\tau$  the e-folding lifetime. They additionally proposed a theoretical scaling relation for main-sequence and giant stars:  $\langle \tau \rangle \propto \langle T_{\text{eff}} \rangle^{-4}$ . However, the CoRoT and Kepler observations disagreed with this which lead Belkacem et al. (2012) to refine the relation, considering the observations, and deriving a stronger effective temperature dependence for the damping rate:

$$\eta \propto T_{\text{eff}}^{10.8} g^{-0.3}. \quad (5.31)$$

To scale this for other main-sequence stars, the damping rate can be estimated from equation 5.29 and the standard value of  $\approx 1.09 \mu\text{Hz}$  at  $\nu_{\max}$ .

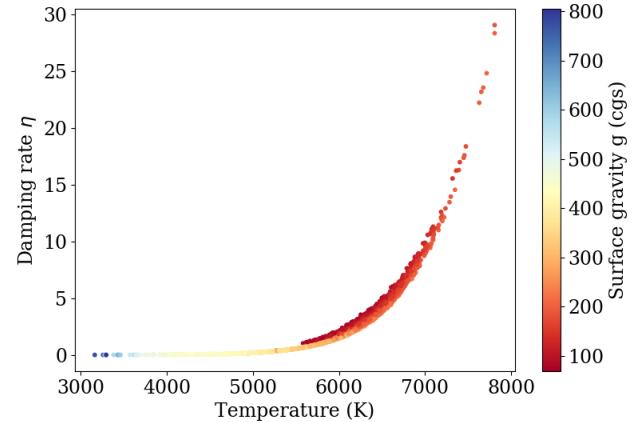


Figure 5.12: The damping rate as a function of temperature and surface gravity, following equation 5.31 (TRILEGAL data).

Subsequently, using equation 5.29 with equation 5.31, the value of the mean width can be estimated. This can be plotted for K dwarfs as in figure 5.13.

This produces a constant value of the linewidth instead of being mode/frequency dependent. The results are of the order of the magnitude that is expected for K dwarfs, ie. smaller than for G-type stars like the Sun.

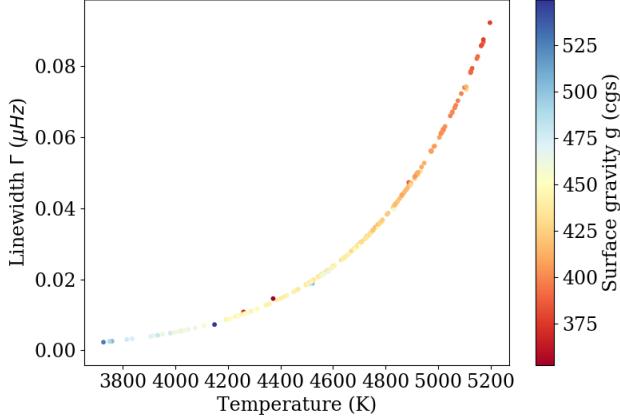


Figure 5.13: The linewidth as a function of temperature and surface gravity, following equation 5.29 for K dwarfs (TRILEGAL data). See Appendix A4 for the plot considering the entire stellar population.

#### The linewidth at $\nu_{\max}$ fit:

[Appourchaux et al. \(2012\)](#) defined a power law relating the linewidth at  $\nu_{\max}$  to the effective temperature:

$$\Gamma = \Gamma_0 + \alpha \left( \frac{T_{\text{eff}}}{5777} \right)^{\beta}. \quad (5.32)$$

Then, [Lund et al. \(2017\)](#) found the fitting parameters using Monte Carlo runs:  $\Gamma \approx 0.07 \mu\text{Hz}$ ,  $\alpha \approx 0.91 \mu\text{Hz}$  and  $\beta \approx 15.3 \mu\text{Hz}$ . Equation 5.32 can be used to calculate the linewidth for K dwarfs.

This yields slightly smaller values than what was obtained from equation 5.31 (comparing figure 5.13 to figure 5.14). Nevertheless, the values are generally similar between those two methods. The values of the parameters  $\Gamma, \alpha, \beta$  were found from fitting red giants and solar-like stars. Therefore, there is an approximation that is being made here when extrapolating this relation for K dwarfs.

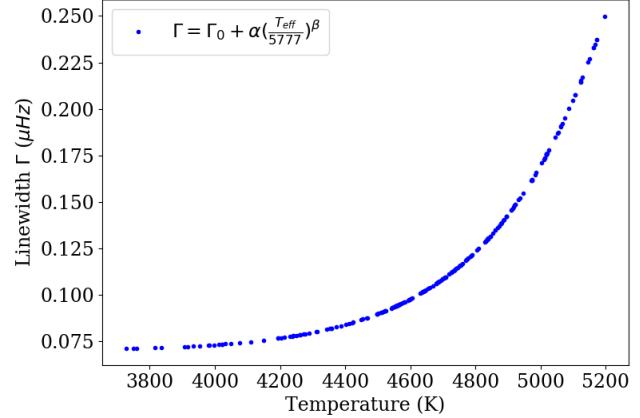


Figure 5.14: The linewidth as a function of temperature following equation 5.32 for K dwarfs (TRILEGAL data). See Appendix A5 for the plot considering the entire stellar population.

These last two methods produced values which are narrower than what is expected for K dwarfs. However, it is a better representation than estimates found using [Appourchaux et al. \(2014\)](#) and [Lund et al. \(2017\)](#) without the Lorentzian dip. Thus, this is the method that was implemented to model the width of the Lorentzian peaks.

#### 5.6.3 The height of the peaks (Maddie T)

To define the height of the Lorentzian peaks, the spectra had to be modulated in power according to the Gaussian envelope. The total power found when integrating the Gaussian envelope must be the same as the power found when integrating the Lorentzian peaks ([Chaplin and Miglio \(2013\)](#)). Equation 5.33 demonstrates how the height of the radial modes  $S_{n0}$  can be determined from the height of the Gaussian envelope,  $H_{\text{env}}$ ,

$$H_{\text{env}}(\nu_{n0}) = \left( \frac{\zeta \pi \Gamma}{2 \Delta \nu} \right) S_{n0}. \quad (5.33)$$

To calculate the heights,  $S_{n0}$ , equation 5.33 can be rearranged and the height of the envelope,

as calculated in equation 5.15, is substituted in. The modulated heights are then used in equation 5.23 to plot the radial modes as Lorentzian peaks. Figure 5.15 shows radial modes plotted as Lorentzian peaks, modulated in power by the Gaussian envelope. The width here was calculated using the previously defined  $T_{\text{eff},\odot}=5777$  K,  $\nu_{\text{max},\odot} = 3100 \mu\text{Hz}$ ,  $\Delta\nu_{\odot} \approx 135 \mu\text{Hz}$  and  $A_{\text{max},\odot} = 2.1$  ppm.

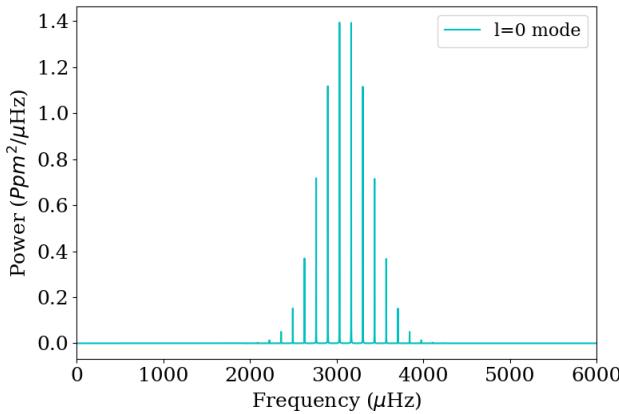


Figure 5.15: Lorentzian peaks for solar  $l=0$  mode with heights modulated by the Gaussian envelope.

The amplitude and hence the power of the radial modes for a particular star are scaled from the solar values using equation 5.11. Nevertheless, the modes of higher angular degree ( $l = 1, 2, 3, \dots$ ) have different amplitudes depending on geometric factors. These come in two forms; cancellation and the inclination of the star's rotation axis.

The geometric behaviour of stellar oscillations is dictated by spherical harmonics  $Y_l^m$  which are functions of the angular degree,  $l$  and azimuthal order,  $m$ . As the star oscillates, matter moving towards the observer is brighter and blue-shifted and matter moving away is dimmer and red-shifted. For increasing angular degree  $l$ , the number of brighter and dimmer sections on the stellar surface increases. This means that, when averaging over the stellar

disc, these sections effectively cancel and the changes in intensity are harder to detect. Consequently, only modes  $l = 0$  up to  $l = 3$  tend to be visible for asteroseismic observations (Ball et al. (2018)).

The visibility,  $V_l$ , quantifies these cancellation effects for each mode of angular degree  $l$ .

$$V_l = 2\pi\sqrt{2l+1} \times \int_0^{\frac{\pi}{2}} P_l(\cos\theta)W(\cos\theta) \cos\theta \sin\theta d\theta \quad (5.34)$$

where  $P_l$  are the Legendre polynomials and  $W$  is a weighting factor which depends on the angular distance from the rotation axis of the star. In general, calculations will use the visibilities of each mode relative to the radial mode ie. the 'relative visibilities'. These ratios can be calculated as in equation 5.35.

$$\tilde{V}_l^2 = \left( \frac{V_l}{V_0} \right)^2 \quad (5.35)$$

The ratio is squared as the visibilities are usually incorporated into frequency-power spectra as opposed to amplitude-frequency.  $\tilde{V}_l^2$  can then be used as in Equation 5.23 to scale the heights of the Lorentzian peaks for modes of different angular degree (Lund et al. (2017)).

The effect on the relative amplitudes of the modes can be seen in figure 5.16. Since the relative visibilities depend on the geometry of a sphere only, typical values can be used as a good model for any star. Here, they are  $(V_1/V_0)^2 \approx 1.5$ ,  $(V_2/V_0)^2 \approx 0.5$ , and  $(V_3/V_0)^2 \approx 0.04$  (Chaplin et al. (2013)). For the final modelled spectra, the Lorentzians for each mode are summed into one continuous function.

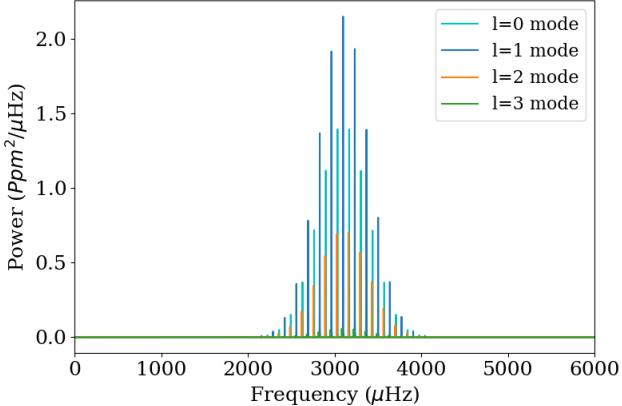


Figure 5.16: Lorentzian peaks for solar  $l = 0$  to  $l = 3$  modes applying relative mode visibilities to the amplitudes.

## 5.7 D (Maddie O and Sam)

**SBW** - To accurately model frequency-power spectra for K and M dwarfs, the small frequency separation between angular degree modes is required. As previously explained, the small separation is scaled by factors of  $D$  which necessitates the creation of an accurate scaling relation for  $D$ .

Based upon data from 100 solar-like main-sequence stars, [Bellinger \(2019\)](#) posits a scaling relation for stellar age, which can be rearranged for  $D$ ,

$$\frac{\text{Age}}{\text{Age}_\odot} \propto \left( \frac{\delta\nu}{\delta\nu_\odot} \right)^{-1.292} \equiv \left( \frac{D}{D_\odot} \right)^{-1.292}. \quad (5.36)$$

Inputting parameters of K dwarfs from the synthetic catalogue suggested that this relation does not hold for these low-mass main-sequence stars; output values of  $D$ , in this case, implied a small frequency separation  $\delta\nu_{02}$  in excess of  $120\mu\text{Hz}$ , much larger than that exhibited in stars of any age, mass or evolutionary stage.

The reason for the breakdown of the [Bellinger \(2019\)](#) equation at low masses is related to the sample upon which it is based. The precision of the exponents of each term suggests they have

been chosen to fit the particular range of masses of the sample, and thus it is unsurprising that K dwarfs, which fall outside this mass range, do not exhibit the same agreement. Physically, this disagreement can likely be attributed to the large age range of red dwarfs, compared with more massive main-sequence stars; the longest-lived, smallest such stars in the synthetic catalogue are in excess of 12 billion years old. As seen in figure 5.3, the variance in  $\delta\nu_{02}$  with age, for stars of a particular stellar mass, is much more pronounced for red dwarf stars than those of greater mass.

**MO** - Following the realisation that the original scaling relation for  $D$  could not be used, a new equation had to be determined. A scaling relationship for  $D$  and the average  $\delta\nu_{02}$  value, is given by

$$D = \frac{1}{6}\delta\nu_{02} \quad (5.37)$$

([Kjeldsen et al. \(2008b\)](#)). An approximate value for  $\delta\nu_{02}$  was determined using an modelled échelle diagram of a K5 dwarf star, giving an average  $\delta\nu_{02}$  of  $15\mu\text{Hz}$ . Using equation 5.37, the new relation for  $D$  would therefore need to produce values of around  $2.5\mu\text{Hz}$  for a K dwarf star. It is also known that  $D$  scales as the dynamic frequency, a measure of the stability of a star ([Chaplin and Miglio \(2013\)](#)). This is due to the dynamical timescale being of similar length to the time for a sound wave to cross a star. The dynamic frequency, in turn scales as  $\sqrt{\rho}$ . As  $\Delta\nu$  also scales with  $\sqrt{\rho}$ , the new scaling relation used for  $D$  is

$$D = D_\odot \left( \frac{\Delta\nu}{\Delta\nu_\odot} \right), \quad (5.38)$$

where  $D_\odot$  is taken as 1.495. This new scaling relation gave appropriate values for a selection of stars, as shown in figure 5.17 with values for  $D$  ranging between  $2\mu\text{Hz}$  and  $2.7\mu\text{Hz}$  for the first 100 brightest K dwarf stars from the catalogue.

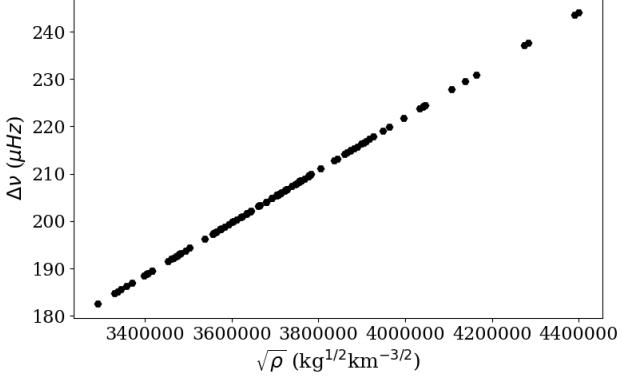


Figure 5.17:  $D$  scales linearly with the root mean density for the first 100 brightest K dwarf stars as anticipated, and the range of  $D$  values is reasonable.

The difference in the spectra being produced by the original relation and the new one was visible, as shown by the differences between figures 5.18 and 5.19. The new scaling relation produced a much more suitable spectrum for a K dwarf star in figure 5.19.

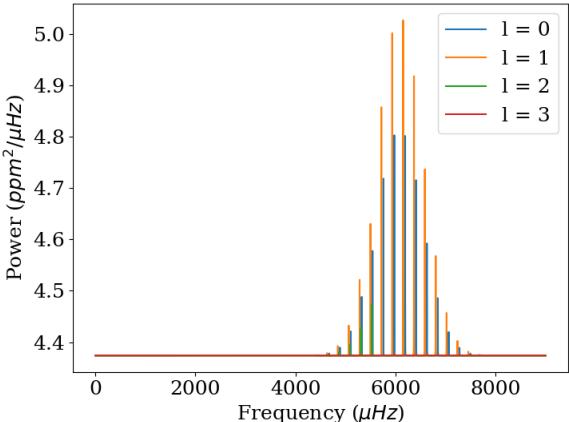


Figure 5.18: The spectra produced with the original scaling relation generated incorrect models for the p-modes. Certain modes were not plotted whilst others disappeared midway through the spectrum. The value for  $D$  produced in this spectrum was  $618.18 \mu\text{Hz}$ .

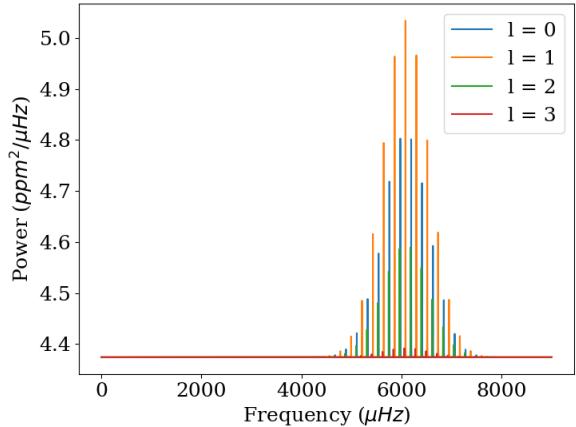


Figure 5.19: The spectra for the same star as in figure 5.18 but with the updated scaling relationship, showing the full oscillation modes. The value for  $D$  produced this time was  $2.41 \mu\text{Hz}$  which was within the range of expected values.

## 5.8 Epsilon (Sam)

The  $\epsilon$  quantity, seen in equation 4.15, is a measure of how a star's intrinsic acoustic waves interact with its surface layers as they reflect into the stellar interior, resulting in observed angular degree modes being shifted to a higher frequency. It is determined by these waves' upper and lower turning points, which are strongly dependent on conditions near the stellar surface and thus on a star's effective temperature (Gough (2003)). As such, White et al. (2012) observes a strong  $\epsilon-T_{\text{eff}}$  relation in unevolved stars. Typical values for evolved and unevolved stars fall in the range  $0.8 \lesssim \epsilon \lesssim 1.5$ , and the shifted frequency of a given angular mode can be obtained from the expression

$$\nu_{\text{nle}} = \nu_{\text{nlo}} + \epsilon \Delta\nu. \quad (5.39)$$

Understanding the mechanisms which determine  $\epsilon$  is a task motivated by the problem of mode identification in relatively massive main-sequence stars. As the p-mode peaks on a frequency-power spectrum are Lorentzians with

widths that scale with  $T_{\text{eff}}$ , massive stars with a high  $T_{\text{eff}}$  can exhibit merged radial and  $l = 2$  modes. In this case, it is often impossible to optically distinguish this merged mode from the  $l = 1$  mode, which is also usually visible; this is explored in [White et al. \(2012\)](#).  $\epsilon$  must be known to some degree of accuracy in relatively massive stars to identify their distinct angular degree modes.

This is not usually an issue in the K and M dwarfs upon which this investigation is focused due to their low effective temperatures. It has been suggested that the known frequencies of different angular degree modes in these stars could be scaled, using the strong  $\epsilon$ - $T_{\text{eff}}$  relation, to find the mode identification of more massive stars ([Bedding and Kjeldsen \(2010\)](#)). This was effectively applied by [White et al. \(2012\)](#) for F-type stars.

It also may be possible to obtain  $\epsilon$  through the use of comprehensive 3D hydrodynamical simulations of the surface layers of stars from given stellar parameters. However, these are expensive and difficult to compute accurately, in large part due to the volatility and low density of these regions and their correspondingly high local density fluctuations. Consequently, an offset exists between calculated and observed values of  $\epsilon$  for stars of all masses. Whilst [White et al. \(2012\)](#) found that it is typically around 0.2, seen in figure 5.20, this offset is not completely consistent, nor has it been related to stellar parameters, and thus it introduces an additional degree of uncertainty into models.

As models predict  $\epsilon$  values which are consistently offset from those observed, empirical data is necessary to determine this quantity accurately; the small amount of bright K dwarf frequency-power spectra collected by Kepler provide the sole source of this data.

Using Kepler data ([Borucki et al. \(2010\)](#)) con-

taining dwarf stars for which  $\epsilon$  has been unambiguously determined, the [Gough \(2003\)](#)  $\epsilon$ - $T_{\text{eff}}$  relation is observed.

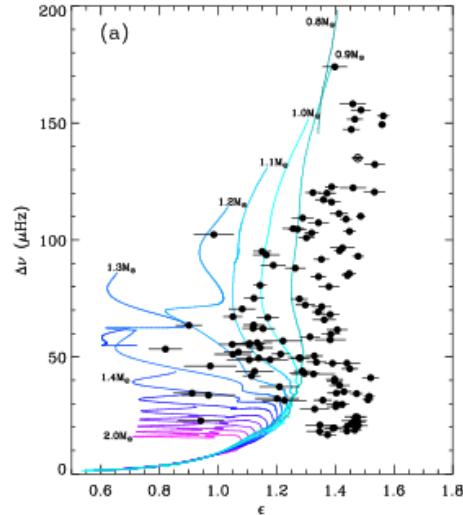


Figure 5.20:  $\epsilon$  diagram of Kepler stars for which mode identities, and thus  $\epsilon$ , have been unambiguously determined. An offset is exhibited between observed data and isochrone models. Taken from ([White et al. \(2012\)](#)).

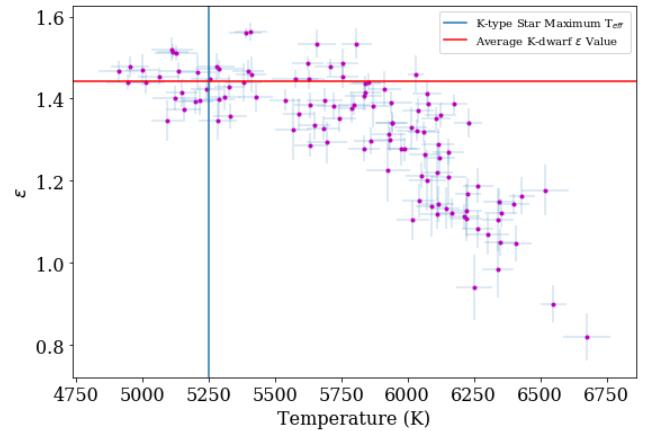


Figure 5.21: Illustration of  $\epsilon$ - $T_{\text{eff}}$  relation in dwarf stars, which appears to flatten at low  $T_{\text{eff}}$ . The vertical, blue line reflects the cut-off temperature of 5250 K; stars with  $T_{\text{eff}} \leq 5250 K$  were considered bright K dwarfs and thus were used in determining  $\epsilon$ . The horizontal, red line shows the value of  $\epsilon$  that was used.

Figure 5.21 demonstrates that the  $\epsilon$ - $T_{\text{eff}}$  relation appears to flatten out at the low  $T_{\text{eff}}$  of K dwarfs. This is further shown in figure 5.20, where isochrone tracks of lower mass stars exhibit less variation than those for higher stellar masses; data from Kepler, also shown on this diagram, appears to corroborate this trend.

Proceeding with the assumption that this relation holds for dimmer K and M dwarfs, a constant value of  $\epsilon = 1.4426$  for these stellar classes was calculated by taking the average of this quantity for all stars with  $T_{\text{eff}} \leq 5250K$ ; here, this cut-off temperature reflects the approximate maximum  $T_{\text{eff}}$  of K stars.

Values of  $\epsilon$  for red dwarfs with  $T_{\text{eff}} \leq 4800K$  are not currently observationally obtainable, meaning that mechanisms affecting conditions at the surface of stars which might arise at lower masses are not accounted for by this approximation. For example, it is known that red dwarfs' convective envelopes become increasingly large at lower masses until they are fully convective at  $\approx 0.35M_{\odot}$ ; this would doubtless affect wave transmission through a star's surface layers and, therefore, change its value of  $\epsilon$ . While reasonable, this approximation of  $\epsilon$  as a constant for K and M dwarfs is not necessarily accurate in all cases.

A discrepancy between the approximation of  $\epsilon$  and its true value will not significantly impact the visibility of p-modes in K and M dwarfs, and thus will have a negligible effect on the validity of the final results.

## 5.9 Granulation (Sharif)

The granulation spectrum is modelled as a zero-centred Lorentzian, as given by equation 5.40.

$$\langle P(\nu) \rangle = \frac{4\sqrt{2}\sigma^2}{1 + (2\pi\nu\tau)^4} \quad (5.40)$$

to calculate this for stars in the TRILEGAL database, first, a granulation spectrum was plotted for the Sun, for which scaling relations would be applied to transform the spectrum to represent the star in question.

To do this, firstly  $\sigma_{\odot}$  and  $\tau_{\odot}$  needed to be found for the Sun.  $\sigma_{\odot}$  is given by

$$\sigma_{\odot} \simeq \frac{\delta I/I}{\sqrt{n}} \quad (5.41)$$

where  $n$  is the number of visible granules on the star and is given by

$$n \simeq \frac{R^2}{r_c} \quad (5.42)$$

where  $r_c$  is the typical convective granule size and is quoted to be 1000 km (Karof (2008)) and the intensity fluctuation,  $\delta I/I$  is quoted at 5%, or 50000 ppm (Chaplin (2019)). This gives a value of  $\sigma_{\odot}$  as 71.87 ppm.  $\tau_{\odot}$  is quoted to have a value of 200 s (Chaplin (2019)). Applying these values to equation 4.10 produces figure 5.22 for the Sun.

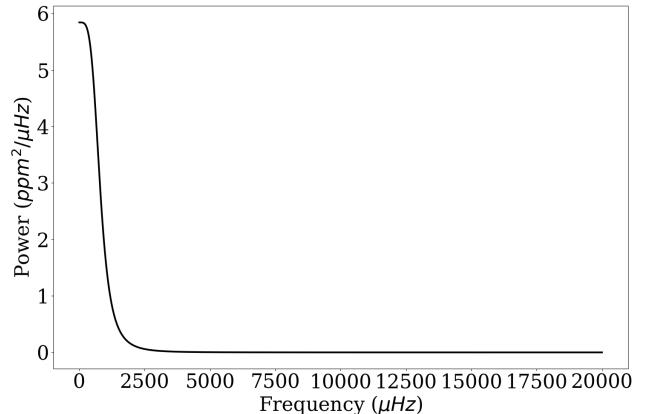


Figure 5.22: Granulation spectrum for the Sun

The scaling relations for  $\sigma$  and  $\tau$  from the Sun are given according to equations 5.43, (Chaplin (2019)).

$$\sigma = \sigma_{\odot} \left[ \left( \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} \right) \left( \frac{R}{R_{\odot}} \right) \left( \frac{M_{\odot}}{M} \right) \right] \quad (5.43a)$$

$$\tau = \tau_{\odot} \left[ \left( \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} \right)^{\frac{1}{2}} \left( \frac{g_{\odot}}{g} \right) \right]. \quad (5.43b)$$

As mentioned before, this method neglects the fact that some stars may exhibit completely different internal conditions. This may mean that if scaling relations are used for stars too different for the ones that they were defined for, they may completely fail. Their validity could have been tested by plotting empirical data for  $\sigma$  and  $\tau$  for stars from the TRILEGAL catalogue and using the `scipy optimise` module to constrain the constants set in a function which has the form of the scaling relations in equation 5.43.

In reality, no such data exists meaning that any derivation of  $\sigma$  or  $\tau$  would have been subject to the assumptions which were under investigation. Therefore, for the sake of this project, it has been assumed that these relations hold, where they can be tested once empirical data for these quantities have been obtained.

Regardless, plotting these quantities with respect to their scaling variables results in figures 5.23 and 5.24.

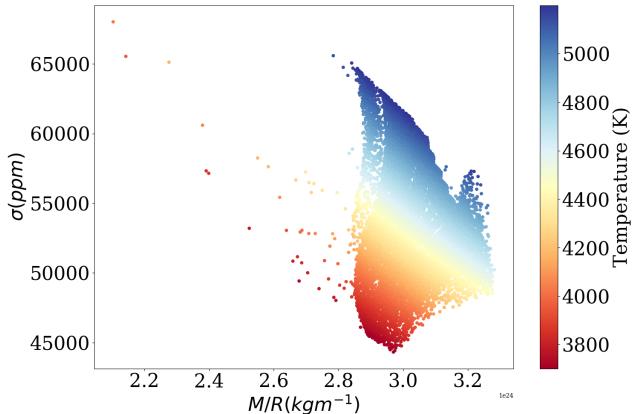


Figure 5.23: How  $\sigma$  varies as a function of  $M/R$  and  $T_{\text{eff}}$  for all the stars in the K dwarf catalogue

As expected,  $\sigma$  is inversely proportional to

$M/R$  and proportional to  $T_{\text{eff}}$  for the stars in the TRILEGAL database.

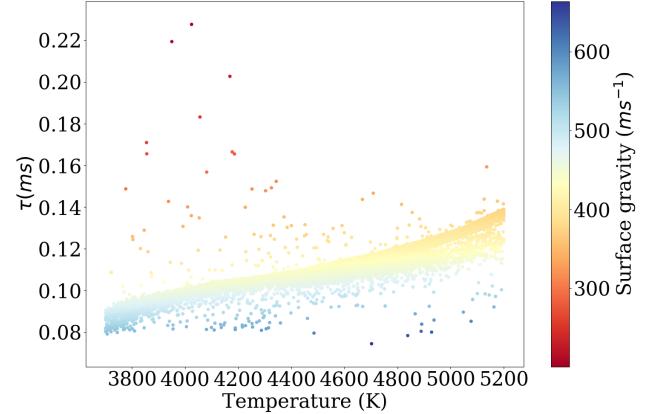


Figure 5.24: How  $\tau$  varies as a function of  $g$  and  $T_{\text{eff}}$  for all the stars in the K dwarf catalogue

As expected,  $\tau$  is proportional to  $\sqrt{T}$  and inversely proportional to surface gravity for the stars in the TRILEGAL database. The dependencies in these relations are not independent of one another, for example, one could expect stars with a greater value of  $T_{\text{eff}}$  to be larger and hence have a smaller value for surface gravity. This could explain the deviation from the  $\sqrt{T}$  relation at large  $T$  since a decreasing surface gravity for larger stars would increase  $\tau$ , resulting in the raised, slightly darker patch of stars between 4800 K and 5200 K. It will be assumed that these scaling relations hold for M and K dwarfs and shall be used to produce granulation spectra.

Applying these relations to the brightest K dwarf from the TRILEGAL data produced figure 5.25.

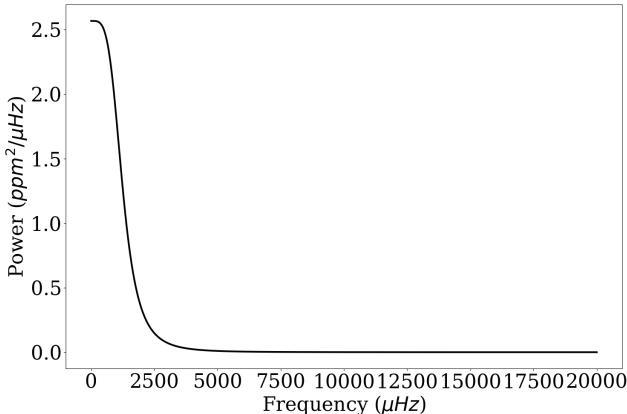


Figure 5.25: Granulation spectrum for the brightest K dwarf,  $m_v = 7.319$

## 5.10 Cadence, Radial Order (Sharif)

To create an accurate representation of the data that will be observed by PLATO, a correct number of data points needs to be used to plot the functions used in the models. This is set by the cadence of the telescope which determines the maximum frequency that can be sampled. Nyquist's theorem says that for a waveform to be properly reconstructed or resolved, there must be at least two samples taken of the waveform per period. If this is not met, an effect called aliasing becomes apparent where a wave is measured to have a lower frequency than it truly does. Equation 5.44 is used to find the upper limit of frequency that PLATO can observe.

$$\nu_{\text{Nyquist}} = \frac{1}{2\Delta t} \quad (5.44)$$

where  $\Delta t$  is the cadence of PLATO and is quoted as 25 seconds Rauer et al. (2016). This sets a maximum frequency of  $20,000\mu\text{Hz}$  where if any observing targets have resonant frequencies greater than this, PLATO would not be able to observe these. This also sets an upper limit on the resolution of the peaks that will be observed due to a finite number of samples being taken during the observation time,  $T$ ,

quoted as 2 years Rauer et al. (2016). The maximum number of bins available in the frequency spectrum is given by  $\nu_{\text{Nyquist}} \times T$ , giving 1261441 bins. These values were put into a linspace function to model the frequency range used for the stellar models.

The Nyquist frequency for the telescope also sets the number of radial modes that can be viewed for each star. Dividing  $\nu_{\text{Nyquist}}$  by the value of  $\Delta\nu$  for each star gives the maximum value of  $n$  that will be measured for each star. The smallest  $\Delta\nu$  measured for a K dwarf was  $145.7\mu\text{Hz}$  which gives a maximum radial order of 138 required to encompass all of the possible p-modes. In practice, a range up to 40 was used which was sufficient.

Additionally, the cadence may place a lower limit on the age of stars in which p-modes can be detected. As shown by equation 5.4, the frequency of maximum oscillation increases with surface gravity which decreases with age as the star expands. Therefore, it may be the case that young enough stars have surface gravities sufficiently large that their p-modes are centred at a frequency greater than what can be observed by PLATO.

## 5.11 Shot noise (Jeff)

The shot noise or photon noise is a result of random variations in the rate of photons landing on the sensor of the telescope. Since the position where the photons land cannot be predicted, there is a random distribution associated with it. If the shot noise level is high, it means that there is a wider spread than what would be expected. In a perfect detection, the photons should be hitting the centre of the sensor every time.

Shot noise is not only related to the generation of photons but is associated with quantum processes, mainly their conversion into elec-

trons within the CCD. The incoming photons hit the detector and release free electrons by the photoelectric effect which are then stored in potential wells within the pixels. The amount of electrons inside the cells at the semiconductor-oxide interface is read by electronics. Thus, the predominant source of shot noise comes from counting the electrons after incident photons have hit the detector. For instance, for 50 received photons, 40 electrons may be emitted into the CCD for the first exposure. After the second exposure, only 35 electrons may be emitted in the CCD. This is affected by the distribution of incoming photons on the detector. This uncertainty around the number of electrons is the shot noise. Once a sufficient number of electrons has been collected the shot noise will reduce significantly. This effect is therefore related to the physical properties of the telescope such as the size of the CCD.

Consequently, the shot noise is one of the main components of the background signal in the frequency-power spectra of the stellar oscillations (see figure 2.1). This value only depends on the apparent magnitude (Marcos-Arenal et al. (2014a)). The formula follows from the apparent magnitude law which relates the fluxes and magnitudes of stars.

$$m_1 = -2.5 \times \log 10 \times \left( \frac{F_1}{F_2} \right). \quad (5.45)$$

Rearranging equation 5.45 leads to:

$$\frac{F_1}{F_2} = 10^{-0.4 \times (m_1 - m_2)} \quad (5.46)$$

where  $F_1$  and  $F_2$  are the fluxes of star 1 and star 2 with respective magnitudes  $m_1$  and  $m_2$ . The shot noise follows Poisson's statistics with the discrete photons (Beenakker and Schomerus (2001)). For a large number of photons, Poisson statistics would follow a Gaussian or a normal distribution. Then, the noise associated to the flux scales as  $\sqrt{F}$ . The RMS noise,  $\sigma$  in time is inversely proportional to  $\sqrt{F}$ . From

Chaplin et al. (2011),  $\sigma$  is given as  $2\sigma^2\Delta t$  where  $\Delta t$  represents the mean cadence in seconds.

Hence, substituting this into equation 5.46 gives:

$$2\sigma_2^2\Delta t = 2\sigma_1^2\Delta t \times 10^{-0.4 \times (m_1 - m_2)}. \quad (5.47)$$

The PLATO mission is due to focus on stars brighter than 11 in V-band magnitudes (Marcos-Arenal et al. (2014a)). Therefore, for  $m_1 = 11 \times 2\sigma_1^2\Delta(t)$ , the shot noise level should be  $18 \text{ ppm}^2 \mu\text{Hz}^{-1}$  (ibid). Thus, the shot noise is defined as:

$$\text{shot noise} = 18 \times 10^{-0.4(11 - m_2)}. \quad (5.48)$$

This shows that as the target gets fainter and smaller this will increase. As the signal to noise ratio gets smaller, it will become increasingly hard to see or analyse the p-modes of the star. This is likely to be one of the reasons why red dwarfs oscillations are hard to detect.

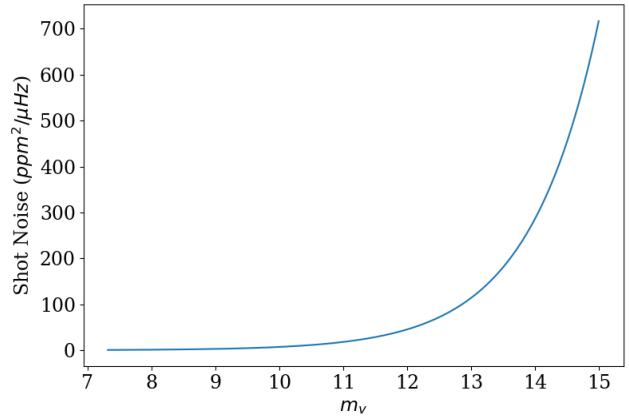


Figure 5.26: Shot noise as a function of  $m_V$

This background signal is frequency-independent and only depends on the apparent magnitude of the target (see figure 5.26). It acts as an offset line that shifts the power of the oscillations (see figure 5.27).

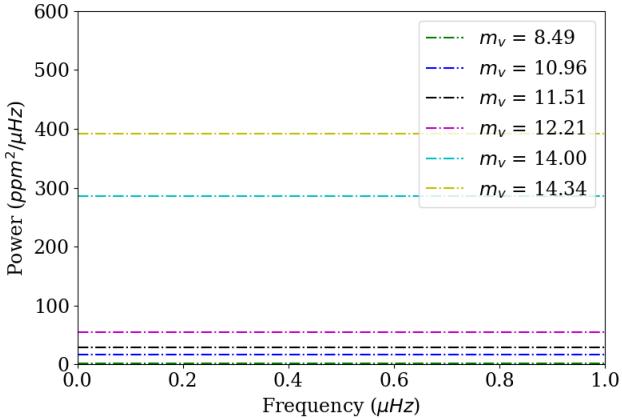


Figure 5.27: Shot noise for K dwarfs of different magnitudes. This shows the offset line

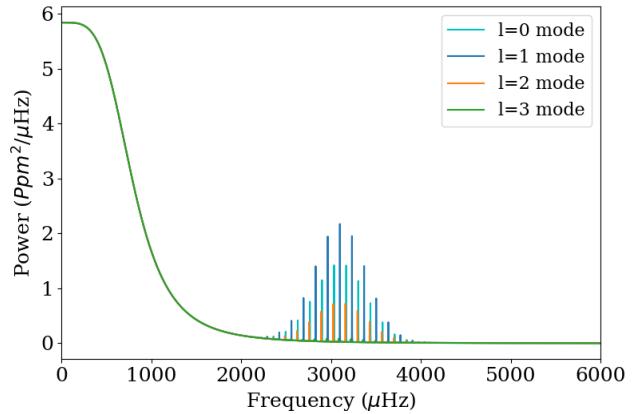


Figure 6.1: Power-frequency spectrum for the Sun

## 6 Models/Results

### 6.1 The Sun Spectrum (Maddie T)

Before building the spectra for the TRILEGAL simulated K and M dwarfs, the focus was applied to model a frequency-power spectrum for the Sun to test the model and initial detection methods.

The Sun is an essential tool to develop asteroseismic techniques and compare with spectra of more distant stars. Viewing the Sun from Earth, there is negligible shot noise and hence the oscillations are much easier to detect. The Sun has also been widely studied and the field of helioseismology provides abundant data and spectra which allow the modelled spectrum to be compared to real-life data.

To quantify a detection for the K dwarfs and M dwarfs observed by PLATO, detection techniques used by the Data Analysis group had to be tested on the initial solar model. This should yield a guaranteed detection of oscillation modes, otherwise indicating a fault in the detection methods.

Figure 6.1 presents the modelled spectrum for the Sun which incorporated the granulation, shot noise and relative visibilities of the modes of different angular degree. Values of  $A_{\max,\odot}=2.1$  ppm and  $T_{\text{eff}}=5777$  K were again used.  $D_{\odot}$  was taken to be  $\approx 1.5$  and  $\epsilon$  was also taken to be  $\approx 1.5$  which is a reasonable assumption based on the  $\epsilon-T_{\text{eff}}$  relation depicted in figure 5.21. Typical values for the solar radius, luminosity, surface gravity,  $\nu_{\max,\odot}$  and  $\Delta\nu_{\odot}$  were obtained from the Astropy constants library (Robitaille et al. (2013)).

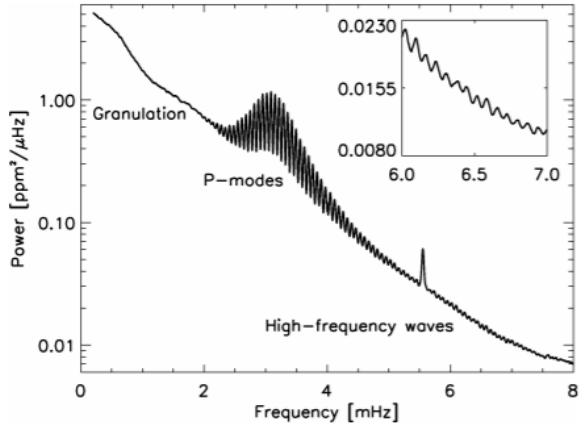


Figure 6.2: Solar frequency-power spectrum from Karoff and Kjeldsen (2008) data. Y-axis is on a log scale.

Figure 6.2 shows a frequency-power spectrum, where the power is on a log scale, plotted using real solar data (Karoff and Kjeldsen (2008)). For determining the detection of p-modes, the modelled spectra will not include artefacts at higher frequencies such as the peak seen in figure 6.2 at  $\approx 5.5$  mHz.

To compare more easily with the modelled spectrum, this is plotted on the same log scale in figure 6.3.

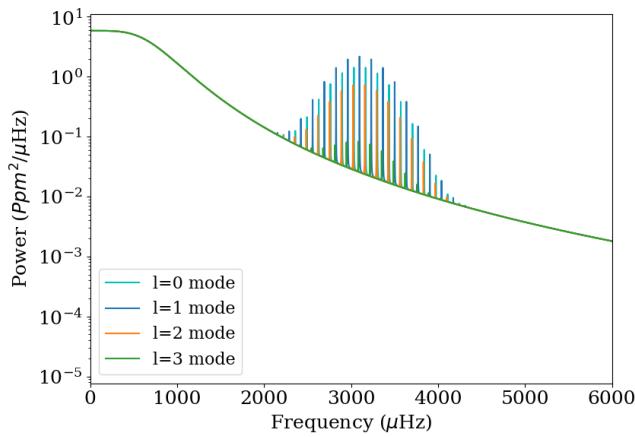


Figure 6.3: Modelled frequency-power spectrum for the Sun with power on a log scale.

It can be seen that the acoustic oscillation modes in figure 6.2 and figure 6.3 both lie between 2000-4000  $\mu$ Hz (2-4 mHz). The maximum amplitude of the p-modes in figure 6.2 is just over 1 which agrees with the maximum amplitude of the radial ( $l = 0$ ) modes in figure 6.3. Consequently, the constructed model can be confidently accepted as a good basic imitation of stellar oscillation spectra.

Solar oscillations within the field of helioseismology have been extensively studied by projects such as BiSON (Birmingham Solar Oscillations Network). This makes the solar oscillation spectrum a critical tool against which to calibrate models for solar-like stars as has been demonstrated by the various previously derived

scaling relations (Kjeldsen et al. (2008a)). In this project, it was the first step in ensuring that not only did the modelling techniques match real-life results, but also, that the detection techniques used by the Data Analysis group worked to a level sufficient at least to detect p-modes with little to no noise.

Nevertheless, PLATO will focus on K and M dwarf stars which are more evolved, generally smaller, cooler and therefore, dimmer than the Sun. This outlines the limitations of the solar model as an indication of what PLATO may be able to detect. Once the Data Analysis group has confirmed that oscillations can indeed be detected from the solar modelled spectrum, this will be developed into more relevant K and M dwarf models.

## 6.2 Models of K dwarfs (Maddie T and Victoire)

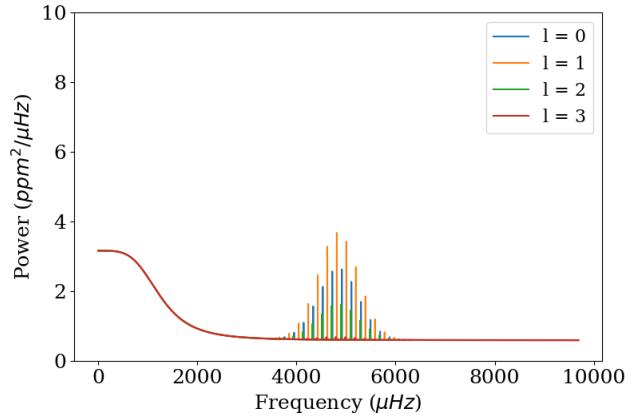


Figure 6.4: Frequency-power spectrum for the brightest K dwarf star ( $m_v = 7.3$ ) in the first catalogue of simulated data.

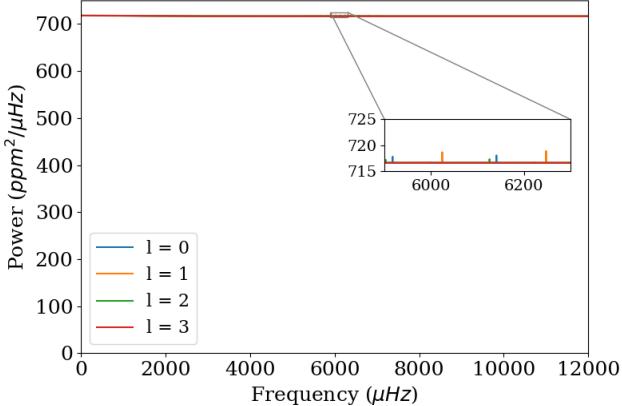


Figure 6.5: Frequency-power spectrum for the dimmest K dwarf star ( $m_v = 15$ ) in the first catalogue of simulated data.

Figure 6.4 shows the spectrum for the brightest K dwarf in the catalogue and figure 6.5 the spectrum of the dimmest one. Table 3 introduces the properties that are known from TRILEGAL data and the ones that were calculated from scaling relations that allowed to define the spectrum.

	<b>K1</b>	<b>K2</b>
Luminosity ( $L_\odot$ )	0.290	0.0900
Radius (km)	498000	426000
Mass ( $M_\odot$ )	0.751	0.629
Surface gravity (cm/s <sup>2</sup> )	39900	45800
Temperature (K)	5020	3990
V Magnitude	7.32	15.0
$\nu_{\max}$ (μHz)	4840	6240
$\Delta\nu$ (μHz)	193	223
$\delta\nu_{02}$ (μHz)	12.8	14.8
$\Gamma$ (μHz)	0.176	0.0730
H at $\nu_{\max}$ (ppm <sup>2</sup> /μHz)	2.57	1.50
Shot noise (ppm <sup>2</sup> /μHz)	0.61	717

Table 3: K dwarfs properties for the brightest (K1) and dimmest (K2) in the catalogue (3 sig. fig.). The first set of properties represents the properties obtained from TRILEGAL. The second set was calculated from scaling relations using the first set.

The intrinsic properties of stars show that they are not at the same evolutionary state and have very different properties. For instance, there is a temperature difference of just over 1000 K. This will have an impact on the calculated values from scaling relations since  $\nu_{\max}$ , the width, the height (indirectly, through the width), etc. depending on temperature.

Across the whole catalogue, the width ranges between 0.249 and 0.07 μHz. As predicted, the values seem slightly smaller than what would be expected for late K dwarfs due to the approximations that were made when extrapolating the equation 5.32.

Both spectra follow the expected Gaussian shape with the  $\nu_{\max}$  defining the position of the peak. The width is significantly smaller for the dimmest dwarf with a value of 0.073 μHz. This, in turn explains why the amplitude of oscillations is higher for the brighter star (cf. equation 5.33).

It can be seen that the shot noise level for the dimmest star is very high (as expected since it is directly related to the magnitude of the star). Therefore, the power per μHz reflects more the impact of shot noise than the actual oscillations.

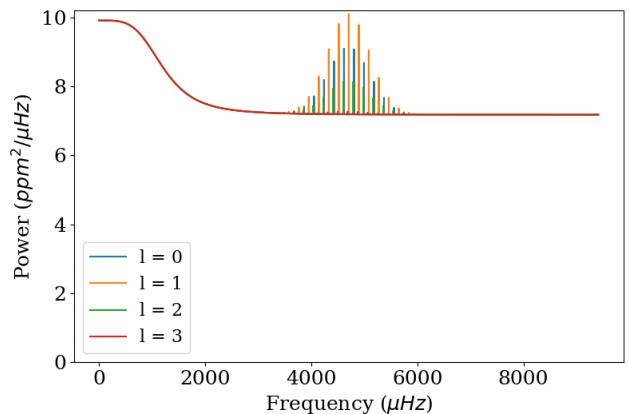


Figure 6.6: Frequency-power spectrum for a 10 magnitude K dwarf star in the first catalogue of simulated data.

The spectra in figures 6.4 and 6.5 represent the two extremes of the simulated PLATO catalogue ranging from a V-band magnitude of 7.3 to 15. Nevertheless, for PLATO’s input catalogue, target stars are specified to be F5 to K7 types with a V-band magnitude  $< 13$  (Heras et al. (2018)). This implies that the mission does not expect to observe oscillations in stars with a magnitude greater than 13. The data analysis section will aim to place a stringent limit on the magnitudes of stars that PLATO will be able to detect oscillation modes for; this will be expected to be brighter than 13. To present a more realistic possible limit to detectable oscillations in K dwarfs, figure 6.6 shows the frequency-power spectrum for a K dwarf of magnitude 10.

	<b>K3</b>
Luminosity ( $L_{\odot}$ )	0.327
Radius (km)	510000
Mass ( $M_{\odot}$ )	0.768
Surface gravity ( $\text{cm/s}^2$ )	39100
Temperature (K)	5100
V Magnitude	10.0
$\nu_{\max}$ ( $\mu\text{Hz}$ )	4710
$\Delta\nu$ ( $\mu\text{Hz}$ )	188
$\delta\nu_{02}$	12.5
$\Gamma$ ( $\mu\text{Hz}$ )	0.206
H at $\nu_{\max}$ ( $\text{ppm}^2/\mu\text{Hz}$ )	2.73
Shot noise ( $\text{ppm}^2/\mu\text{Hz}$ )	7.18

Table 4: K dwarfs properties for the  $m_v=10$  star (K3) (3 sig. fig.)

Table 4 shows the intrinsic properties of this star and the values used for plotting the spectrum which can be calculated from these properties.

The spectra here do not yet include noise due to the intrinsic random nature of photometry. This makes it difficult at this stage to predict which stars will allow for oscillations to be de-

tected. From looking at the spectra alone, the only indication of detectability will be the ratio of the height of the oscillations to the shot noise.

For K1 for example, the maximum height of the radial modes is  $2.57 \text{ ppm}^2\mu\text{Hz}^{-1}$  and the shot noise due to intrinsic brightness is  $0.61 \text{ ppm}^2\mu\text{Hz}^{-1}$ . This gives a ratio of height to shot noise of 4.2. For K2 however, a ratio of 0.002 is achieved which implies this star will be much harder to detect. K3 gives a middling ratio value of 0.38. These values indicate that the relative detectabilities of different stars however, they do not reveal anything about the detection limits of PLATO. The following data analysis will aim to quantify the limit of detection.

Figure 6.7 visually demonstrates the increase in  $\nu_{\max}$  and  $\Delta\nu$ , and the subsequent fall in the power of maximum oscillation with decreasing temperature and luminosity. The height of the granulation also decreases with decreasing temperature.

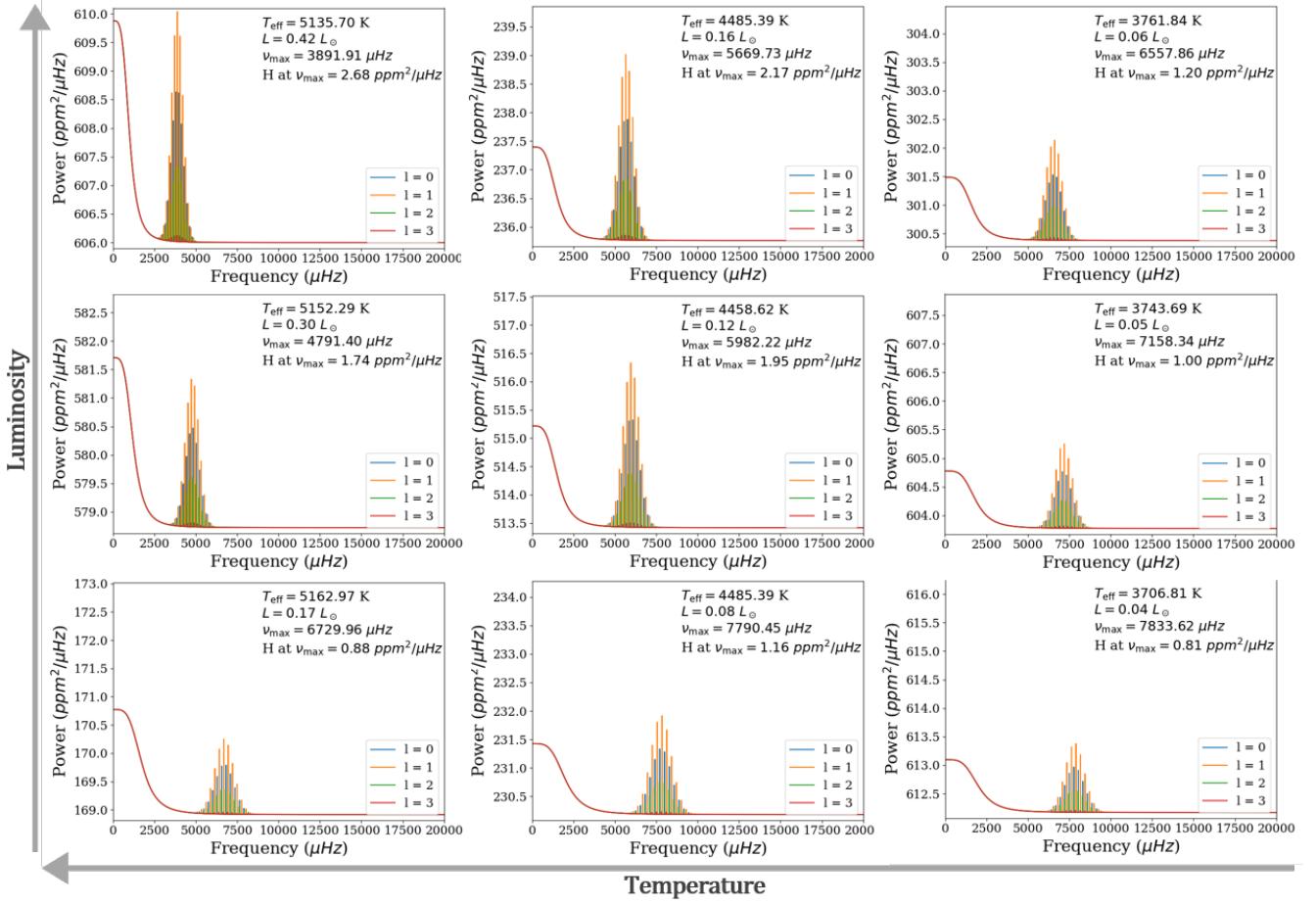


Figure 6.7: Frequency-power spectra of nine K dwarfs across a range of temperatures and luminosities.

### 6.3 Models of M dwarfs (Chamini)

	M1	M2
Luminosity ( $L_{\odot}$ )	0.023	0.017
Radius (km)	291323	273008
Mass ( $M_{\odot}$ )	0.428	0.408
Surface gravity (cm/s <sup>2</sup> )	66374.31	71121.35
Temperature (K)	3476.2	3324.3
V Magnitude	8.02	15.0
$\nu_{\max}$ ( $\mu\text{Hz}$ )	9680.8	10607.5
$\Delta\nu$ ( $\mu\text{Hz}$ )	325.4	347.7
$\delta\nu_{02}$ ( $\mu\text{Hz}$ )	21.6	23.10
$\Gamma$ ( $\mu\text{Hz}$ )	0.0704	0.0702
H at $\nu_{\max}$ ( $\text{ppm}^2/\mu\text{Hz}$ )	0.441	0.324
Shot noise ( $\text{ppm}^2/\mu\text{Hz}$ )	1.15	716.5

Table 5: M dwarfs properties for the brightest(M1) and dimmest (M2) in the catalogue. The first set of properties are the ones from TRILEGAL. The second set was calculated from scaling relations the first set.

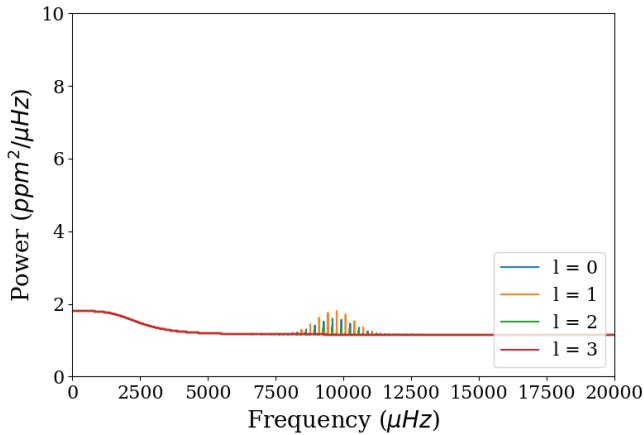


Figure 6.8: Frequency-power spectrum for the brightest M dwarf star ( $m_v = 8.02$ ) in STEP01. The y-scale matches that of Figure 6.4 for comparison.

Figure 6.8 shows the spectrum for the brightest M dwarf in the catalogue and Figure 6.9 the

spectrum of the dimmest one. Table 5 shows the properties of the respective spectra.

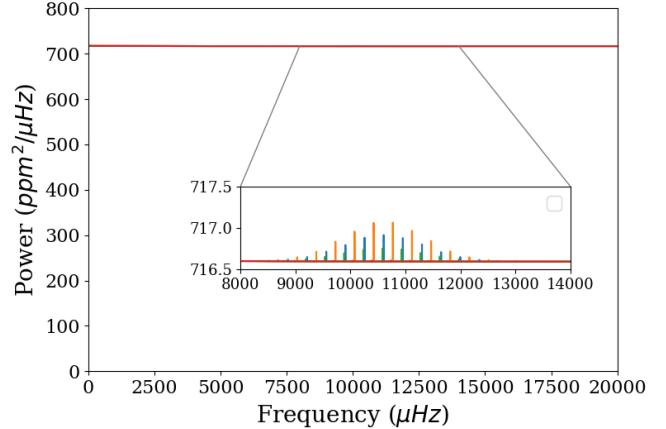


Figure 6.9: Frequency-power spectrum for the dimmest M dwarf star ( $m_v = 15$ ) in STEP01.

The  $\nu_{\max}$  of M dwarfs in STEP01 ranges from 6789.88 to 16080.97  $\mu\text{Hz}$  (as seen in Figure 6.11). These relatively large figures (compared to K dwarfs' range of 3891.91 to 8323.62  $\mu\text{Hz}$ ) are in accordance of equation 5.4, as  $\nu_{\max}$  is proportional to the surface gravity,  $g$ , and inversely proportional to the effective temperature,  $T_{\text{eff}}$ , and M dwarfs have some of the lowest temperatures and highest surface gravities in the main sequence. The large frequency separation,  $\Delta\nu$ , which ranges from 233.89 to 539.80  $\mu\text{Hz}$  in STEP01 M dwarfs, is proportional to  $\nu_{\max}$  (Stello et al. (2009)). As  $\Delta\nu$  is proportional to the width of the Gaussian curve governing the overall p-modes (equation 5.21), the relatively large  $\Delta\nu$ s exhibited by M dwarfs lead to larger widths, and smaller peaks. The linewidth remains consistent as it ranges from 0.070 to 0.071  $\mu\text{Hz}$  while the maximum height of the peaks above the granulation and shot noise ranges from 0.101 to 1.26  $\text{ppm}^2/\mu\text{Hz}$ .

Dividing the maximum heights by each shot noise leads to the fractional height above the shot noise to range from 0.0002 (Figure 6.10) to 0.3822 in the brightest star (Figure 6.8). These were predicted to be the least and most likely

M dwarf detections in STEP01.

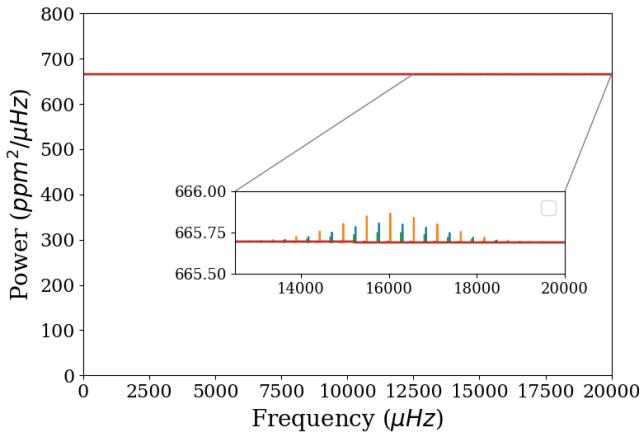


Figure 6.10: Frequency-power spectrum for the smallest amplitude relative to the shot noise for an M dwarf star in STEP01.

dwarfs, is comparable to the bottom right spectrum of Figure 6.7, the coldest and dimmest of K dwarfs, as expected.

	<b>M3</b>
Luminosity ( $L_{\odot}$ )	0.005
Radius (km)	168102.99
Mass ( $M_{\odot}$ )	0.222
Surface gravity (cm/s <sup>2</sup> )	103753
Temperature (K)	3104.56
V Magnitude	14.92
$\nu_{\max}$ (μHz)	16012.62
$\Delta\nu$ (μHz)	534.7
$\delta\nu_{02}$	35.54
$\Gamma$ (μHz)	0.070
H at $\nu_{\max}$ (ppm <sup>2</sup> /μHz)	0.116
Shot noise (ppm <sup>2</sup> /μHz)	666

Table 6: "Known" and predicted M dwarf properties for star corresponding to the smallest  $H_{\max}$ -Shot noise ratio (Figure 6.10) in the TRILEGAL STEP01 simulation.

Like Figure 6.7, Figure 6.11 (using the same incremental scale) visually shows the relationship between temperature, luminosity and the  $\nu_{\max}$ ,  $\Delta\nu$ , width and height of the p-mode oscillations above the shot noise. The top left spectrum, one of the most luminous and hottest of the M

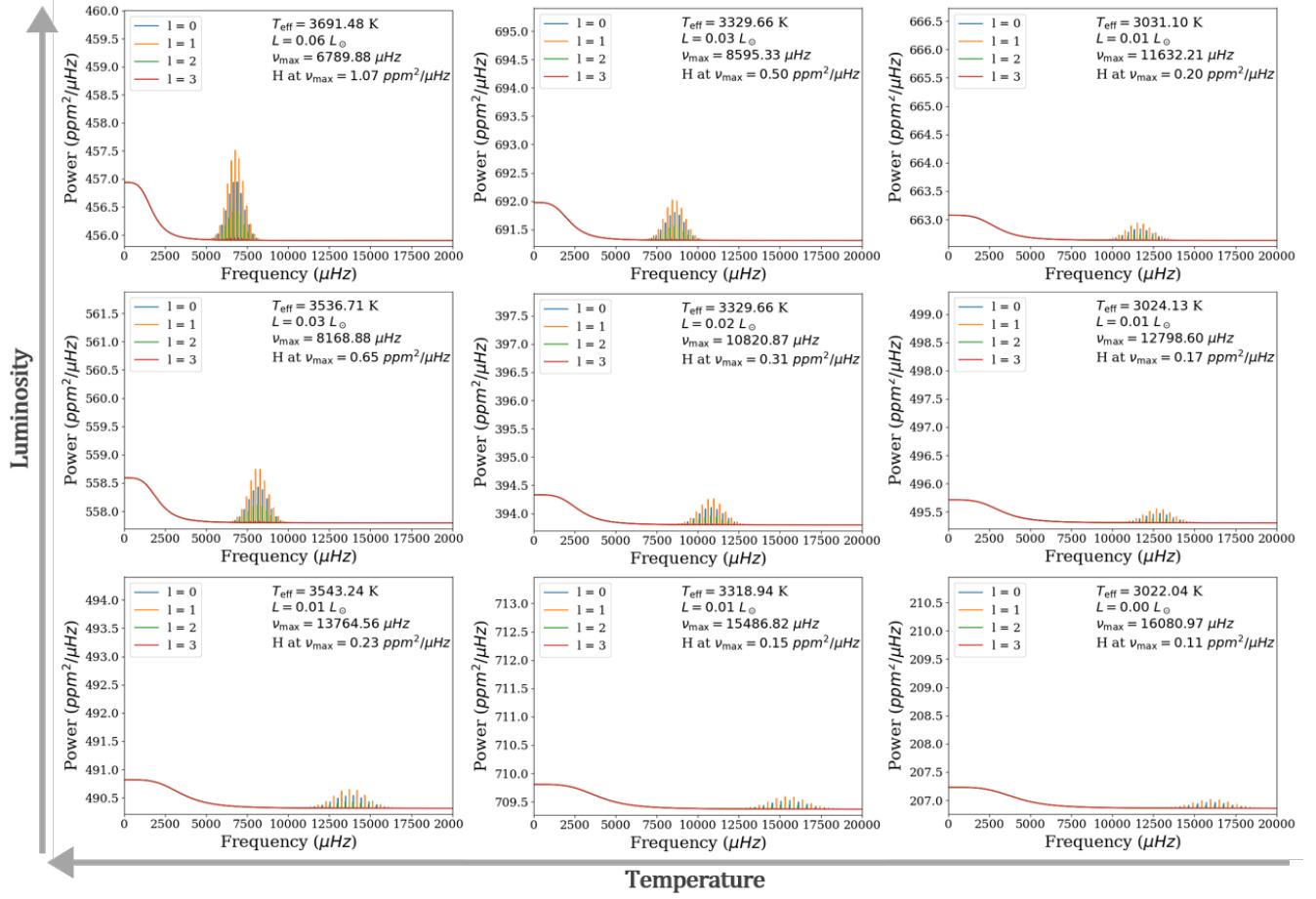


Figure 6.11: Frequency-power spectra of nine M dwarfs across a range of temperatures and luminosities.

## 6.4 Motivation for the targets selected (Chamini)

With 61229 K and M dwarfs determined across the three catalogues, the targets selected to be passed on for Data Analysis needed to be justified. As the random noise added to the spectra and potentially concealing the p-modes was proportional to the shot noise, it was decided that targets with the smallest apparent magnitudes would be the most likely to produce detections, and the 40 brightest K dwarfs and M dwarfs from STEP01 were initially sent over.

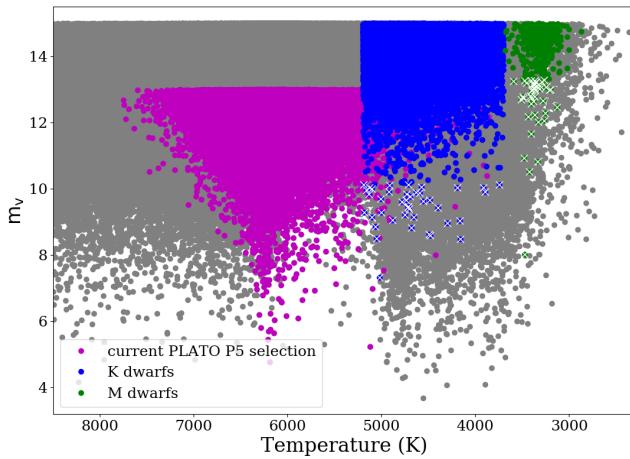


Figure 6.12: Graph showing the magnitudes and effective temperatures of STEP01 and the brightest 40 K dwarfs and M dwarfs chosen for analysis in white.

These stars were randomly distributed across the K and M dwarf regions of the CMD (Figure 6.13) and represented a range of temperatures, masses and luminosities.

Next, the brightest K dwarf star was then artificially projected to various distances to simulate evenly spaced apparent magnitudes which were then analysed to determine the magnitude at which the random noise from the shot noise overwhelmed the p-mode spectra and rendering it undetectable.

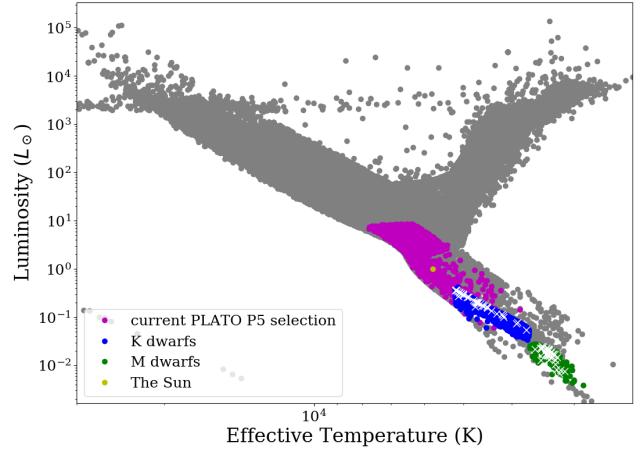


Figure 6.13: HR diagram of STEP01 and the brightest 40 K dwarfs and M dwarfs (in apparent magnitude) chosen for analysis in white.

Through trial and error, it was determined that the two main factors influencing detections were the added random noise which was determined via the shot noise, and the maximum height of the p-mode oscillations above the background noise, which was inversely proportional to the  $\nu_{\max}$  (Figure 6.14).

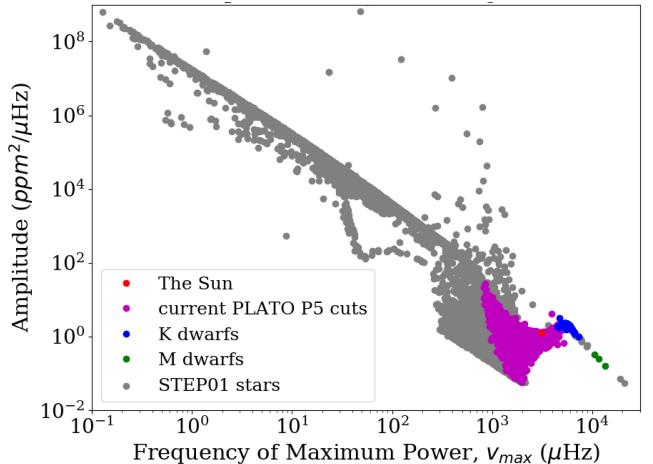


Figure 6.14: The inverse relationship between  $\nu_{\max}$  and  $A_{\max}$ . As shot noise is of the order  $\sim 10^0 - 10^2$  for  $6 < m_v < 15$ , the p-mode oscillations in dwarfs are particularly hard to resolve.

From figure 5.1 and figure 6.17 it can be seen

that  $\nu_{\max}$  increases with increasing surface gravity and decreasing temperature.

A combination of a low level of shot noise, and a small  $\nu_{\max}$  (therefore a larger peak in maximum oscillations), leads to a more accurate detection of the  $\nu_{\max}$  itself. To determine the limits on these parameters for seismology, the distribution of  $\nu_{\max}$  against the V-band apparent magnitudes that governed shot noise (Section 10.2) was studied (Figure 6.15 for K dwarfs and Figure A6 for M dwarfs).

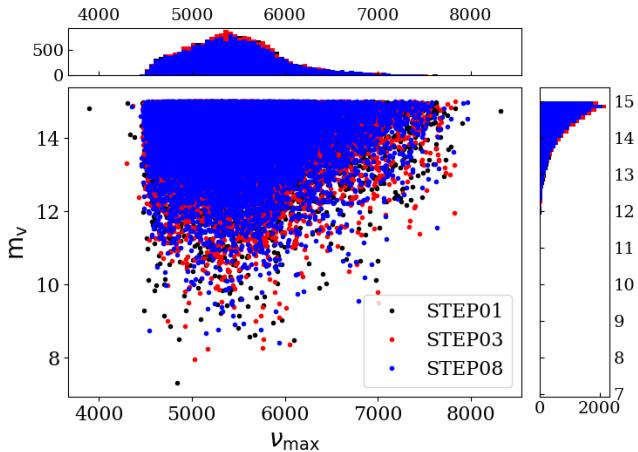


Figure 6.15: Distribution of  $\nu_{\max}$  of K dwarfs in STEP01, 03, & 08 against their apparent visual magnitudes. Histograms showing the distribution of apparent magnitudes and  $\nu_{\max}$  show a sharp exponential decrease in the number of targets at decreasing magnitudes and a modal  $\nu_{\max}$  of  $\sim 5400 \mu\text{Hz}$ .

Due to a lack of bright, low  $\nu_{\max}$  targets, an artificial grid of evenly spaced  $\nu_{\max}$ s from the target list, strategically placed at a range of magnitudes was created. As only the visual magnitude of the chosen simulated star was altered, this could be compared to “placing” the same star at a varying distance away from PLATO. From trial and error, various grids with different ranges of  $\nu_{\max}$  and  $m_v$  were created. These stars were then analysed to create a grid mesh of spectra that could be used to determine the

limits in detection, and then be compared to Figure 6.15, the actual simulated population of K dwarfs. For example, Figure 6.16, with a maximum  $\nu_{\max}$  of 5300 and 0.25 increments in brightness was used to narrow down the limits in  $m_v$ .

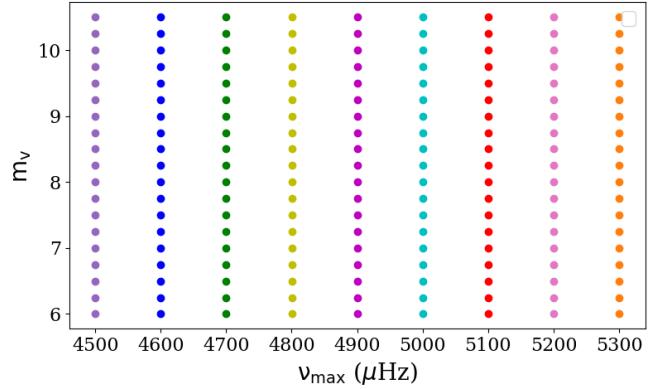


Figure 6.16: Sample graph of a grid of evenly spaced  $\nu_{\max}$ s placed at an incrementally increasing series of brightness. Each colour represents a simulated K dwarf from TRILEGAL that has been projected to several “artificial” magnitudes.

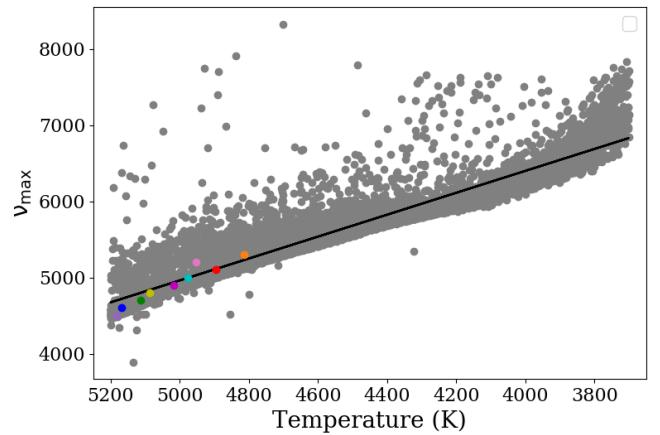


Figure 6.17: The effective temperature,  $T_{\text{eff}}$  and  $\nu_{\max}$ s of the stars chosen for the final artificial grid of K dwarf targets, compared to that of the K dwarfs identified in the STEP01 simulation from TRILEGAL. The line of best fit ( $\nu_{\max} = -1.4374 T_{\text{eff}} + 12148$  (to 5 s.f.)) is shown in black.

From equation 5.4, and figure 5.1 the inverse relationship between  $\nu_{\max}$  and Temperature can be used to realise those limits in Effective Temperature, which can then be used to determine the final limits of detection.

After determining the limits of detection in  $\nu_{\max}$ , the values can then be approximately converted into best estimates for the minimum temperature (Figure 6.17) and maximum surface gravity (Figure A8) expected for resolvable oscillations by using the best-fit lines.

## 6.5 Check of the spectrum (Maddie O)

When checking the scaling relations, however, it was important not to check them against the properties of the stars that were used in the calculations of the individual parameters themselves, such as with  $\nu_{\max}$ . The value of  $\nu_{\max}$  scales proportionally to  $\log(g)$  and inversely proportional to  $T_{\text{eff}}$ . However, as both of these variables are used to calculate  $\nu_{\max}$  so were not effective in ensuring the values being produced were accurate. As  $\nu_{\max}$  decreases as the star evolves, due to the surface gravity changing more rapidly than the effective temperature of a star, it must be inversely proportional to the  $\log(\text{age})$  of the star.

As shown in figure 6.18, this generally held for a random selection of K dwarf stars. However, the scaling relation quickly became less accurate for M dwarf stars, as demonstrated by figure 6.19. In this figure, there is a less well-defined correlation between age and  $\nu_{\max}$ , implying that alternative scaling relations may be more appropriate for the smallest of main-sequence stars. The reasoning for the failings in the scaling relations for M dwarfs could be explained by low mass stars having a varying "sound speed profiles" (Chaplin and Miglio (2013)), which would impact on the value of  $\nu_{\max}$  as this is proportional to the acoustic cut-

off frequency.

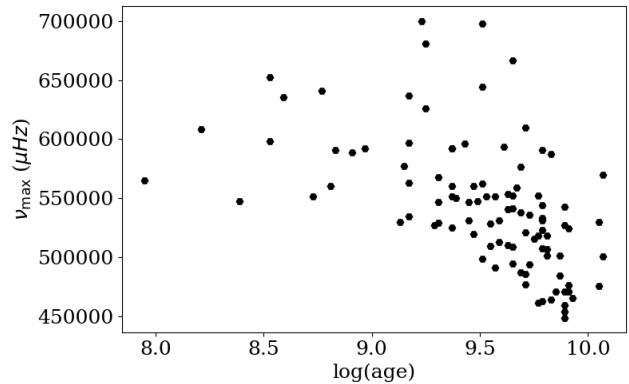


Figure 6.18: For K dwarfs, the scaling relation used to calculate  $\nu_{\max}$  appears to hold reasonably well, with the value of  $\nu_{\max}$  trending generally downwards as the stellar age increases.

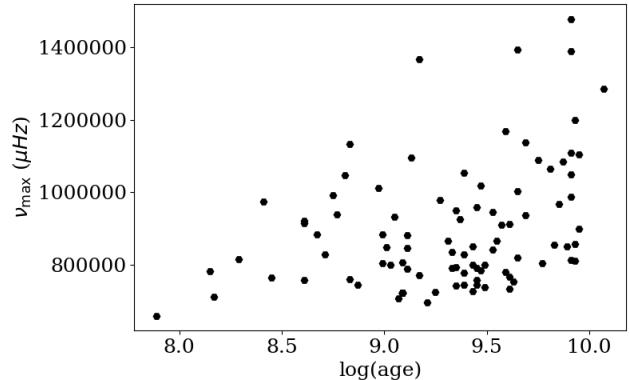


Figure 6.19: For M dwarfs, the scaling relation used to calculate  $\nu_{\max}$  does not appear to produce as accurate values as hoped, with the value of  $\nu_{\max}$  appearing to either have no trend or even a slight upwards trend as age increases.

The large frequency separation  $\Delta\nu$  and  $\nu_{\max}$  should also be related according to

$$\Delta\nu \propto \nu_{\max}^{0.77}. \quad (6.1)$$

as discussed in Stello et al. (2009). For a selection of K and M dwarfs this appeared to hold

with the simulated data, even for the smallest of stars available, as demonstrated in figures 6.21 and 6.20.

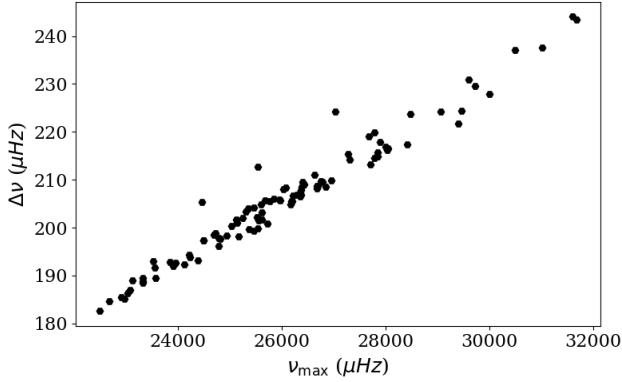


Figure 6.20: The relationship between  $\Delta\nu$  and  $\nu_{\max}$  determined by [Stello et al. \(2009\)](#) does hold for a range of K dwarf stars.

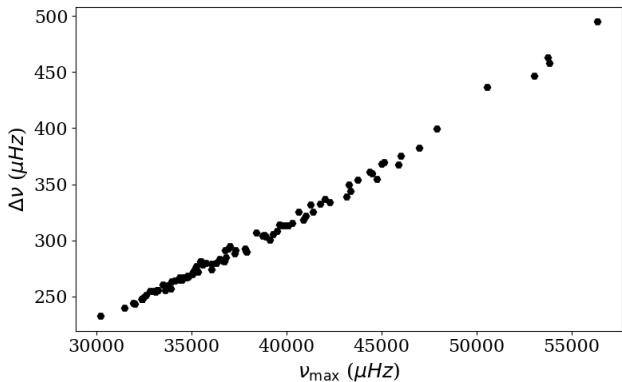


Figure 6.21: For M dwarfs, the relationship between  $\Delta\nu$  and  $\nu_{\max}$  also appears to hold, with a direct correlation between the two values for even the lowest mass stars.

The scaling relations for  $\Delta\nu$  and  $\nu_{\max}$  were the two main ones that needed to be checked to ensure they worked for K and M dwarf stars as these were the ones that had not been tested for stars other than solar-like stars or red giants. Overall the scaling relations held well for K dwarf stars but began to deviate from expected values once reaching the lowest mass stars, the M dwarfs. Unfortunately, the height of the

Lorentzian peaks did depend on these values so this inaccuracy in the scaling relations may have reduced the detectability of the p-mode oscillations in the lowest mass stars. Without knowing how  $\Delta\nu$  and  $\nu_{\max}$  change in low mass stars compared to solar-like stars, however, it would be difficult to determine new relations that work better so this is something that may need to be investigated further.

## 7 Discussion

### 7.1 Approximations Made (Maddie T)

The spectra provided by the TSM group for data analysis represented the best attempt at modelling the p-mode oscillations for K and M dwarfs that PLATO might observe. A spectrum was constructed by modelling constituent components separately and adding these together. As is unavoidable, the final spectra provided were imperfect compared to real data due to approximations made during modelling.

Firstly, for multiple components, scaling relations were used which do not hold for all stellar types and are instead extrapolated for K and M dwarf, late main-sequence stars. For  $\nu_{\max}$ , the relation was found to not hold for the hottest F dwarf stars which implies that the relation may be imperfect for red dwarfs, and hence the K and M dwarfs targeted here. An incorrect  $\nu_{\max}$  would shift the centres of the oscillation modes for each spectrum and affect detectability at the analysis stage.

The linewidth of the Lorentzian peaks was another value extrapolated from a scaling relation found from fitting red giants and solar-like stars. This resulted in the output values being smaller than expected for K dwarf spectra. This was the best approximation of the three methods for estimating the width however high-

lights another imperfection in the modelled spectra.

The value of  $\epsilon$ , used when calculating mode frequencies, was difficult to model as values obtained from models in the past were frequently offset from observed values. Here, the value was assumed to be constant for K and M dwarfs. Nevertheless, an offset in the value of  $\epsilon$  would only shift the positions of the modes in frequency and not affect power or visibility.

Determining the value of  $D$ , which was also used to calculate mode frequencies, involved another approximated scaling relation. This relation assumed that the dynamic time was similar to the time a sound wave takes to cross a star. Nonetheless, the  $D$  parameter affects the spacing of the modes so if this value was slightly offset, it should not affect the visibility of modes, crucial to quantifying a detection limit.

Finally, the modelled spectra assumed that the stars are non-rotating which is not realistic and not what will be observed by PLATO. The non-rotating model negates any effects on the visibility due to different angles of the stellar rotation axis. This may make oscillation modes significantly more or less detectable depending on the angle. The effects of rotational splitting were also discussed; the addition of which would ultimately improve the model.

## 7.2 Satellite Jitter (Sam)

There are also factors which contribute to the star's frequency-power spectrum that could not be accounted for due to time constraints and the limitations of the data available in the synthetic catalogue.

Intrinsic noise in all observational data obtained using a satellite is difficult to quantify. The main contributor to this noise is satellite

jitter, which increases the noise level of individual exposure frames taken by a satellite. This effect arises due to subtle, high-frequency variations in the direction in which the satellite's telescope is pointing; if this occurs on much shorter timescales than the 25s cadence of PLATO, images will move significantly on the telescope's CCD. As the pixel response of a given CCD is non-uniform over its surface, this jitter ultimately causes a small but non-negligible and inconsistent drop in measured flux, an additional intrinsic source of noise in observations.

The significance of this effect must be evaluated for each frame when observing a target if satellite jitter is to be accounted for. As such, it requires significant computational time and was thus not feasible in this investigation. However, the fact that quantifying this source of noise is both possible and effective in more accurately simulating PLATO's observational data is demonstrated in [Marcos-Arenal et al. \(2014b\)](#).

## 8 Intro DA (Luke)

The task the Data Analysis (DA) group set out to achieve was to discover whether or not it was possible to make detections of p-mode oscillations in K and M dwarf stars. This was done to determine the feasibility of collecting asteroseismic data for these stars, which could be used to place constraints on their physical properties and ages, furthering the fields of exoplanet study and characterisation and that of stellar formation and evolution.

To achieve this, models would be taken from the target selection and modelling (TSM) group and add a random noise component, simulating observational noise and creating noisy realisations of the data. The effect of doing this would act to obscure both the defined individual oscillatory modes and the general shape they form within the data, making successful detections far more challenging. Since the effect increases with decreasing absolute magnitude and increasing distance, the challenge of making detections in the target stars is compounded further.

To overcome this, it would be required to devise different methods of making detections and test their functionality and effectiveness using pre-existing Kepler data and solar oscillation models. This would allow the methods to be refined and improved upon before their application to the target K and M dwarf star models, with the design being to smooth out and resolve any issues before testing on largely unknown quantities. As such, any issues encountered could be more effectively approached.

## 9 Theory (Luke)

### 9.1 How can detections be made

When considering possible methods to be used in quantifying the presence of p-mode oscillations in the frequency-power spectrum, significant consideration was given to what parameters would be available to the PLATO mission data analysts. This was necessary since although any parameter from the models produced by the TSM group could be accessed, many of these would be found from further analysis of successful p-mode detections. Some examples of this are effective surface temperature,  $T_{\text{eff}}$ , mass,  $M$  and radius,  $R$ , which can be combined to get the average stellar density  $\rho$ . All of these, and a few additional physical parameters not listed above can be extracted from measurables such as the large frequency separation,  $\Delta\nu$ , the small frequency separation,  $\delta\nu_{02}$  and the frequency of maximum power oscillation,  $\nu_{\text{max}}$ . This meant that only directly observable parameters such as magnitude, luminosity and temperature could be used, which placed restrictions on the available detection techniques. However, additional techniques could be utilised that required the use of ‘forbidden’ parameters such as  $\nu_{\text{max}}$  to use as a check for whether the methods chosen and developed were producing realistic results.

Detections of p-mode oscillations of a star in a frequency-power spectrum can be made in several ways. The easiest and perhaps most obvious way of doing so is to look at the noisy spectra to simply see by-eye if the p-modes can be seen. Since this method is incredibly subjective, it may be influenced by prior knowledge of the rough or true location of the p-modes. One of the primary objectives was to develop a quantitative method of characterising whether or not a detection of the oscillations could be made from the noisy spectrum.

The method proposed was to use the python SciPy module `curve_fit` to fit curves to the overall Gaussian shape of the oscillation modes, with the theory being that due to the random nature of the noise, this should be the only Gaussian shape that the code could find. As such, providing that the code could make a fit, a positive detection would be confirmed. Testing of this showed that, although the overall method was sound, there were some flaws, meaning that several adaptations and iterations had to be made. These will be discussed in later sections.

Another method that can be used to make detections of p-mode oscillations from a frequency-power spectrum is the power-integration method. This detection method relies on there being an excess of power in the spectrum around where the oscillations are centred compared to that of the background. This is expected since there are specific components not related to random noise processes in this region that contribute to the power.

The frequency-power for any given star can be modelled as a combination of the following elements shown below in figure 9.1.

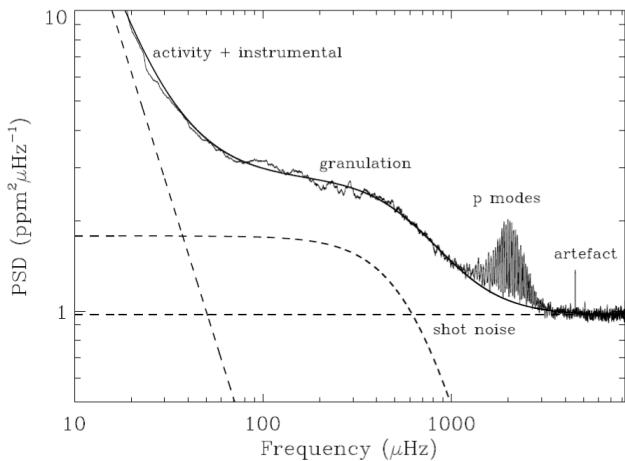


Figure 9.1: Components of a stellar frequency-power spectrum [Basu and Chaplin \(2017\)](#)

This can be simplified into two components;

a background component, encompassing stellar activity, instrumental effects, granulation and shot noise, and an oscillation component.

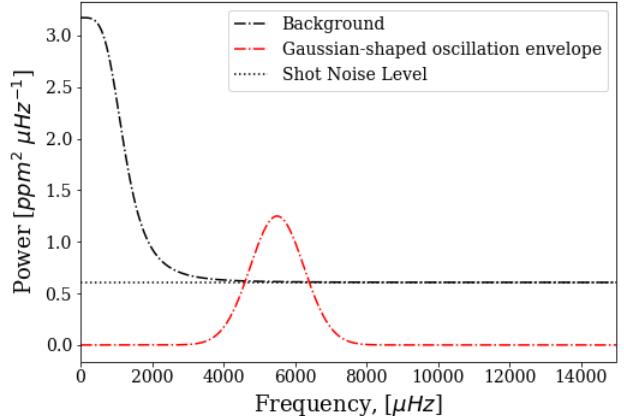


Figure 9.2: Model showing the two main components of a frequency-power spectrum and the underlying Gaussian shape the p-mode oscillations fit to.

The power-integration method then calculated the power in a given region under each of these curves and produces a ratio of these two values. Testing this on stellar data where oscillations are known to have been observed will give a different value for the ratio than that for a star where it can be seen in the model that the oscillations are obscured. As such, a cut-off threshold can be established for this method, above which a detection is made and below which it is said that the p-mode oscillations cannot be distinguished from the noise. Doing so requires careful fitting, as well as a choice of integration limits, which are discussed further in Section 13.

Both detection methods were possible in theory so their feasibility was explored on test data-sets to see if they worked in practice. It was anticipated that each method would have strengths and flaws and would be able to provide different parameters. These parameters could then be checked against those of the model provided by the TSM group to ascer-

tain if the detection methods were working as expected.

Since it was known that if a true detection or true non-detection was being made the methods should agree with one another, utilising both also doubled as a way to troubleshoot each of the separate methods in the event of a non-agreement. The exception to this would be in the event where the results of one or both methods were close to their respective cut-off thresholds, for which a decision would have to be made as to which method produced more reliable and accurate results.

A further motivation of using multiple methods to detect the presence of oscillations in the data was that each star could be given a statistical confidence level corresponding to the percentage above or below the threshold that the detection had been made at. Combining these from each method with an appropriate weighting factor would quantify the feasibility of making asteroseismic measurements in the model stars being observed, based on a set of arbitrarily chosen, but consistent limits.

## 9.2 Parseval's Theorem

Parseval's theorem is fundamentally a statement of energy conservation. In the field of asteroseismology, it means that the total amount of power contained in a spectrum must be conserved when a transformation is made between the time and frequency domains. Such a transformation is needed since the photometric data is taken in the time domain, yet the oscillations can only be observed in the frequency domain. In the time domain, the mean-squared power,  $\sigma^2$  of white noise, ie. the background, is given by the equation:

$$\sigma^2 = \frac{1}{N} \sum_{j=1}^N y_j^2 \quad (9.1)$$

where  $N$  is the number of discrete observations in the time domain and  $y_j$  is the power contained within each individual observation.

When transformed into the frequency domain, this mean-squared power is equally distributed between the frequency bins. To ensure that the signal is sampled in a way that means no information is lost and no spurious signal created, Nyquist sampling is used, discussed earlier in more depth in Section 5.10.

In the frequency domain, the average power per bin  $\langle P \rangle_{bin}$  is given by;

$$\langle P \rangle_{bin} = \frac{\sigma^2}{(bin)} = \frac{2\sigma^2}{N} \quad (9.2)$$

where  $\sigma^2$  is the mean-squared power of the white noise defined by equation (9.1). The bins in the frequency domain are separated by  $\Delta_T$ , the so-called frequency resolution, equivalent to the inverse of the total time of observations. As such, a more useful expression of power is per unit change in frequency  $\langle P \rangle_\nu$ , often given in units of  $\text{ppm}^2 \mu\text{Hz}^{-1}$ , as;

$$\langle P \rangle_\nu = \frac{\langle P \rangle}{\Delta_T} = \frac{2\sigma^2}{N\Delta_T} = \frac{2\sigma^2 T}{N} = 2\sigma^2 \Delta t \quad (9.3)$$

To ensure the power remained the same in both domains, a scaling factor,  $A_f$  was introduced to this expression for the power, such that  $\langle P \rangle_\nu = 2\sigma^2 \Delta t A_f$ . This scaling factor was calculated and tested on multiple KIC targets whose data was downloaded using the python package Lightkurve [Lightkurve Collaboration et al. \(2018\)](#) from the NASA Kepler database which reduced the difference in the total power between the two frames down to  $10^{-14}\%$ . This small difference was deemed to be insignificant, and likely due to the cropping of significant outliers from the time-domain spectrum.

### 9.3 Importance of preserving and extracting measurables from successful detections

When developing detection methods, it was essential to ensure that all measurables either remained unchanged or could be recovered from any given process during each point in the detection method. This was necessary as should p-mode oscillations be detected successfully in a given star, the useful physical information about the structure of the star would be encoded within these. As this information is what is desired from the photometric measurements made by PLATO, it is therefore essential that the characteristics of the individual oscillations and that of the larger p-mode oscillation structure are not lost during the data processing and manipulation stages used in making detections.

Measurables such as amplitude and location of oscillations could be found from one of the proposed methods, Gaussian Fitting to the overall shape of the p-mode oscillations, which would return one of the global parameters,  $\nu_{\max}$ , the frequency of maximum power in the oscillations. This would be found by the Gaussian fit since it would coincide with the centre of the fitted Gaussian. To extract a second of the global stellar parameters, the large frequency separation,  $\Delta\nu$ , a separate code would be developed, which would analyse the stars in which p-mode oscillations had been detected and find this parameter. The methods used to extract these parameters are discussed in more detail in Section 15.

## 10 Noise (Dan H)

### 10.1 Statistical Noise

The spectrum provided by the TSM group is the almost complete and accurate simulation of true data, however, something yet to be added to the spectra are the effects of random noise. When performing scientific observations, the physical processes being observed do not generally conform to precise patterns. The laws and theories physicists propose to model processes instead attempt to generalise these complex and chaotic patterns, resulting in computationally simpler equations. Therefore, to properly simulate realistic data that PLATO plans to observe, the DA group was required to add fake noise.

The power spectra provided by the TSM group was generated using a series of equations to model the surfaces and internal structures of stars. These equations were developed from the data recorded from long observations of the fluxes of various stars; for example the information on star surface conditions derived from [Brown et al. \(1991\)](#). None of the real stars observed in the physical data behaved in the same way, owing to random processes outside the star such as dust particles blocking photons or extra photons being detected from the cosmic background or distant galaxies when observing; therefore, the equations derived from these observations utilised the average results from all of the observed stars to understand how a idealised star would behave. Averaging over large numbers of observations allows for these random factors to be discounted. This means that theoretically, if infinite observations of the same star was made, then the noise would be able to be completely discounted. Hence, for the added statistical noise to be realistic, the noise adding process needed to give the true spectrum as an average result over infinite tests.

There have been several attempts at modelling accurate, noisy, solar-like spectra. One previous simulation by [Anderson et al. \(1990\)](#) uses the form,

$$O_i^{\text{sim}} = -y_i \ln R_i \quad (10.1)$$

where  $y_i$  is the TSM generated power values, and  $R_i$  is a random number,  $0 \leq R_i \leq 1$  generated using a uniform distribution. The negative  $\ln$  factor ensures that the input will always be multiplied by a number,  $0 \leq x \leq \infty$ . Negative  $\ln$  graphs always result in a positive number if the  $x$  value is between  $0 \leq R_i \leq 1$ , and owing to the exponential nature, any number below  $\frac{1}{e}$  will result in value  $>1$ . This is appropriate to model the effect of random noise on data, as random processes can result in increases or decreases in data observed, therefore, the range of factors produced by this data is appropriate.

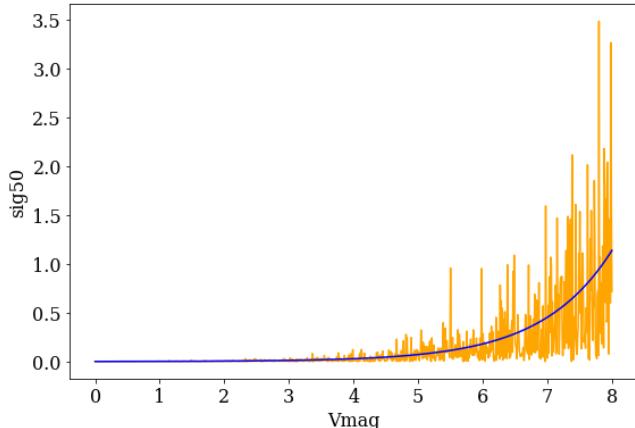


Figure 10.1: When the noise algorithm is applied to a simple exponential graph, the range of results also increases exponentially with the data.

A uniform distribution is used to generate random data to ensure there is no bias. Uniform distributions mean that every value within its range is equally likely to be selected, unlike other distributions (e.g Gaussian distributions) which have a varying probability. This is important in ensuring that the data still follows

the input spectrum. Equation 10.1 generates a random factor for each data point which means that the final result is affected by the initial input; larger values resulting in a much larger range of noisy data when multiplied together.

This can be seen in figure 10.1, with the noisy data range increasing with the input values. This increase in range means that despite the larger values of noise being applied to the data, the overall average is still equal to the input data as there is equal probability for values both above and below the data line.

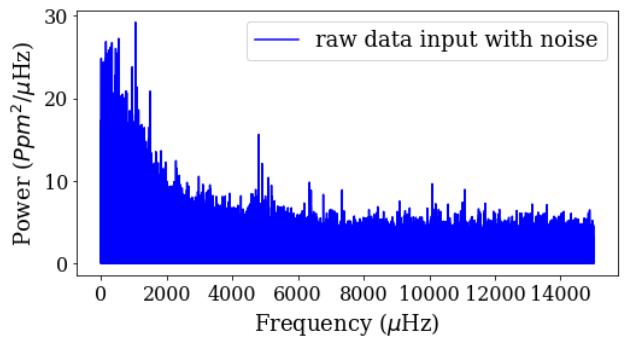


Figure 10.2: Noise algorithm applied to a single stars spectrum

Once applied to the overall stellar spectrum of the stars, (as seen in figure 10.2) a data set is achieved which appears to be affected by random processes. The noise addition process also affects the granulation curve visible at the low end of the spectrum, as its higher raw values can result in higher levels of noise. This process can hide the features of the spectrum, such as the p-modes of the star. This means that the investigation into the visibility of various star's p-modes is more accurate, as in physical recorded data, these features are not always obvious and visible.

## 10.2 Effects of Shot Noise

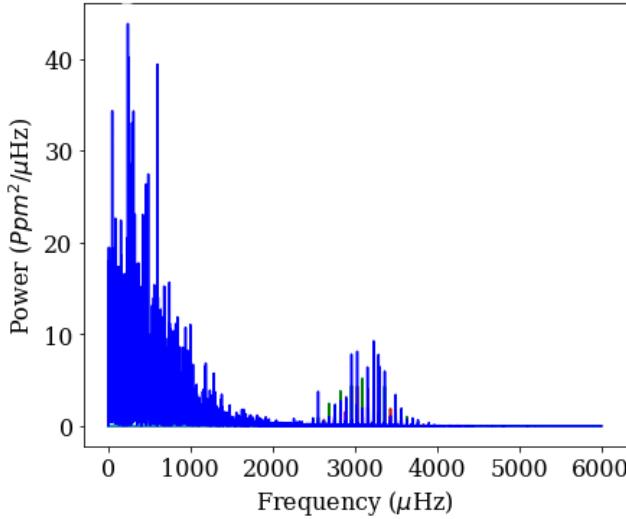


Figure 10.3: Random Noise algorithm applied to the full solar spectrum with no shot noise

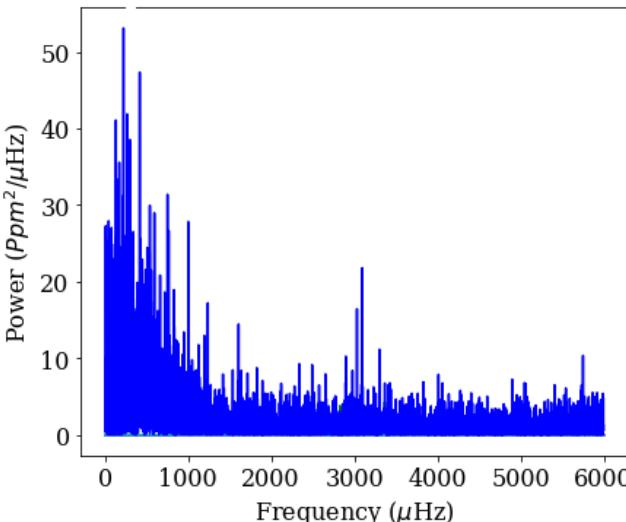


Figure 10.4: Random Noise algorithm applied to the full solar spectrum with a shot noise level from a  $m_v$  of 8

Statistical noise is not the only form of noise that the data from PLATO will include; shot noise must also be taken into consideration. As discussed previously in section 5.11, Shot noise

is an intrinsic property of stellar spectra, and must, therefore, will affect on the visibility of various stars.

When shot noise is applied to a stellar spectrum, it adds a base level of power, which is the minimum reading possible. This causes the spectrum to shift up and down depending on the intrinsic level of shot noise. If shot noise was not a property of stars, then the granulation peak and the separate Gaussian envelope of the p-modes would be easily distinguished. The granulation curve would flatten off before reaching this section, meaning the Gaussian envelope would appear against a base level of zero power surrounding it. As this is not the case, there are several effects that the shot noise has on the detection of the data.

This base level of power provided by the shot noise causes some problems when attempting to make a detection. The p-modes have a constant level of power which does not shift up or down with the shot noise level and, therefore, can become drowned out by the noise if the level becomes too high. As seen from figures 10.3 and 10.4, if the level of shot noise increases, then the Gaussian envelope becomes difficult to see by eye. However, analytical methods (discussed in sections 11, 12 and 13) allow the envelope to be distinguished from the surrounding noise when detection by eye is impossible.

As this shot noise level rises with increasing levels of  $m_v$ , it is possible to determine an approximate level at which the signal of the star's oscillation modes become completely drowned out by utilising the Gaussian fitting methods discussed in section 12. Figure 10.5 shows the effect of increasing the apparent magnitude on the difference in centres between the known centre of the Gaussian envelope and the centre determined via Gaussian fitting. As ex-

plained in section 12, this is a good measure of the level of visibility of a stellar object. Figure 10.5 shows a very clear spike around a level of  $m_v = 8$ . From this result, it is possible to see that there is a definite point at which a star is determined to be no longer visible and, therefore, unable to completely analyse the p-mode data of the spectra. This is only a test with the solar spectrum however, as other stars may have stronger or weaker p-modes within their spectrum which can move the limit at which the p-modes become invisible up or down accordingly, therefore, it is not possible to take a  $m_v = 8$  as a definite limit for all stars.

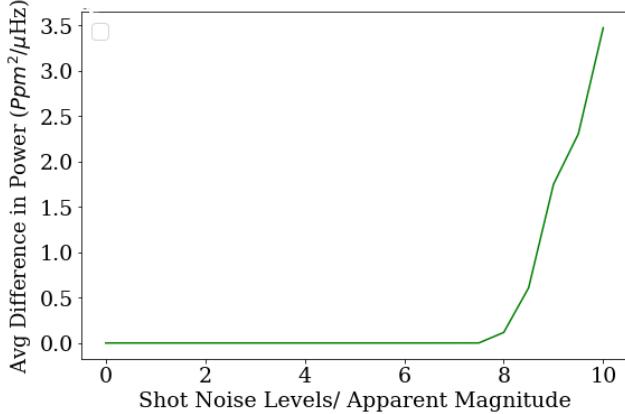


Figure 10.5: Increasing shot noise level caused by increasing the apparent magnitude used for the solar spectrum has a direct effect on the difference between the centres of the true Gaussian and fitted Gaussian as explained in section 12

## 11 Binning (Paige and Dan L)

### 11.1 Binning Basics (Dan L)

Binning is a method of processing data by smoothing a function. This is done by collecting a discrete number of data points into a series of groups - bins. Once the data is collected into these bins of a predetermined size,

with the group of data points as a whole in each bin is assigned a single value. This value is most commonly the median value of each bin. Binning in this way reduces the impact that the noise has on the signal as the median rarely takes an extreme value due to its independence of outliers, unlike other statistical averages such as the mean. This allows for a much smoother data set which increases the clarity of the p-modes within the spectrum which in turn makes them easier to detect.

While binning data has clear advantages, it is important to understand that over-binning can also present issues. If the bin size attached to the binning process is too large then the spectrum resulting from it will be too smooth. In this scenario, some of the intricacies of the data can be lost. In the context of this investigation, if the bin size is too large it could follow that the individual p-modes cannot be resolved and hence any observation that is attempted would result in a failed detection.

### 11.2 Binning Exponential Statistics (Dan L)

A Poisson point process, such as the arrival rate of photons, can be analysed using the theory of exponential statistics. In the context of this project, it can be said that the Lorentzian peaks are exponential. From this, a Gaussian shape can be created within the using binning. The exponential form of the data can be described by the chi-squared distribution with two degrees of freedom,  $2k$ . By increasing  $k$ , the data will take a form more akin to a Gaussian. An increase in binning, meaning that the number of data points within each bin increases, will lead to an increase in the binning factor, as per equation 11.1.

$$k = 2 \times \text{binning factor} \quad (11.1)$$

Where the binning factor is equal to the number of data points that are contained within each bin. For example, with 100 data points and 20 bins, the binning factor would be equal to 5.

The Erlang distribution, shown in equation 11.2, takes two input parameters, one being the number of degrees of freedom. The output of the Erlang distribution is a probability density function. An exponential distribution describes the time between two adjacent events of a random process with a constant mean, as per L (1970). Conversely, the Erlang distribution allows an insight into the time taken between any given event within the random process and the  $k^{\text{th}}$  following event. This removes the adjacency constraint in standard exponential distributions.

$$f(x; k, \lambda) = \frac{\lambda^k x^{k-1} e^{-\lambda x}}{(k-1)!} \quad (11.2)$$

where  $\mu = 1/\lambda$  is the scale factor,  $\lambda$  is called the rate parameter and  $k$  is the number of degrees of freedom. The scale factor is set at a value of 2. This simplifies the Erlang distribution to a chi-squared distribution that depends only on  $x$  and  $k$ . Increasing  $k$ , which is done by increasing the size of the bins, will lead to the initial exponential form of the data taking on the desired Gaussian shape. This is depicted graphically in figure 11.1.

### 11.3 Illustration of Binning (Paige)

The process of binning, as detailed above, was illustrated through multiple plots to demonstrate its effects and how it has aided in the identification of p-modes throughout this project.

To show how binning an exponential distribution leads to a Gaussian plot, a function was

coded using the Erlang distribution equation (equation 11.2).  $\mu$ , the scale factor, was set to equal two for the initial distribution to be simplified to one that is chi-squared with  $2k$  degrees of freedom, as described above. This is the exponential form,  $k=2$ . The function was then binned in a way which increased the bin size and therefore the degrees of freedom such that a more Gaussian-like distribution was produced.

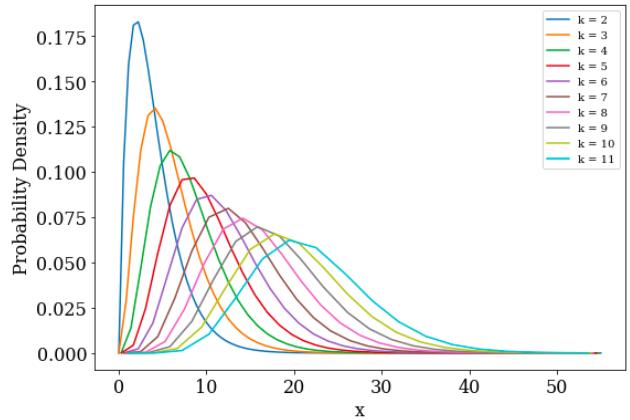


Figure 11.1: A finite illustration of the effects binning has on an initially exponential distribution.

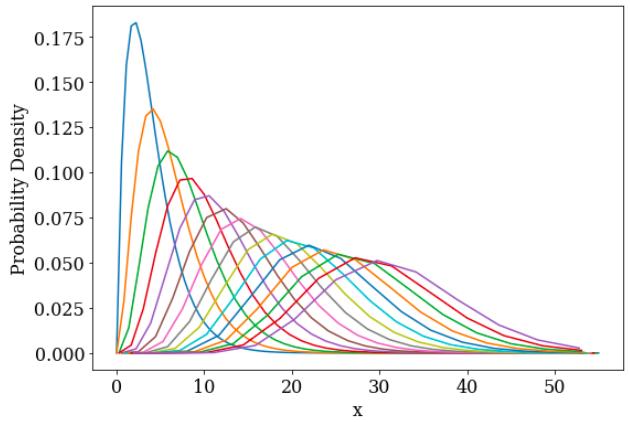


Figure 11.2: A more detailed illustration, which contains more iterations, of the binning effects, allowing the plots to continue and show a clear Gaussian distribution.

Figure 11.1 shows the exponential plot,  $k=2$ ,

becoming more Gaussian as binning takes place. This contains a finite number of plots so that each is distinctly coloured and can be identified. Using a 'for' loop, the bin sizes varied from 18 to 100, such that, with 100 data points the binning factor ranged between 1.00 and 5.56 and in turn the degrees of freedom increased from 2 to 11.

Figure 11.2, no longer has distinct colours but displays more plots, up to  $k = 16$ . This further shows how the exponential function evolves into a Gaussian.

An example of this binning method working can be seen in figures 11.3 and 11.4.

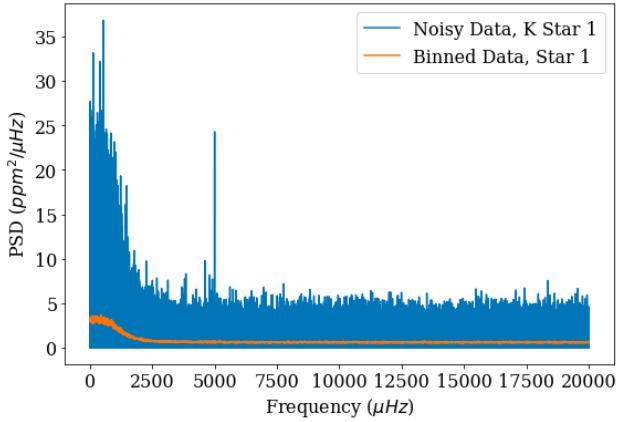


Figure 11.3: Graph to illustrate how binning the data results in a smoother plot.

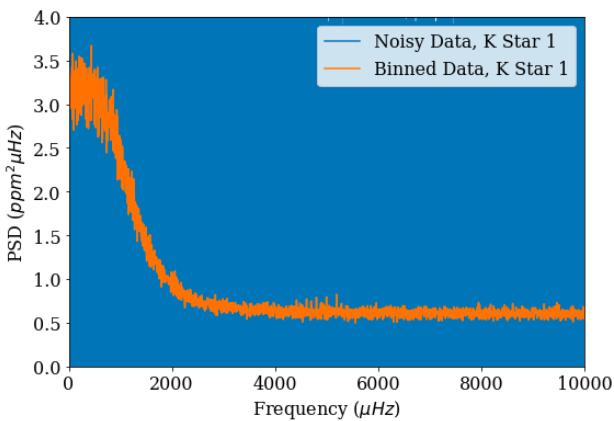


Figure 11.4: Cropped version of figure 11.3 .

It can be seen that binning the data smooths and reduces the noise's effect as described and makes it much easier to pinpoint the location of the p-modes.

## 12 Gaussian Fitting (Edward and Dan H)

### 12.1 Curve Fitting Processes (Dan H)

Gaussian fitting is a powerful analytical tool which can be used to determine the features of the stellar plots generated by the TSM group. It allows for the detection of the p-mode envelope, even though it may be indeterminate to the human eye. In combination with the binning algorithms discussed in section 11, Gaussian curve fitting can be used to determine a pass/fail criteria for star p-mode detection which will exact limits on the PLATO telescope to be found.

Gaussian fitting relies on a simple Python function known as `curve_fit`. This function inputs a set of  $x$  and  $y$ -axis data and any mathematical function. The `curve_fit` function then attempts to find a set of input values for the mathematical function which output a set of values as close to the given  $y$ -axis data as possible. By re-inputting the generated parameters into the mathematical function, a new curve can be generated which fits the data based on the mathematical function.

For the fitting, the Gaussian function was utilised.

$$y = H \exp\left\{-\frac{(x - x_0)^2}{2W^2}\right\} \quad (12.1)$$

where  $x_0$  is the centre,  $H$  is the height and  $W$  is the full-width half maximum of the Gaussian.

It is known from the TSM group that the p-mode envelope is generated via a similar Gaussian function, and therefore, an exact Gaussian should be easy to fit to it in the final spectrum. As the TSM group provided the features of the Gaussian envelope within the data, the results generated can be compared with the true values. The closer the values of the fitted Gaussian, then the more likely a detection of the star. However, not all the values produced in the curve fit are as useful as one another when noise is factored in.

The most important of the generated values is the  $x_0$  or centre of the Gaussian. If there was no noise present, then fitting the Gaussian would be a simple process and would output the exact curve, however, shot noise and granulation present a problem. These create a background level of noise which means that the data becomes obscured (as seen in section 10). When fitting the Gaussian with noise added, the values of the curve fit can be shifted away from the true values. The height of the Gaussian may be reduced or increased as the statistical noise can increase or decrease the values of data. The full-width half max will be increased as the high level of shot noise causes the curve to appear flattened against it. However, these noise levels affect the final measurement, the centre of the Gaussian, to a much lesser degree. Noise can only affect the y-value magnitude, but not its position on the x-axis; meaning this value should be preserved. Therefore, if this value is incorrect when fitting the Gaussian, then it was not affected by the addition of noise, but rather the poor fitting of the curve.

This provided with the first selection criteria, the difference in the centres of the Gaussian. By comparing the fitted value to the known true value, a difference between them was obtained. The larger this difference, the poorer the Gaussian fit and the less visible the star.

However, when attempting to fit a Gaussian to the data, it cannot automatically find the exact position of the envelope. With the addition of the granulation curve, the shape of the data cannot be modelled by a single Gaussian curve. To overcome this problem, a series of solutions was developed.

## 12.2 Gaussian Fitting Methods

### 12.2.1 Multiple Generic Gaussian Fitting (Dan H)

To develop these fitting methods, the DA group used a full spectrum of the Sun, which could be altered to appear at various distances (visual magnitudes) away from the Earth. Having control over the distance allowed level of shot noise affecting the plot to be changed, which allowed for the success of the method with stars of different brightness to be determined.

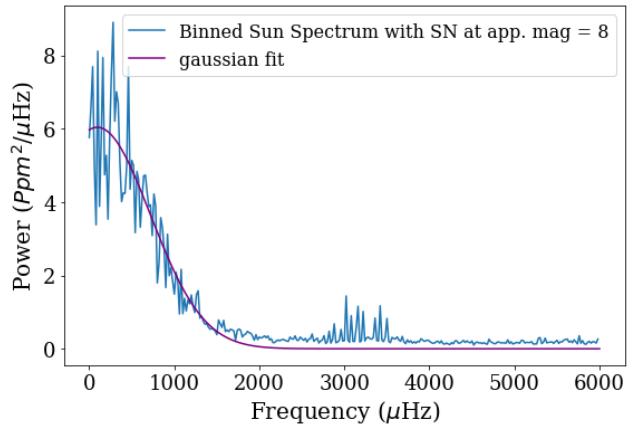


Figure 12.1: Attempt at fitting a Gaussian plot to the solar spectrum

The first attempt at Gaussian fitting of a full spectrum utilised a process of multiple generic Gaussian fitting. When testing the data with the solar spectrum, attempting to fit a Gaussian over the complete data set (once passed through the binning process), results in the curve\_fit fitting a Gaussian to the granulation curve instead of the true Gaussian. This occurs because the Gaussian envelope has a much

smaller magnitude than the granulation curve and, therefore, is ignored. As you can see in figure 12.1, the purple Gaussian fit has ignored the p-mode envelope entirely. This meant that the method required to determine the shape of the Gaussian envelope would need to be adjusted from a simple curve fit.

The first way to improve this fit is to alter the Gaussian function used to plot it. The function shown in equation 12.1 assumes that the function flattens off to zero at  $-\infty$  and  $\infty$ . In reality, the shot noise level creates a raised zero level which is not at  $y = 0$ . Therefore, by adding a  $y_0$  value to the equation as following, the Gaussian plot can be shifted up or down to account for this base level of noise.

$$y = H \exp\left\{-\frac{(x - x_0)^2}{2W^2}\right\} + y_0. \quad (12.2)$$

The second attempt to improve the fit was through applying a second Gaussian on top of the original Gaussian. By taking equation 12.2 and adding it to itself with a second set of variables, an equation is obtained for a combination of two separate Gaussian fits with double the parameters. As the full stellar spectrum is equivalent to the granulation curve with the added p-mode envelope, by adding together two separate Gaussian plots, it would be modelled properly. If this process had succeeded, the curve\_fit process would fit the data to both of the predicted Gaussian plots simultaneously. However, in practice this failed, as the program always placed the second Gaussian on the peak of the first, creating a single large plot once again and missing the p-mode envelope. This method was abandoned but the improved Gaussian formula was to be continued to be used. This then led to the second fitting technique.

### 12.2.2 Single Generic Gaussian Fitting with Cut-off (Dan H)

For the curve\_fit process to detect the true Gaussian envelope, it has to ignore the granu-

lation curve on the lower frequency end of the plot. One method of achieving this is through the use of a cutoff frequency. This would simply cut the spectrum at a point past the granulation curve and then proceed with the Gaussian fit with the remaining data. As the granulation curve is no longer included it cannot be accidentally fitted to and the fitting program will therefore only find the true Gaussian envelope.

To find this cutoff frequency a single Gaussian fit was initially made to the full spectrum, as before. This will give a fit to the granulation curve, which closely follows its shape. The shape of solar spectra means that the p-mode envelope is usually found after the granulation curve flattens off, or very near to the bottom of its curve. Applying the same logic to the newly produced curve fit, it can be seen that the desired section from the data is the bottom of the first Gaussian and the flat areas at higher frequencies. To determine where this point is two other features of the generated Gaussian were utilised; the height, and the  $y_0$  value.

$$Y_{\text{cutoff}} = C(H - y_0) + y_0. \quad (12.3)$$

Using equation 12.3 the  $y$ -value required can be found. By subtracting the  $y_0$  value from the height of the Gaussian,  $H$ , the overall height of the Gaussian above the noise level is obtained. Next, this value is multiplied by a cutoff factor,  $C$ , to reduce it to the value above the noise level required, before adding the  $y_0$  value to this new value. A script then checks the  $y$ -values of the curve to find the  $x$ -value closest to this point, before removing all the data prior to this point. After some testing, a value of approximately 0.1 is appropriate for the cutoff factor as this removes most of the granulation curve, without removing any of the p-modes on nearer stars. Curve fitting can now be applied normally to

the cut data, as there is only a single curve to detect; the Gaussian of the p-mode envelope.

Figure 12.2 shows an example of this process on the original solar spectrum provided by TSM. The cut and binned plot now only contains frequency values above 2000, which results in a spectrum without the granulation curve. When a subsequent Gaussian plot is applied to the cut data it fits the true Gaussian envelope (shown in purple) with only a small error between their centre values (in dotted lines). From this, it can be seen that this method is much more powerful than the double Gaussian as it obtains an exact value for the accuracy of the fit, and therefore, the likelihood of a detection.

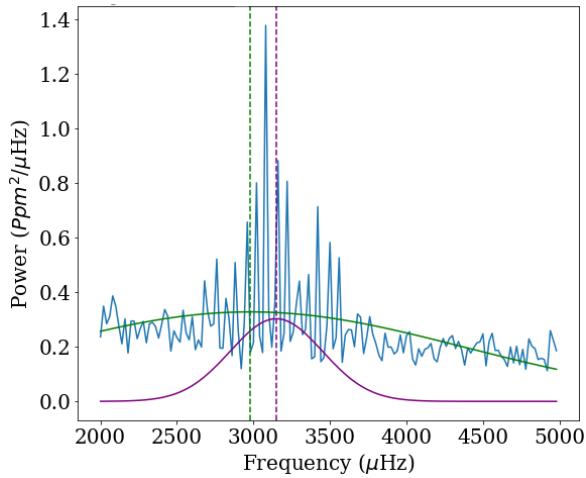


Figure 12.2: Cut Solar spectrum, with original Gaussian and centre points plotted

By repeating the statistical noise addition process around 50 times for each star, an average difference for 50 trials was obtained. This average became one of the criteria for determining if a star was visible. However, this process still had flaws.

Firstly, If the initial Gaussian plot was inaccurate and created a very high power Gaussian at a very negative number, then the cutting algorithm had a chance to not cut the data. The max height of this Gaussian would be larger

than 10 times the height of the data and would therefore, simply return the same Gaussian twice. To address this problem a background removal process was applied to the data. This process will be explained more completely later on in the section, but the process simply involves the subtraction of the original Gaussian fit from the data. This flattens out the data and allows much easier detection of the Gaussian envelope. Using this improved process it became possible to detect the centres of many of the stars provided in the first data sets by the TSM group.

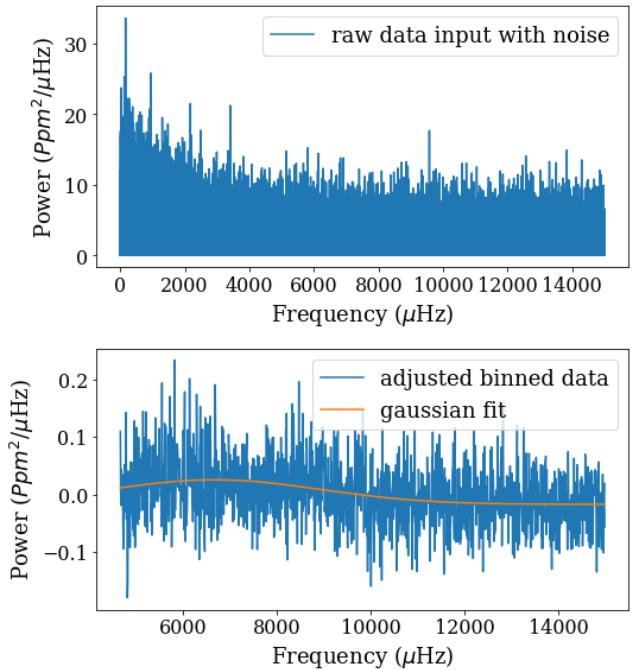


Figure 12.3: Plots for the raw star data, and the cut and binned data with a plotted Gaussian fit

The second major problem with this process however, is less easily solved. As both fits are Gaussian, the initial fit does not completely conform to the shape of the granulation curve. The granulation curve has an approximately Gaussian shape, but it is not exact. To obtain better results, another method of detection was created.

## 12.3 Background fitting and Gaussian fitting (Edward)

The second method that was used to detect the envelopes, was to fit a line to the granulation then subtract that line from the data and fit a Gaussian to the flattened data. As covered previously in the report, every star has a unique background so each must be fitted independently. The `curve_fit` program in python was used to fit a line to the background. The equation for the power of the background is ([Lund et al. \(2017\)](#)),

$$\text{power} = \frac{A}{1 + (2\pi\nu\tau)^2} \quad (12.4)$$

where  $A$  and  $\tau$  are constants. It was found that restricting  $A$  to be positive improved the fitting of the background.

[Equation 12.4](#) was used to fit the background as it has few values that the `curve fit` program would have to change, unlike other equations that can fit the background. This is important as the `curve fit` program reduces in accuracy as you increase the number of values it fits.

A modified version of [Equation 12.4](#) was used which accounted for the shot noise of the star,

$$\text{power} = \frac{A}{1 + (2\pi\nu\tau)^2} + \text{shot noise.} \quad (12.5)$$

Shot noise was calculated for each star independently using the  $m_v$  of the star, which can be measured by PLATO and so can be used for analysis. [Equation 12.6](#) is an altered form of [Equation 5.48](#) found earlier in the report,

$$\text{shot noise} = 18 \times 10^{(-0.4 \times (11.0 - m_v))}. \quad (12.6)$$

Initially, when analysing spectra from the Sun the shot noise component of the background equation was not needed for the fitting. As the  $m_v$  of the Sun is so low that the power approaches zero at high frequencies. However,

for K dwarfs,  $m_v$  is much higher so must be included.

Once a line has been fitted for the background it is subtracted from the spectrum. Then a Gaussian is fit to the flattened data. One other option would be to fit the Gaussian to data including the fitted background inside the Gaussian fit. This method was tried initially and found to be less accurate than the subtraction method.

A Gaussian was then fitted to the flattened data. The Gaussian in the form,

$$y(x) = Ae^{\frac{-(x-B)^2}{2C^2}} \quad (12.7)$$

where  $A$ ,  $B$  and  $C$  are the height, centre and width of the Gaussian respectively. The `curve_fit` program could alter the values of  $A$ ,  $B$  and  $C$  to find the best fit for the data. Certain restrictions were placed on the range the program could change  $A$ ,  $B$  and  $C$ . The height,  $A$ , was prevented from being negative as the envelope should be positive. The centre,  $B$ , was limited to be between the minimum and maximum value of the frequency of the data provided, as it was known the centre of the envelope must be contained within the range of the data.

Another possible method which could be used would be to fit both the background and the Gaussian simultaneously. This would have the advantage of being quicker as only one fit would have to be done per star. However, a disadvantage of this method is the large number of values the `curve_fit` program has to fit, this will lead to a reduction in the accuracy of the fit. An advantage of the method where the background and the Gaussian are fit individually is that errors in the fitting can be identified more easily. This was useful when ensuring the fits were completely accurate.

## 12.4 Finding $\nu_{\max}$ (Edward)

The  $\nu_{\max}$  of a given star was calculated multiple times, adding the noise separately each time and then averaging over all calculated  $\nu_{\max}$  values. The Gaussian fitting appeared to produce a successful detection for every star, even those with a  $m_v$  large enough that no clear detection could be made by eye. This was occurring because the curve\_fit program would fit a Gaussian to the noise if no envelope was present to fit to. It was also found that when the curve fit program could not find the envelope, the final position of the centre of the Gaussian was very close to the initial value of the centre. If an envelope was visible, a range of initial values for the centre of the Gaussian centred around the true value of the centre of the Gaussian would each return an answer approximately equal to the true centre of the Gaussian. This observation was used to find detections that could not be made by eye. Using this method to categorise detections it was learnt that the ability to detect a star depends both on the  $m_v$  and the  $\nu_{\max}$  of a star not just the  $m_v$  which was previously assumed.

Figure 12.4 shows a positive detection. This can be seen as there is a plateau from 3500  $\mu$ Hz to 5500  $\mu$ Hz. The plateau shows the range of initial guesses for the centre of the Gaussian which converge on an approximately equal value for the fitted centre.

Figure 12.5 shows no detection, this can be determined as there is no plateau and the data more accurately fits the  $y = x$  line shown on the figure. A possible detection is the middle point between figure 12.4 and 12.5 where it can not be decided which category the graph belongs in.

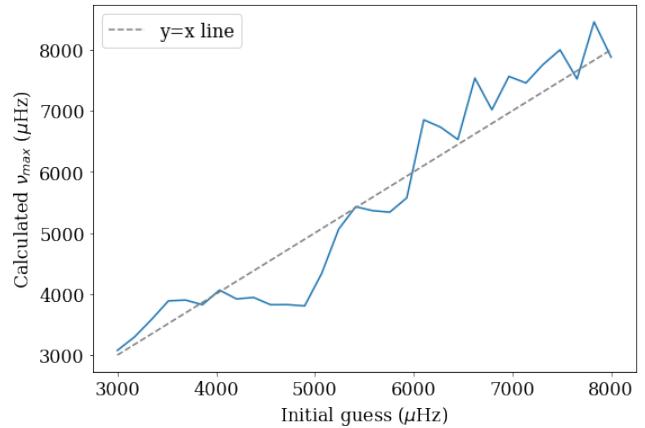


Figure 12.4: Graph showing how the fitted value for the centre of the Gaussian changes with an initial guess for the centre of the Gaussian, when there is a detection

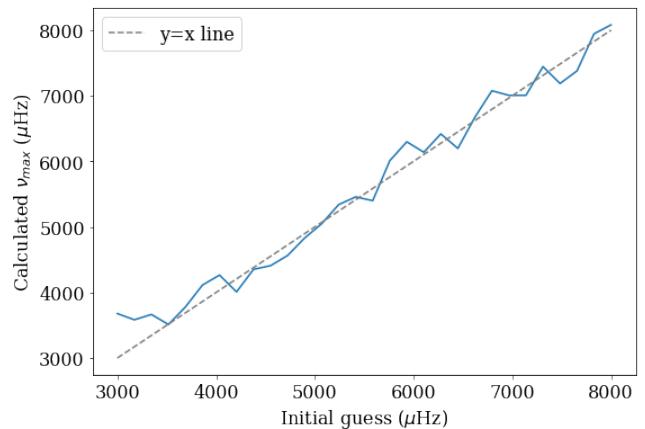


Figure 12.5: Graph showing how the fitted value for the centre of the Gaussian changes with the initial guess for the centre of the Gaussian, When there is no detection

Initially this method of detection was used for 40 K dwarfs provided by TSM. Each star was fit using 30 initial conditions for the centre of the Gaussian with 50 trials from each initial condition for the centre. The graphs which were produced by this process were then sorted manually into three categories: definite detection, possible detection and no detection. The graphs were sorted by 6 people separately to ensure that human bias was removed. The cat-

egory with the most votes was then selected for each star. If there was a tie maybe was selected to ensure no false positives.

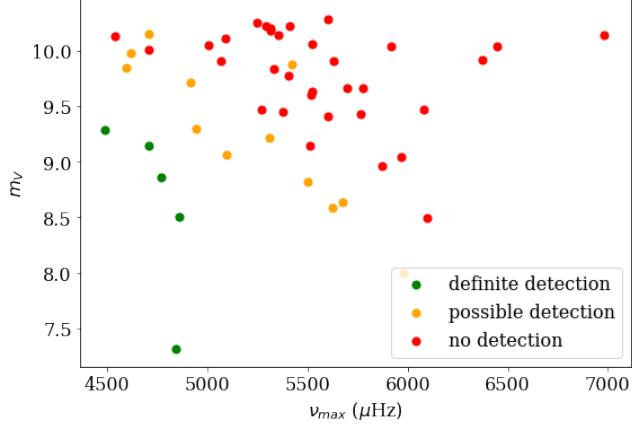


Figure 12.6: Graph showing K dwarfs provided by TSM plotted on a graph of  $\nu_{\max}$  against  $m_v$  where the colour indicates if a detection could be made

Figure 12.6 shows that the ability to detect a star's oscillations depends both on the  $m_v$  and the  $\nu_{\max}$ . Unfortunately, this data is inaccurate as there was a fault in the generation of the initial data this leads to the number of detections being lower than depicted in Figure 12.6.

The initial data provided by the TSM was generated using stars from a simulated galaxy, however, these stars left areas on  $m_v$  and  $\nu_{\max}$  graph unfilled as can be seen in Figure 12.6. This meant there was no data for stars in the unfilled regions, meaning it could not determine if there would be a detection for a star in those regions.

To solve this problem The data provided by TSM was altered in a grid shown below in Figure 12.7.

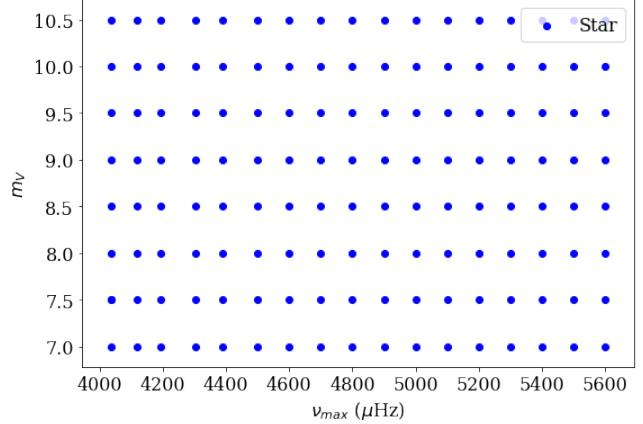


Figure 12.7: Graph showing a grid of stars for analysis

Seventeen stars with  $\nu_{\max}$ 's approximately 100  $\mu\text{Hz}$  apart from each other with the lowest being 4000  $\mu\text{Hz}$  and the highest being 5300  $\mu\text{Hz}$  were selected from the initial data set. Then the visual magnitudes of each of these 17 stars were altered to set values ranging from 7 to 10.5 with a step size of 0.5. This produced 136 stars in a grid shown in Figure 12.7.

These stars were then run through the second method detection described previously. Then sorted manually by five people into the categories definite detection, possible detection and no detection. The most popular answer for each star was selected as before and the data was then plotted in a colour grid with the K dwarfs overlaid.

As can be seen in Figure 12.8 there is one star in the definite detection area. The data forms the shape which was expected with low  $\nu_{\max}$ s and  $m_v$ s giving detections and larger values giving no detection. To provide a more detailed result the grid was remade with more stars, the range of the  $\nu_{\max}$  and  $m_v$  values of the stars were also changed to be more focused on the area of change from definite detection to no detection, to avoid large areas of no detection.

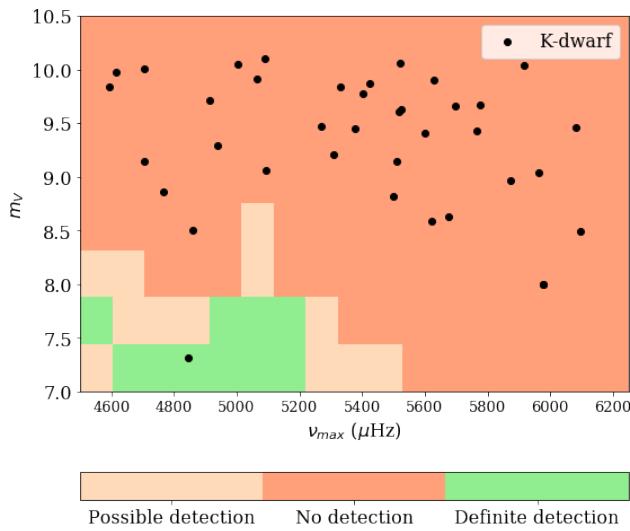


Figure 12.8: Background colour represents whether the star corresponding star to that location could be detected, the black dots represent K dwarfs

This lead to the creation of a new set of grid data. 14 stars with  $\nu_{\text{max}}$ 's approximately 100  $\mu\text{Hz}$  apart from each other with the lowest being 4000  $\mu\text{Hz}$  and the highest being 5300  $\mu\text{Hz}$ . The  $m_v$ 's of each of these 14 stars were altered to set values ranging from 6 to 10.5 with a step size of 0.25. This produced 266 stars in a grid shown below in figure 12.9.

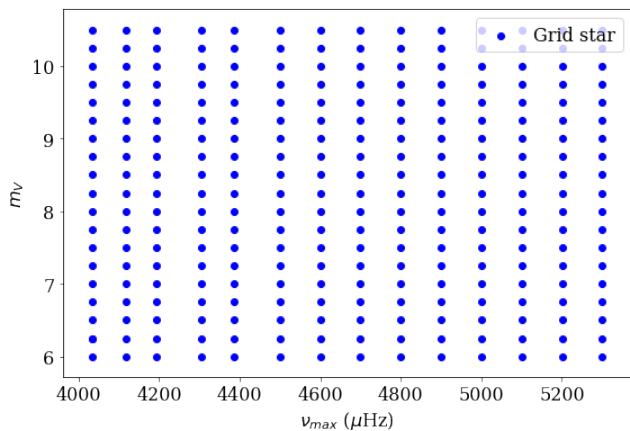


Figure 12.9: Stars generated in a grid shape by TSM for analysis

The 266 grid stars were then run through the code with 30 initial guesses for the centre of the

Gaussian and 25 trials for each initial guess. The graphs which this process produced were then categorised by three people separately. The data that can be seen below in Figure 12.10 shows the results of the categorisation of the stars as a colour grid. The black dots overlaid are the K dwarfs provided by TSM, the M dwarfs provided by TSM can not be seen on the graph as their  $\nu_{\text{max}}$ 's and  $m_v$ 's are too large to be displayed on the graph. But it can be assumed that they would not be able to be detected.

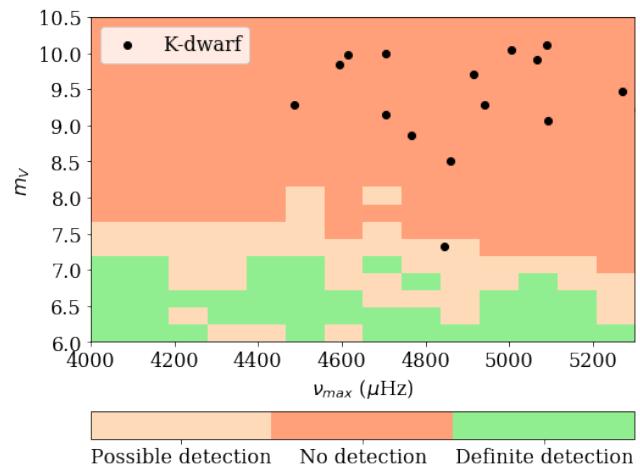


Figure 12.10: Background colour represents whether the star corresponding star to that location could be detected, the black dots represent K dwarfs

As can be seen in figure 12.10 there are no stars in the definite detection area, only one star is in the possible detection area and the rest are in the no detection area.

Upon further inspection into the generation of the grid data it was decided that the first five columns of stars in the grid with  $\nu_{\text{max}}$ 's of 4000  $\mu\text{Hz}$ , 4100  $\mu\text{Hz}$ , 4200  $\mu\text{Hz}$ , 4300  $\mu\text{Hz}$ , 4400  $\mu\text{Hz}$  respectively had original visual magnitudes which were too high before being altered making them unusable in the grid. These K dwarfs were anomalies so analysing their oscillations could not be extrapolated to other K

dwarfs of different visual magnitudes. These columns were removed, Figure 12.11 shows the final data with the unusable stars removed.

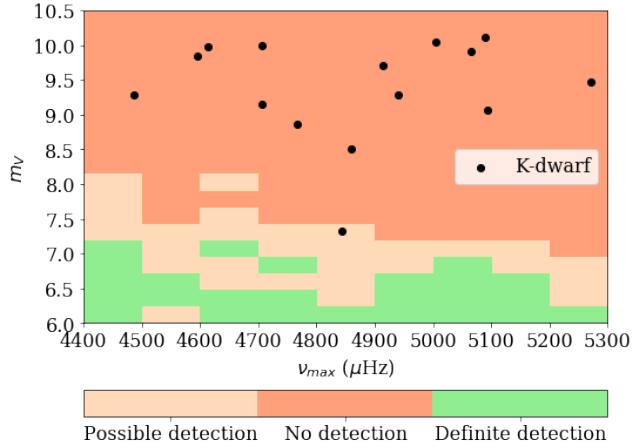


Figure 12.11: Background colour represents whether the corresponding star to that location could be detected, the black dots represent K dwarfs. the graph is a cropped version of Figure 12.10

This cropped version of the data does not change the number of stars in the definite detection and possible detection areas. However, it does reduce the area in which it can be said that there could be a detection.

This method does have potential downsides as the area around the grid stars is assumed to act the same as the star at its centre in regards to detecting oscillations. This could be counteracted by a more detailed grid around areas of interest. Another issue with this method is the colour grid assumes the stars will be at the centre of the rectangles, however, they are not exactly. The stars selected were the ones closest to the value of  $\nu_{\text{max}}$  which was needed. As the  $\nu_{\text{max}}$  of the stars could not be changed this lead to the initial grid not being perfectly spaced.

One additional possibility to improve the method would be to write code which will categorise the graphs plotting the initial centre of the Gaussian against the fitted centre of the Gaussian, shown in Figure 12.4 and 12.5. Into

the categories definite detection, possible detection and no detection. This code could work by analysing the  $y$ -values of the graph and evaluating whether the  $y$ -values are contained within a given range of  $y$  for a given length of the  $x$ -axis. The constraints would be provided by a person initially and then the code could analyse all the stars with greater consistency than a human could. This would also allow a larger number of stars to be categorised which would lead to greater accuracy when determining the cut off for detection.

## 12.5 Temperature method (Edward)

One technique which could be used to boost detections is averaging the spectra of stars with similar temperatures. This can be done as stars with similar temperatures have approximately identical spectra. Doing this leads to a reduction in the noise of the data which can then be further reduced by binning the data. This method does lead to the inability to make judgements of specific stars only certain temperature ranges of stars. However, it does provide results for groups of stars with similar detections.

# 13 Power Integration Fitting (Paige and Dan L)

## 13.1 Fitting to the Background and P-modes (Dan L)

The Gaussian fitting process described in section 12, while successful for stars with p-modes well within the detection thresholds of the model, can encounter difficulties when fitting to stars with a less accessible p-mode envelope. The low signal to noise ratio that corresponds to a star with a greater apparent magnitude and hence a lower brightness, such as K and

M class dwarfs, leads to poor resolution of p-modes over noisy background when plotted on a frequency-power spectrum. The function which attempts to automatically fit a Gaussian curve to the data was found to occasionally do so in an incorrect position when the resolution is low. The Gaussian fit which arises from this imperfect process could lead to inaccurate solutions for the centres, amplitudes and standard deviations of the p-mode envelope. As a result, these fits should be discarded to prevent them from skewing the data with inaccuracies.

The frequency-power spectrum generated by each star can be characterised by the power of the background compared to the power of the Gaussian envelope containing the p-modes. Stars of a higher apparent magnitude, and hence those which provide a lower flux, show a higher level of shot noise. Following from this, they also exhibit less visible p-modes. To quantify which detections are real and which indicate an incorrect fit, the power contained by the background signal and the p-modes must be compared. The power of these two functions is measured by integrating the mean background level and the Gaussian envelope as per equations 13.1 and 13.2:

$$P_{\text{Background}} = \int_{\nu_{\max} - \frac{FWHM}{2}}^{\nu_{\max} + \frac{FWHM}{2}} \frac{a}{(1+(2\pi\nu b)^4)} d\nu \quad (13.1)$$

$$P_{\text{Gaussian}} = \int_{\nu_{\max} - \frac{FWHM}{2}}^{\nu_{\max} + \frac{FWHM}{2}} A e^{-\frac{(x-\nu_{\max})^2}{2\sigma^2}} d\nu \quad (13.2)$$

where the integers  $a$  and  $b$  are previously determined by the 'scipy.optimize.curve\_fit' function. The Gaussian is characterised by its amplitude,  $A$ , its centre,  $\nu_{\max}$ , and its standard deviation,  $\sigma$ . Each of these parameters is determined by the function that fits the Gaussian envelope. Given that the Gaussian envelope gives a stronger signal relative to the background for brighter stars, it follows that the

ratio of power under the Gaussian to the power under the background decreases as the detection limit is approached. This means that as p-mode detections become increasingly difficult the power ratio should decrease too. It was concluded that, if some minimum acceptable value for this ratio could be specified, a cut-off ratio, any individual fit with a ratio below this value should be further investigated. These data points would be assumed to have an improper fit and, therefore, would be marked as a fail or simply undetectable and should be discarded.

For a reliable determination of what this threshold value should be, it was essential that any process involved could be carried out without using any results acquired from the automatic Gaussian fitting code. A false positive reading from the Gaussian fit would pass incorrect parameters to the subsequent ratio calculation and any results acquired would also be incorrect.

## 13.2 Initial Investigations of Power Ratios (Dan L)

Initial calculations of the power contained within the Gaussian envelope and the power contained within the background were carried out on the early solar test data provided by the TSM team. Information regarding key values such as  $\nu_{\max}$ , the height of the Gaussian envelope at  $\nu_{\max}$ , the background level as well as the FWHM of the Gaussian envelope were already available due to extraction by the data analysis team. These parameters could be used to construct the graph shown in figure 13.1 which highlights how the p-modes stand above the background. The Gaussian envelope that contains these p-modes, the background level and combined background and p-mode signal can be seen in figure 13.2.

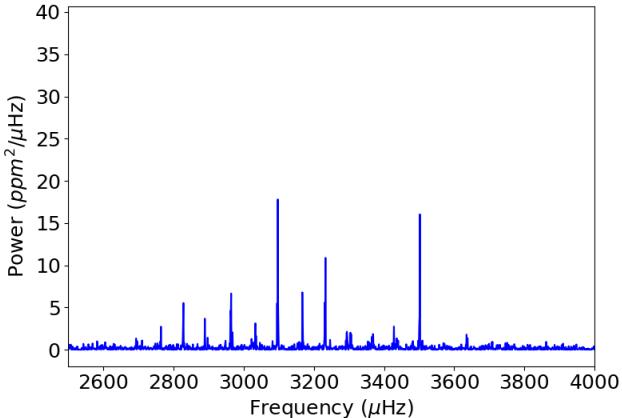


Figure 13.1: A zoomed in view of the p-modes for solar test data

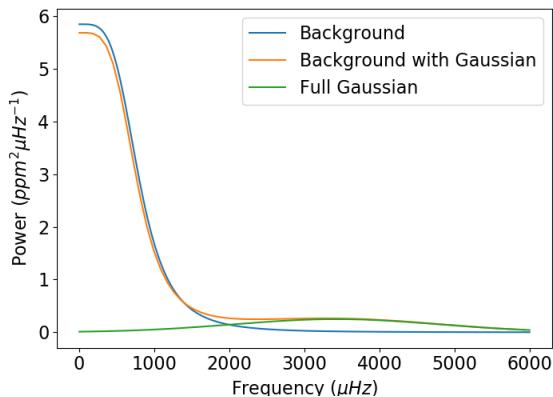


Figure 13.2: A graph showing how the background and Gaussian envelope combine to create the measured frequency-power spectrum

Each line within figure 13.2 is represented by an equation written within the written code, subsequently, they're plotted on the same pair of axes. These equations could then be inserted into the ‘scipy.integrate’ function to compute the power contained within both the Gaussian and the background. Following from this, the power ratio assigned to this plot could be found by computing the ratio of power under the Gaussian to the power under the background. It was decided, through analysis of figure 13.1, that the p-mode information was largely contained up to a frequency of roughly 500 $\mu$ Hz.

To ensure that no information was lost the initial integral limits were set at 0 to 6000. The ratio is then defined by,

$$X = \frac{\int_0^{6000} Ae^{\frac{-(x-b)^2}{2c^2}}}{\int_0^{6000} \frac{a}{(1+(2\pi xb))^4}} \quad (13.3)$$

where X is the power ratio, A is the Gaussian amplitude,  $\sigma$  is the FWHM of the Gaussian and a and b are constants determined by the granulation fitting code. Equation 13.3 yielded the first power ratio result of  $0.153 \pm 4.233$ .

This process was an important first step towards implementing a power ratio threshold that would discount false positive detections. While the method described above provides valid results, it was an initial test which gave an estimate for the ratio magnitudes that were expected, as well as providing an introduction to the required process. However, the coding process was reliant on the  $\nu_{max}$  of the Sun and other parameters defined above. While these values were accessible in the case of Sun data, they would not be readily available for stars within the simulated galaxy created by TSM. It follows that a second method which could provide a calculation of the power ratio without these parameters was required. This would result in the ability to reliably compute a power detection threshold. One proposed method of achieving this was through a manual fit. If a Gaussian could be plotted in the right place without requiring any information from the code involving the automatic Gaussian fitting process, a threshold ratio for successful detections could be reliably sourced. A function that allowed variation of its amplitude,  $\nu_{max}$  and FWHM was used to achieve this.

### 13.3 Optimising the fit using Sliders (Paige)

To optimise the p-mode fitting and make it independent of information provided by the TSM

team, a slider function was defined. This function used equation 13.2 for a Gaussian curve and allowed three parameters to be varied: amplitude,  $A$ , centre,  $\nu_{max}$ , and standard deviation,  $\sigma$ . It was thought that manually fitting the Gaussian to detectable stars and comparing this to others that are much dimmer and ambiguous would result in the determination of a cut-off ratio. It was then expected that this value could be incorporated into the automatic Gaussian fitting code so that any undetectable or poorly fitted stars could be discarded. This method of manual fitting was also considered beneficial as it could be used to provide an average of  $\nu_{max}$ . This  $\nu_{max}$  could then be compared to that outputted by the automatic code and also the true  $\nu_{max}$  from TSM to further determine whether the fit was correct and whether a detection had been made.

The slider function was first developed using the initial frequency-power spectrum of the Sun. As seen below, in figures 13.3 and 13.4, the slider function was plotted along with a binned frequency-power spectrum and fitted background.

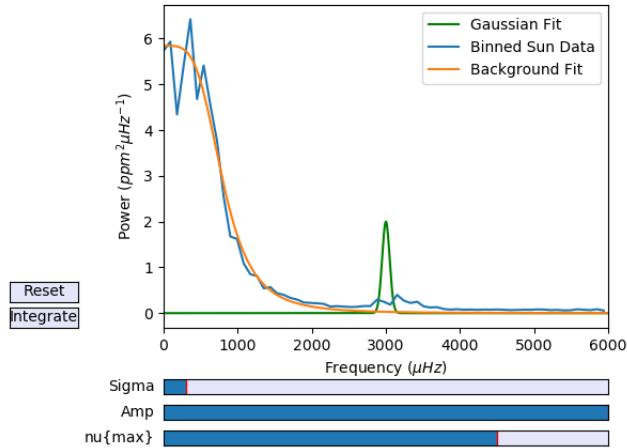


Figure 13.3: The initial slider function which allows manual fitting of a Gaussian to a binned frequency-power spectrum of the Sun. In this plot the Gaussian is in its initial position.

To improve the accuracy of p-mode detection, this manual method allowed for a precise, by-eye Gaussian fitting. Each plot consisted of the three sliders and two buttons: Reset and Integrate. The reset button was implemented so that multiple fits could be plotted to the same frequency-power spectrum several times. This repetition would result in average values and errors, improving the reliability of the method. To remove some bias from the fitting process, the slider did not display numerical values. An integrate button was defined to improve the efficiency of the function. When pressed, this button calculated the integral for both the background and Gaussian respectively, over the entirety of the frequency-power spectrum. The method for determining ratios was therefore: manually fit the Gaussian, click integrate, click reset, repeat several times.

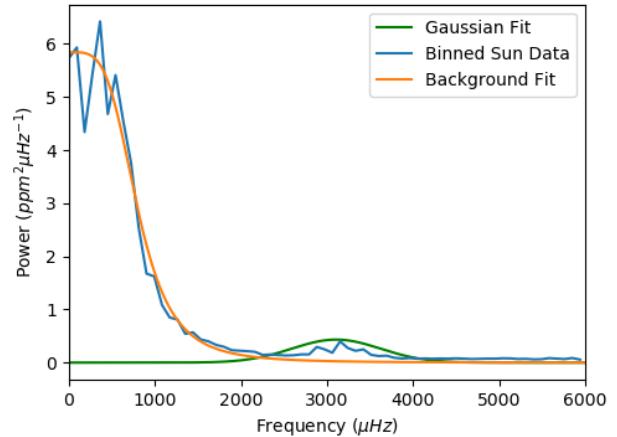


Figure 13.4: The initial slider function as in figure 13.3. In this plot the Gaussian has been manually fitted to the p-modes.

This method was first conducted for the Sun in table 7.

	Power under Gaussian ( $\text{ppm}^2 \mu\text{Hz}^{-1}$ )	Error ( $\text{ppm}^2 \mu\text{Hz}^{-1} 10^{-10}$ )
Fit 1	495.760	5.710
Fit 2	480.55	0.019
Fit 3	446.73	0.002
Fit 4	351.56	3.920
Fit 5	552.59	2.430
Fit 6	412.21	0.022
Fit 7	505.30	0.260
Fit 8	418.55	0.001
Fit 9	491.27	0.268
Fit 10	390.41	6.820

Table 7: Power beneath the Gaussian for the Sun, with errors.

The average power beneath the fitted Gaussian was calculated as  $454.89 \pm 1.99 \times 10^{-8} \text{ ppm}^2/\mu\text{Hz}$ . While the power beneath the background was outputted as  $5287.22 \pm 6.82 \times 10^{-10} \text{ ppm}^2/\mu\text{Hz}$ . From this, the ratio of power was found to be  $0.086 \pm 3.773 \times 10^{-12}$ .

Upon evaluation of this initial slider fitting, two key issues were highlighted. Firstly, the integration should not take place over the entirety of the spectrum. Detectable oscillations, for each star are centred on  $\nu_{\max}$ . Taking the integrals over the whole range of the spectrum merely adds more noise and increases the risk of missing detections. Instead, it is more accurate to consider the oscillations as extending a finite distance either side of  $\nu_{\max}$ . When determining the range over which integration should take place, the following equation was considered (Ball et al. (2018)),

$$\Gamma_{\text{env}} = 0.66(\nu_{\max})^{0.88}. \quad (13.4)$$

Equation 13.4 illustrates how the width of the Gaussian envelope depends on its  $\nu_{\max}$ , where  $\Gamma_{\text{env}}$  is full width half maximum (FWHM). The larger the  $\nu_{\max}$ , the wider the range to be considered. Using equation 13.4, the new integral limits were calculated as  $\nu_{\max} \pm \Gamma_{\text{env}}/2$  and the fitting function was also made to output the

average  $\nu_{\max}$ ,  $\Gamma_{\text{env}}$ , and limits. Within this method, the  $\nu_{\max}$  used is the value obtained from each fit of the slider. From this point, both the power beneath the background and power beneath the Gaussian was calculated with the new integral range.

A second, much more significant issue was the inability to see the majority of p-modes by eye. The bin size had to be reduced to approximately 500 and this resulted in the ability to observe oscillations in only one star. It was a star which had already been classified as a definite detection - star 1 of 20 K dwarfs.

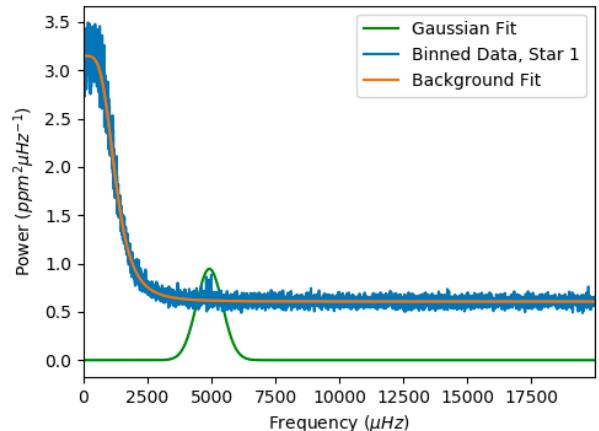


Figure 13.5: The slider function fitted to Star 1.

	Star 1	Errors
Average $\nu_{\max}$ ( $\mu\text{Hz}$ )	4813.10	9.86
Average $\Gamma_{\text{env}}$ ( $\mu\text{Hz}$ )	1148.36	2.07
Average Upper Limit ( $\mu\text{Hz}$ )	5387.28	10.89
Average Lower Limit ( $\mu\text{Hz}$ )	4238.92	8.82
Average Power Ratio	1.67	$2.210 \times 10^{-14}$

Table 8: Key values from the power ratio analysis of Star 1.

Figure 13.5 shows the by-eye detectability of Star 1's p-modes. The analysis was carried out

on this star in the form of 10 manual fits as seen for the Sun data in table 8.

This ratio of 1.67 seemed large compared to the 0.00003175 obtained from the automatic Gaussian fitting code. Upon investigation, it transpired that the automatic code used the previous integral range, ie. the entirety of the spectrum and Gaussian. These limits resulted in an exceptionally large background power due to the initial granulation Gaussian. Unfortunately, changing the limits within the larger, automatic code presented challenges - it was time-consuming and inefficient. Therefore, while it is acknowledged that this power ratio method would be much more accurate by using the aforementioned limits,  $\pm \text{FWHM}/2$ , to make the ratios comparable, the integration needed to happen over the whole spectrum and Gaussian fit.

By reverting to the initial limits, a ratio of 0.091 was calculated for Star 1. While, as expected, the ratio had reduced, it was still not comparable to that calculated in the automatic Gaussian fitting code.

Due to this, further evaluation was conducted. The initial concept of the power ratio method was severely limited by the inability to manually fit Gaussians - it could no longer be used to determine whether the automatic code had fit the p-mode envelope in the correct location by comparing the  $\nu_{\max}$ . Because of this, two further methods were considered in the aim of determining a cut-off ratio: increasing the shot noise (SN) level and investigating its effect on the power ratio, and using scaling relations as an estimate to check for concordant results.

### 13.4 Increasing the Shot Noise Level (Paige)

An additional method to determine the cut-off ratio was to increase the SN level to represent

different levels of noise compared to the p-mode envelope. The SN levels of each categorised ‘yes’ star were first considered to determine approximate magnitudes required in the investigation. The values were found to range between approximately  $0.60 \text{ ppm}^2/\mu\text{Hz}$  and  $3.75 \text{ ppm}^2/\mu\text{Hz}$ . Within this range, 10 different backgrounds were plotted on the frequency-power spectrum of Star 1 - these can be seen in figure 13.6.

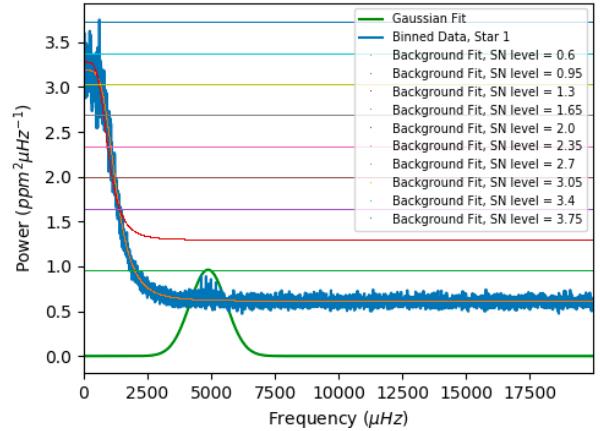


Figure 13.6: The slider function fitted to Star 1 with a range of SN levels.

Integration was conducted for each background several times and an average power ratio for each SN level was calculated.

SN $\text{ppm}^2/\mu\text{Hz}$	Power Ratio	Error ( $\times 10^{-13}$ )
SN = 0.60	0.910	0.628
SN = 0.95	0.065	0.454
SN = 1.30	0.055	0.385
SN = 1.65	0.044	0.303
SN = 2.00	0.036	0.250
SN = 2.35	0.031	0.213
SN = 2.70	0.027	0.185
SN = 3.05	0.024	0.213
SN = 3.40	0.021	0.147
SN = 3.75	0.019	0.628

Table 9: Power ratios for star 1, with errors, over a range of SN levels.

As seen in table 9, the smallest ratio obtained was 0.019. This is still not comparable to the

ratio obtained for Star 1 via the automatic Gaussian fitting code, especially as star 1's SN level is approximately  $0.60 \text{ ppm}^2/\mu\text{Hz}$ .

Through further analysis, it was noted that the automatic code's key aim was solely detection of the p-modes and the precision of the fit was not considered important. This is reiterated when considering figure 13.7, the frequency-power spectrum for Star 1. It can be seen that the fitted Gaussian is minimal in contrast to the visible p-modes. The height of the Gaussian, in figure 13.7 is less than  $0.05 \text{ ppm}^2 \mu\text{Hz}^{-1}$ , while that of the Gaussian plotted manually is approximately  $1.00 \text{ ppm}^2 \mu\text{Hz}^{-1}$ . It is worth noting that while the ratios, unfortunately, aren't comparable, the manual fit is accurate. The manual average  $\nu_{\max}$  is  $4783.10 \mu\text{Hz}$ , while the approximate true  $\nu_{\max}$  is  $4800.00 \mu\text{Hz}$ . In comparison, the fitted  $\nu_{\max}$ , obtained from automatic Gaussian fitting function is  $5500.00 \mu\text{Hz}$ . This proves that the initial idea of investigating misfit Gaussians and improving their position could have been solved with the manual fitting, power ratio code.

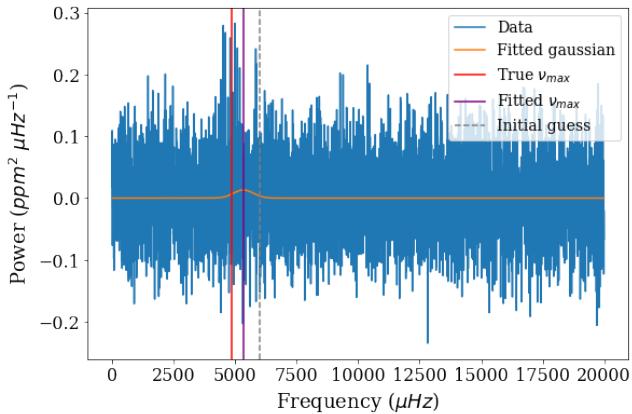


Figure 13.7: The frequency-power spectrum for star 1, illustrating the size of the fitted Gaussian

The second plot, figure 13.8, further illustrates the size of the Gaussian.

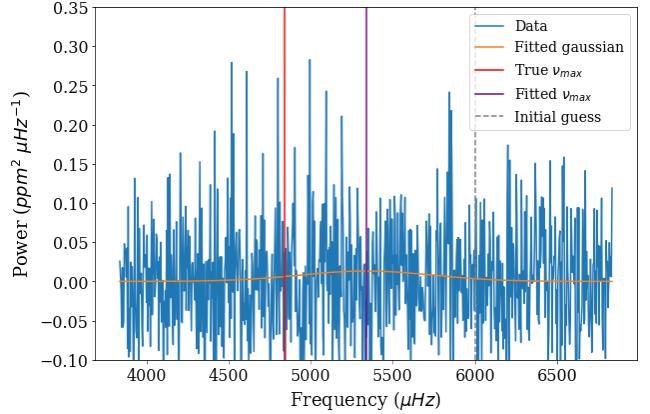


Figure 13.8: A cropped version of figure 13.7 to further illustrate the size of the fit.

Despite the knowledge that the ratios would not be comparable, the results for the power ratio method were still considered hypothetically. It was thought that a ratio of approximately 0.019 or bigger would mean a definite detection. From this, further work using SN levels relating to 'maybe' and 'no' stars was planned to determine a cut-off ratio. However, when this began, it was quickly noticed that the SN levels of 'maybe' and 'no' stars were similar to those of 'yes' stars. For instance, taking 5 'no' stars and investigating their SN levels showed a surprising range from only  $1.792$  to  $4.150 \text{ ppm}^2 \mu\text{Hz}^{-1}$ . While the upper end of this range is much larger than the SN levels seen in 'yes' stars, the lower end is comparable. Considering that three out of these 5 stars were within the same SN range as 'yes' stars, the method of increasing SN levels on only Star 1 seemed inaccurate.

### 13.5 Using Scaling Relations as an Estimate to Check for Congruent Results (Dan L)

Another way of validating the cut-off ratio of stars is to use solar scaling relations. The integral for the Gaussian envelope which contains the p-modes is dependent on its amplitude, width and centre, as per equation 13.2. For the cut-off point of observations, it can be

assumed that the amplitude of the p-mode envelope is approximately equal to the amplitude of the noisy background, or conversely, the p-mode envelope will not be visible if the background noise level has a higher amplitude than the p-modes themselves. It therefore follows that the following approximation can be made,

$$A_{\text{Gauss}} \approx A_{\text{Background}}. \quad (13.5)$$

Previous use of the code has allowed categorisation of stars into 3 categories. Stars which show definite signs of a p-mode envelope, stars which show no distinction from the noisy background anywhere across the frequency-power spectrum and those which cannot be detected 100% of the time yet show some sign of a Gaussian envelope which fall into the ‘maybe’ category. Tighter definitions were needed to dictate whether a star can be observed or not to further categorise ‘maybe’ stars into ‘yes’ or ‘no’ groups. Each of these stars had a corresponding  $\nu_{\text{max}}$  attributed to them. By using the scaling relation shown in equation 13.4 it was also possible to find the corresponding full width half maximum (FWHM) of these stars.

The use of this equation means that two of the three parameters required for the integration of the Gaussian were readily available for the known stars. The final parameter required for this computation, the amplitude of the Gaussian, could be inferred using the approximation shown in equation 13.5, where the amplitude of the background could be determined by utilising the slider function.

The slider function was modified to incorporate the scaling relation for the FWHM of the Gaussian envelope. The new function was then tested on Star 1 of the K-dwarf catalogue. While the  $\nu_{\text{max}}$  of every star was known, Star 1 also had p-modes which stood higher above the background than any other star within the data set. It was therefore decided that this star

would be a good test to ensure that the width determined by the scaling relation was reliable. This plot can be seen in figure 13.9. It can be seen here that the width of the Gaussian function appears to be accurate. The integral for this scenario was carried out 7 times, returning a power ratio of 0.933173. While the calculation of this value is a step in the right direction, the threshold ratio cannot be inferred from this star since it lies in the ‘yes’ category. Now it can be proved that the scaling relation works, the fit needs to be tested on a different star.

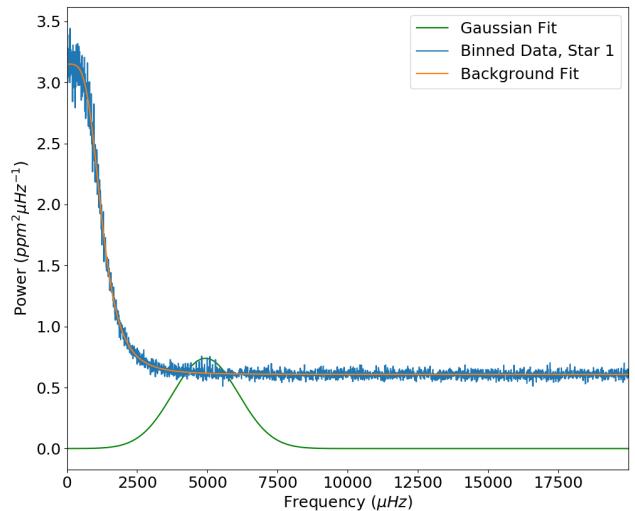


Figure 13.9: A graph highlighting the fitting process used on K-dwarf 1, reliant only on  $\nu_{\text{max}}$  and the amplitude.

Once the check on the validity of the scaling relation had been completed the next step was to pursue a power ratio for a star that was in the ‘maybe’ category of detections. To do this, the  $\nu_{\text{max}}$  and corresponding indexing number of a star in this category was required. The first step of this process was to isolate all stars in the ‘maybe’ category. As previously mentioned each star in the catalogue as a whole had been divided into the ‘yes’, ‘no’ and ‘maybe’ categories and consequently the stars in each category were assigned a value of 2, 1 or 0 respectively. This generated the list of stars in the ‘maybe’ category as shown in the appendix 15.

Each of the stars in this list has a corresponding  $\nu_{\max}$  associated with them. Once the star numbers had been identified the corresponding  $\nu_{\max}$  values could easily be indexed. To correctly identify the  $\nu_{\max}$  of any star using the slider function, values were re-added to the sliders themselves. This allowed for more precise alignment with the correct  $\nu_{\max}$  which could otherwise not be seen. The moderated slider function was then used on Star 6. This fitting process can be seen in figure 13.10. Once the Gaussian had been fitted, the ‘integrate’ function contained within the slider window was used, generating a power ratio of 0.920556. Given that this ‘maybe’ detection could be confirmed as a positive it would be important to set the threshold ratio value lower than this to avoid the exclusion of any true positive readings.

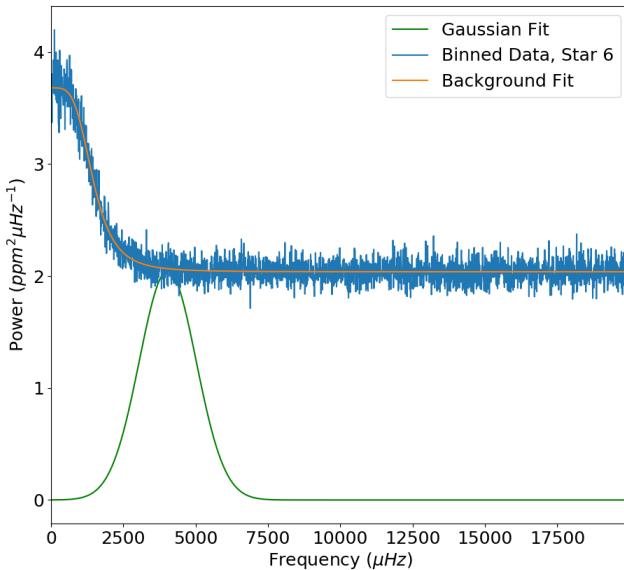


Figure 13.10: The modified slider function showing a fit for star 6 using its known  $\nu_{\max}$  and the scaling relation shown in equation 13.4.

## 13.6 Power Ratio Evaluation (Paige)

Despite the effort and consideration which went into this method, its time-consuming nature

meant that a cut-off ratio couldn’t be determined. Instead, a ratio of 0.019 was deemed a value which dictated whether a ‘yes’ star was a detection. Rather than being a cut-off ratio, it was simply another method of verifying a detection had been made. However, by considering the average  $\nu_{\max}$  calculated using the slider, it was clear that the theory behind how this method could have helped was sound. Complications came in the form of integral limits, imprecise Gaussian fitting in the automatic code and inability to see p-modes by eye. Each of these issues provided varying degrees of difficulty to overcome them.

To overcome the issues with integral range, more time would be required to incorporate the more accurate limits into the automatic Gaussian fitting code.

While the imprecision of the Gaussian fitting could also be improved with time, the power ratio method could instead implement a way to make its fits more comparable. The final data used was flattened. While the ratio outputted by the automatic Gaussian code was done so before flattening occurred, to make the results more comparable, the data could be flattened in the slider method. This would result in lower amplitude, manual fittings and would, therefore, be much more similar to the small Gaussian fit in the automatic code. In further studies, this would be a beneficial route to investigate. Furthermore, in a simpler approach, the amplitude of the slider Gaussian could be fixed at 0.02, as this is the approximate amplitude of the small Gaussian seen in the automatic code. This would allow the width and centre to be varied while limiting the amplitude, therefore, making the final ratio more comparable.

Unfortunately, the inability to detect many of the p-modes by eye had a large impact on this method. Increasing SN levels to represent the

variety of different background levels compared to p-mode heights seemed beneficial in practice. However, when noting the SN levels for 'maybe', 'no' and 'yes' stars, they were all similar. In later work, the range of noise should be further considered and potentially also the relationship between visual magnitude, SN levels and distance should be investigated and added to the method.

## 14 Detection Thresholds (San)

Several different methods were used to quantify successful detections; these involved analysing if a star's p-mode oscillations were still detectable after the noise was added. These methods were important, as the Gaussian fitting could be fitted anywhere near the given initial guesses, regardless of whether the p-modes were visible or not, so identifying what was a correct detection helped rectify issues of this nature.

### 14.1 $\sigma$ test

One of the simplest ways to quantify the detections for theoretical models was a comparison of the true  $\nu_{\max}$  to the fitted  $\nu_{\max}$  from the Gaussian; this could even be checked from just the graphs, as if a fitted value was far away from the true value, it would be visible. Rather than analysing the flat difference, usage of the standard deviation allowed a better comparison between the fitted and true  $\nu_{\max}$  values relative to the Gaussian.

The difference between the true and fitted  $\nu_{\max}$  was computed, and compared to the value of  $1\sigma$ . The value of  $\sigma$  required knowledge of the FWHM value for the peak in the Gaussian fit, and could be found from the equation ([W. R. Leo](#)):

$$\sigma = \frac{\text{FWHM}}{\sqrt{2 \ln 2}}. \quad (14.1)$$

The value of  $1\sigma$  was chosen, as it was small enough that it would be easy to see if the Gaussian was being fitted in the wrong place. Likewise, the value is also large enough such that the variability arising from performing many trials (when fitting the Gaussian) is taken into account. A value within  $1\sigma$  of the fitted  $\nu_{\max}$  was desired to ensure accuracy, and was used as the detection threshold for this method (as seen in figure 14.1); a true  $\nu_{\max}$  further than  $1\sigma$  from the fitted  $\nu_{\max}$  would return a 'fails' output. This would indicate that the Gaussian was being fitted to noise rather than the p-modes of oscillations. For this test,  $1\sigma$  was chosen as a reasonable threshold. To obtain a list of stars consistently within the  $1\sigma$  limit, a large number of iterations were computed (2500 trials). Stars that consistently fit these criteria were then compared to the 'detectable' stars outputted from varying the Gaussian & grid comparison test below. Stars suggested as being 'detectable' from both these methods were likely to be real detection in a PLATO observation.

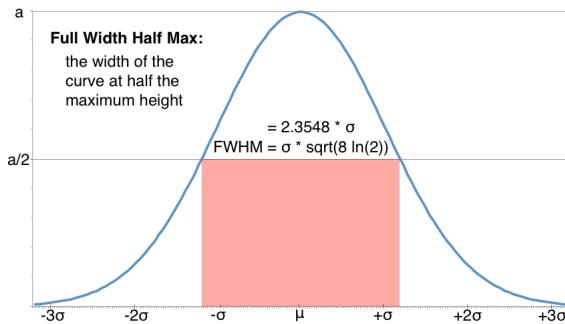


Figure 14.1: Shows a Gaussian distribution with the positions of  $1\sigma$  & length of FWHM; the length of the FWHM is approximately 2.35 times the standard deviation. Using the standard deviation can give a reasonable idea of detectability (Credit: [Brainmap.org](#)).

This method was not perfect, as there was always a risk of incorrect fitting for the Gaussian. The frequency arrays also had to be considered, as they weren't completely continuous, mean-

ing the positioning of the FWHM and  $\sigma$  could be offset from their true values- this could offset the detection criteria, even if it was a small amount. Knowledge of the FWHM required values of the indices in the frequency array at the peak and points of half-maxima. But once again, as the values in the frequency array were not continuous, the indices returned were not always at the exact locations, deviating the values of the desired parameters from their true value. Even for these detection limits, there existed a small error.

Another issue comes from the usage of the true  $\nu_{\max}$ . For experimental modelling, this is acceptable as parameters like  $\nu_{\max}$  are known, but for observations by PLATO, this value would not be known, making the method harder to carry out. This could be partially rectified by using known  $\nu_{\max}$  values from similar stars, or using scaling relations from equation 5.8 (would require the mass and radius values). But, in practice the, modelling process followed might be less straightforward for real observations by PLATO than for model stars. So it was important to use other methods alongside this to quantify the detections.

## 14.2 Varying the Gaussian: ‘Plateau’ test and Grid Stars comparison

A problem with the Gaussian fitting is the bias of the fit towards the initial input guesses (of the centre and widths). Unless the original signal had a strong, distinguishable set of p-modes, the Gaussian would be fitted to a section of noise and would give a false detection. This was especially prevalent for fainter stars, where the noise was more significant.

As previously mentioned, a grid of 266 fake stars (originating from 14 initial stars, see figure 12.10) with a varying  $\nu_{\max}$  and apparent

visual magnitude ( $m_v$ ) were produced by the TSM group and passed on for analysis. For each star in the grid catalogue, the initial guesses for the centre of the Gaussian were varied, and the fitted  $\nu_{\max}$  found. After plotting the calculated value against the initial guesses, the grid stars were manually categorised accordingly to three levels of detectability ('definite', 'possible' or 'no detection'). Shapes of graphs like those seen previously in figures 12.4 & 12.5, showed if a Gaussian fit had been accurately done or not.

The point of the grid was to provide a clear measurement of the cutoff for detectability based on the properties of  $\nu_{\max}$  &  $m_v$ . The range of the straight section in the graph (like in figure 12.4) provided a measure of detectability, as the wider the plateau was, the more likely a star's oscillations were being detected and not the noise. By using the grid stars, there was an approximate guideline for detectability that could be applied to the catalogue of K and M dwarfs.

Next, a graph of each K and M dwarf's visual magnitude against its true  $\nu_{\max}$  was plotted, with the colour coding for detectability of the grid stars overlaid on the same graph. By putting this information on the same graph, it could be seen how 'detectable' the K and M dwarfs were compared to a standard model. This was also helpful as it gave an idea of whether the  $\nu_{\max}$  and  $m_v$  of a star could affect its detectability.

Another advantage of this method was that it ensured that if a fitted  $\nu_{\max}$  was close to the true value, it could be checked whether this was just a coincidence or if the Gaussian was fitted correctly (as it looked over a large range of initial guesses).

As with any method, this one did not come without disadvantages. The method required

a manual input stating whether the star was detectable or not. This introduces an error as no observer is the same, so what is a ‘definite’ or ‘potential’ detection is not necessarily consistent. By asking multiple observers to follow this procedure for the same set of stars and finding the average, this subjectivity is reduced. The method is effective, in this case, for the small numbers of stars nevertheless, the process could become tedious if the catalogues contain thousands of stars. A more automated process could improve efficiency.

The initial guesses only varied the centre of the Gaussian - they did not change the height or widths. There is a possibility that altering these parameters could have changed the detectability for some stars, as the shape of the signal was altered by the binning. Implementing differing parameters for each star would be an improvement for this method, allowing for a better fit to be adjusted to the signal.

### 14.3 Power Integration Ratio

As previously mentioned, integration under the background curve and Gaussian fit can help quantify the detectability. Whilst varying the parameters of the Gaussian manually (like the centre, width etc.) to fit the curve, the ratio of the power under the fitted Gaussian to the power under the background curve (these two curves can be seen in figure 13.4) was found. Stars found to give positive detections from the previous two methods were tested via this method; this involved integration and the limits taken were FWHM/2 either side of the fitted  $\nu_{\max}$ . For a successful detection, this ratio must be above a threshold value (e.g for star 1, this was found to be 0.019). Further refinement on the cutoff value came via variation of the shot noise on the star’s signal and by making the background signal’s amplitude equivalent to the Gaussian peak value (this is also detailed in the Power Integration Fitting section

above).

Unlike other methods, the Gaussian was fitted manually, requiring the judgment of the researcher(s), so this method assumed that the Gaussian was in the correct position when finding the cutoff. If this was used as the first check, it could give false detections as it would find a successful detection for a ‘star’, which had a fitted  $\nu_{\max}$  very far away from its true  $\nu_{\max}$ . Similarly to the grid method, the manual input could be very cumbersome for a larger number of stars. Due to the requirement of manual input and the risk of human error, it was key that this method was the final check for detectability, and used after the filtering of results from the  $\sigma$  & initial guesses tests.

From the usage of all three methods, detections of stellar oscillation modes could be confidently confirmed or denied. Initially, the standardised grid gives approximate  $m_v$  and true  $\nu_{\max}$  limits for a star observed by PLATO, and comparing with the results from the  $1\sigma$  test provides stronger evidence of a positive or negative detection. The ratio test serves as a final check for the positive detection of a star’s oscillations. Overall, some control in the variability of the Gaussian fitting and the effects of noise was found, as confirmation from all three tests is required for a positive detection. By the limits on detectability provided in these methods, it is hoped that PLATO can identify and discard false positives in the same way.

## 15 Measuring Global Parameters (K-Ryan)

Along with verifying and establishing the detection thresholds, the subsequent progression was to return global parameters that would give an insight into the properties of the stars. This was a significant step as, following a confirmed detection, returning measurable parameters became just as important as confirming detec-

tions. Without these parameters, the subsequent analysis of the star and possible orbiting exoplanets wouldn't be possible. The two most important parameters returned were  $\nu_{\max}$  and  $\Delta\nu$  which are detailed below.

## 15.1 Measuring $\nu_{\max}$ and $\Delta\nu$ and from Gaussian Fitting

The frequency at which the Gaussian envelope is a maximum,  $\nu_{\max}$ , was returned directly from the Gaussian Fitting function - detailed in full above - and came out of the curve fitting function in the python module `scipy`. This function also returned the full-width-half-maximum (FWHM) which is the full width of the Gaussian envelope at half the maximum height.

to calculate the large frequency spacing,  $\Delta\nu$ , the  $\nu_{\max}$  and FWHM were fed back into a function that calculated the large separation, making use of the python module `peak.utils` - specifically the functions '`indexes`' and '`pyplot`'. '`Indexes`' returned the indices of the frequency-array corresponding to the middle of the detected peak based on two values: threshold and minimum distance. These refer to the threshold above which a peak is considered as a peak and took a value from 0-1. Minimum distance defined the minimum distance between peaks and depended on a user-defined quantity (which was appropriate to the specific case). Both of these quantities were relative, with the minimum distance being simpler to specify as all that needed to be considered was the balance between too many and too few peaks. The threshold was relative to the most intense peak and needed to be carefully selected. Due to the relative ambiguity of the data (the data contained all p-modes meaning the number of peaks originally in the spectra was unknown and could not be matched to), determining an adequate threshold was important as again, a balance needed to be achieved between too many and too few peaks.

The function first centred the peak-finding operation on  $\nu_{\max}$  and the width of this was defined as roughly  $4 \times \text{FWHM}$ ; the upper and lower bounds were defined as  $\nu_{\max} + 4 \times \text{FWHM}$  and  $\nu_{\max} - 4 \times \text{FWHM}$  respectively. The frequency-array was cropped to be this size and the function '`indexes`' used only this range of values - 4 was decided by testing integers from 1-6 and analysing where the peak envelope roughly ended.

The next step was to determine the mean number of peaks the '`indexes`' function found. Due to the random nature of the noise, each time the noise was added it would be different and some iterations had extremely high noisy values. These high values were only sometimes detected as peaks so the number of detected peaks varied each time noise was added. Thus, it was reasonable to average the number of peaks using a semi-arbitrary threshold value of 0.375 (from testing using the original solar data, this value gave the most consistent detection of 13 peaks) and it was averaged by adding noise to the data 100 times and detecting the number of peaks found using the `indexes` function; the mean number of peaks were then found. To find the optimal threshold value, noise was again added 100 times and in each iteration, the `indexes` function was run with an increasing threshold value, starting at 0.25 and ending at 0.5. Over the 100 iterations, the threshold values which returned an index value equal to the mean number of peaks were averaged over and this returned the optimal threshold value. This was done because an arbitrary threshold value could be added and could be too high or too low due to the randomness of the noise and a varying heights of the peaks due to the noise. Finding the thresholds that returned the same number of peaks as the average proves that the random noise affects the '`indexes`' function as different thresholds combined with different iterations of noise detected the same number of

peaks. This process served the dual purposes of also constraining the threshold value such that on any iteration of noise, the number of peaks and thresholds adjusted accordingly. The large separation was then calculated by finding the difference in frequency between the peaks and multiplied by 2 as every other peak was a peak of the same angular degree - this being the definition of large frequency separation. The function itself then returned the average number of peaks detected and the large separation with errors on both values.

## 15.2 Measuring $\Delta\nu$ for Solar-Like Spectra

At the time of creating the large separation function, models of M and K dwarfs were not yet available meaning analysis was performed on Sun-like stars. The  $m_v$  of the Sun model was varied, effectively moving the distance of the star which subsequently affected the shot noise levels. For the Sun, it was known that the  $\nu_{\max} \sim 3000 \mu\text{Hz}$  so this was put into the large separation function with the  $m_v$  varying; the large separation was calculated for different  $m_v$ . The main result from this test was that up to a  $m_v \sim 6$ , the large separation was consistently calculated to be  $\sim 67 \mu\text{Hz}$  and then above  $m_v$  of 6, the large separation increased. Due to this having been tested on the Sun first, the accuracy could be cross-checked with known values as the true  $\Delta\nu$  is known to be  $135\text{-}136 \mu\text{Hz}$  [Mullan \(2013\)](#). The value returned from the function was roughly half; this was due to the function calculating the distance between successive modes instead of only the modes of the same angular degree, however, this was quickly rectified by multiplying the measured  $\Delta\nu$  by two to account for this. Subsequently, the measured  $\Delta\nu$  fell within the accepted range,  $135\text{-}136 \mu\text{Hz}$ , up to  $m_v \sim 6$ . The values deviated after this  $m_v$  due to the peaks becoming less prominent in comparison to the surrounding noise meaning it became harder for the func-

tion to distinguish between a peak and the random noise. Figures 15.1 and 15.2 are examples of the Sun-like spectra with  $m_v$  1.33 and 8 analysed determining  $\Delta\nu$  to be  $135.55 \pm 11.18 \mu\text{Hz}$  and  $139.86 \pm 22.41 \mu\text{Hz}$  respectively. The Solar-like spectra then were mainly used to optimise the function and test the limits to see where, and under what conditions the function stopped working or failed to return suitable values.

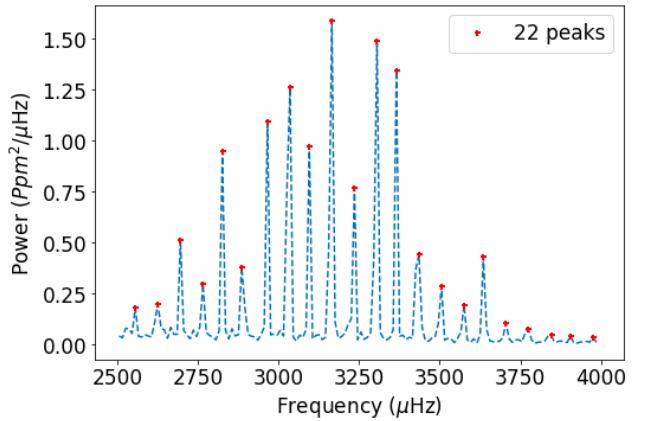


Figure 15.1: Solar-Like spectrum with  $m_v$  1.33 and a measured  $\Delta\nu = 135.55 \pm 11.18 \mu\text{Hz}$

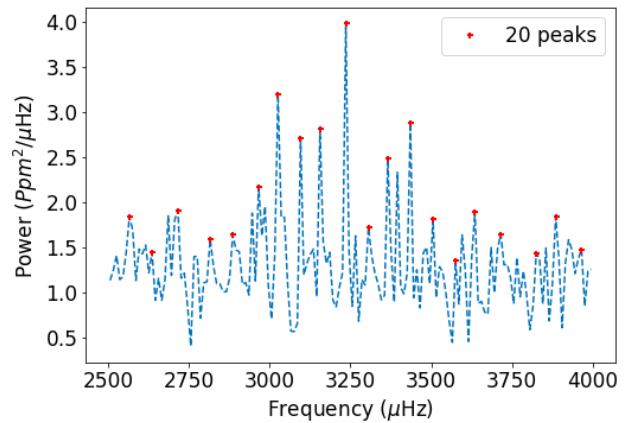


Figure 15.2: Solar-Like spectrum with  $m_v$  8 and a measured  $\Delta\nu = 140.29 \pm 25.95 \mu\text{Hz}$

It is clear from figure 15.2 that as the p-modes become drowned out by the noise, distinguishing the peaks from the noise was much more difficult; the main consequence of this being that

the large spacing then is calculated incorrectly. This was a known issue approaching the subject which was the reason behind having the centre of the Gaussian and the FWHM be an input to limit the region in which to detect peaks.

$m_v$	$\Delta\nu$ ( $\mu\text{Hz}$ )	$\sigma_{\Delta\nu}$ ( $\mu\text{Hz}$ )	#Peaks
-2	136.00	12.00	21
-0.33	136.92	13.99	21
1.33	135.55	11.18	22
3	135.31	10.55	22
4.67	135.10	11.00	22
6.33	137.36	17.80	20
8	140.29	25.95	19
9.67	142.89	27.96	18
11.33	144.37	28.68	19
13	145.42	29.93	19

Table 10: Large Separation for Sun-like stars with a varying  $m_v$  and the mean number of peaks detected.

Table 10 is the primary data collected from the Sun-like spectra with a varying  $m_v$  to emulate the same star but at different distances.

Once the DA group received their first M and K dwarfs, the large separation function was run on all of the stars in the catalogues and at this stage, the stars whose p-modes were prominent enough to be detected by this function were stars 1 and 4. Their  $\Delta\nu$  were successfully measured to be  $189.19 \pm 3.74 \mu\text{Hz}$  and  $179.00 \pm 10.33 \mu\text{Hz}$  and their spectra can be seen in figures 15.3 and 15.4 respectively - the vertical green line marks  $\nu_{\max}$  and the horizontal red line represents the threshold.

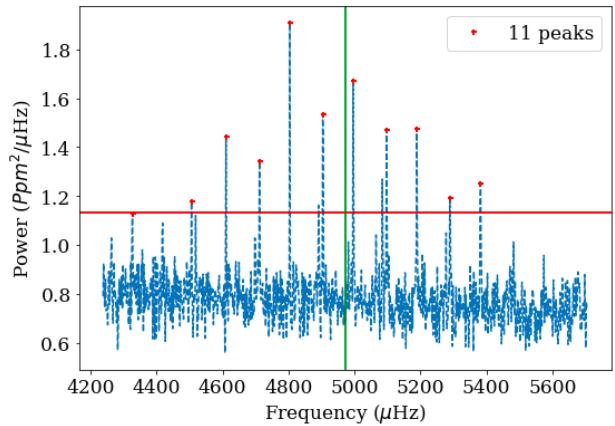


Figure 15.3: First K dwarf with  $m_v$  7.319 and a measured  $\Delta\nu = 189.19 \pm 3.74 \mu\text{Hz}$ .

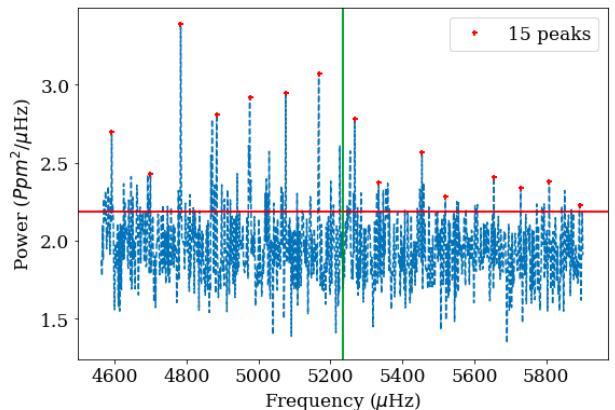


Figure 15.4: Fourth K dwarf with  $m_v$  8.5 and a measured  $\Delta\nu = 173.60 \pm 9.30 \mu\text{Hz}$ .

When the finalised catalogues of M and K dwarfs were received, the large separation function was then run on the stars deemed to be definite and maybe detections. It was not run on all the stars in the catalogue as the stars that couldn't be detected would not have had any visible p-modes to analyse because they could not be seen over the noise. The Gaussian Fitting placed the definite and possible detections in a separate catalogue and this function was run on those stars as the p-modes of these stars were detectable.

# 16 Results and Analysis (K-Ryan and San)

## 16.1 Shot noise & Magnitude for the catalogue stars' range (San)

The catalogues provided by the TSM group contained 40 K & M dwarfs, so there were 80 stars available to analyse for their detectability. It was clear even from the offset that fainter stars should not be considered as their shot noise increases significantly, making the p-modes less clear in the spectrum, even after binning. From the shot noise equation, the power often causes the shot noise to increase rapidly, so the list of detectable stars is much lower than the total number of model stars and is biased towards the lower magnitude end of the catalogues. Looking at the range of  $m_v$  in the catalogues, the brightest star (K dwarf) was  $m_v +7.319$  and the faintest (M dwarf) was  $+12.578$ , so even stars at the lower end are affected by the shot noise- this can be seen previously in figure 5.26.

## 16.2 $1\sigma$ test observations (San)

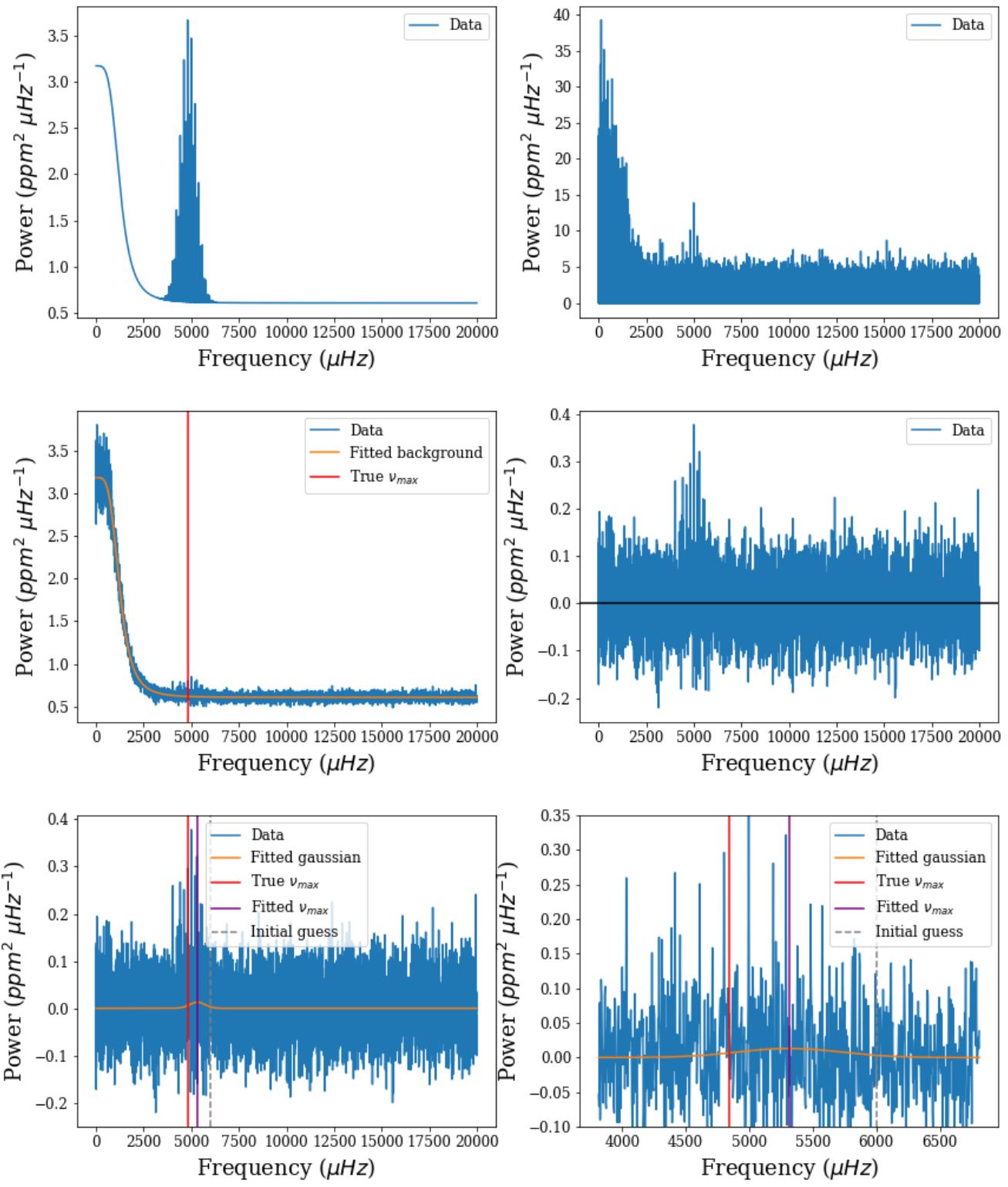
As stated above, for each star there were 2500 iterations run to fit a Gaussian, and this was done as there is a variability arising from adding noise to the original signal. A high number of iterations also improves the reliability behind the results. For this purpose, an initial guess of 6000  $\mu$ Hz, a width of 750  $\mu$ Hz and a bin size of 250 was chosen; the idea behind this bin size was that it should be high, but still close to the large frequency separation of the stars. As the Sun has a value of around 135  $\mu$ Hz ([Mullan \(2013\)](#)), both K & M dwarfs should have a value of a similar order so this value meets these conditions.

### 16.2.1 K dwarfs (San)

Analysing table 11, it is clear the highest percentage fits come from the top two stars (star 1 in figure 16.1), and this was expected as these were the brightest stars analysed, so their signals were the easiest to detect; stars 3, 5, & 6 were also worth keeping in mind. Although all stars were checked with the grid stars method afterwards, these stars were given particular attention to as they were the most likely candidates for a ‘detectable’ signal. It was also noted (though not always true) that when the difference between the fitted and true  $\nu_{\max}$  was lower, the percentage fit tended to be better- this makes sense mathematically as an accurate fit improved the probability for a positive result.

Star no.	Percentage fit to $1\sigma$ (%)	Fitted $\nu_{\max}$ ( $\mu\text{Hz}$ )	True $\nu_{\max}$ ( $\mu\text{Hz}$ )	$m_v$
1	65.56	5314.31	4843.53	+7.319
2	50.80	6003.09	5978.78	+7.999
3	42.80	6186.97	6096.24	+8.495
4	23.60	5932.21	4859.73	+8.500
5	42.04	6037.29	5621.63	+8.587
6	41.08	6032.94	5676.26	+8.634
7	38.12	5997.71	5498.10	+8.821
8	3.44	5979.15	4766.09	+8.859
9	38.72	6166.74	5873.71	+8.963
10	39.28	6163.29	5965.03	+9.042
11	24.08	6009.48	5092.87	+9.062
12	35.36	6080.80	5512.04	+9.144
13	13.64	6090.71	4705.57	+9.149
14	28.68	6011.72	5308.38	+9.212
15	7.52	6128.48	4487.58	+9.288
16	16.56	6076.81	4940.40	+9.292
17	35.16	6092.07	5600.31	+9.407
18	37.08	6159.33	5766.50	+9.429
19	28.48	6055.31	5376.66	+9.445
20	37.80	6151.12	6081.52	+9.464
21	24.45	6129.26	5270.63	+9.470
22	0.45	6213.70	5515.85	+9.606
23	31.05	6110.96	5524.11	+9.630
24	35.20	6155.22	5697.21	+9.658
25	37.05	6204.39	5775.14	+9.665
26	14.05	6074.98	4914.87	+9.710
27	26.55	6153.35	5403.34	+9.771
28	24.85	6079.93	5331.66	+9.838
29	6.90	6198.97	4594.75	+9.841
30	27.20	6112.43	5423.91	+9.872
31	15.55	6232.07	5630.70	+9.905
32	16.60	6176.79	5064.80	+9.909
33	32.10	6246.09	6371.07	+9.912
34	6.20	6151.55	4613.84	+9.975
35	4.00	6207.25	4705.57	+10.002
36	29.10	6222.90	6444.84	+10.036
37	40.20	6235.57	5917.83	+10.041
38	14.10	6103.30	5005.09	+10.045
39	28.85	6183.44	5522.20	+10.054
40	16.15	6212.13	5089.35	+10.104

Table 11:  $1\sigma$  test done on the K dwarf catalogues showing the percentage of times from 2500 iterations, that the true  $\nu_{\max}$  was within  $1\sigma$  of the fitted  $\nu_{\max}$  ( $\nu_{\max}$  values are to 2d.p.). A higher value suggests the star's p-modes were more detectable, and were compared with detectable stars found from the other detection methods.



1639 fits made that lie within 1 sigma of  $v_{\max}$     65.56 % detections rate in Star 1 at visual magnitude of  $v = 7.319$  based on 1 sigma fitting  
 861 fits made that lie outside of 1 sigma of  $v_{\max}$

The average fitted  $v_{\max}$  is 5314.31 which is 470.78 away from the true value of 4843.53  
 The mean power under the background is 15676.223 whilst the mean power under the oscillation Gaussian is 0.200

Figure 16.1: This is star 1, which had the highest percentage fits within  $1\sigma$ , so was more likely to be a ‘detectable’ star; also shows details of the integration under curves.

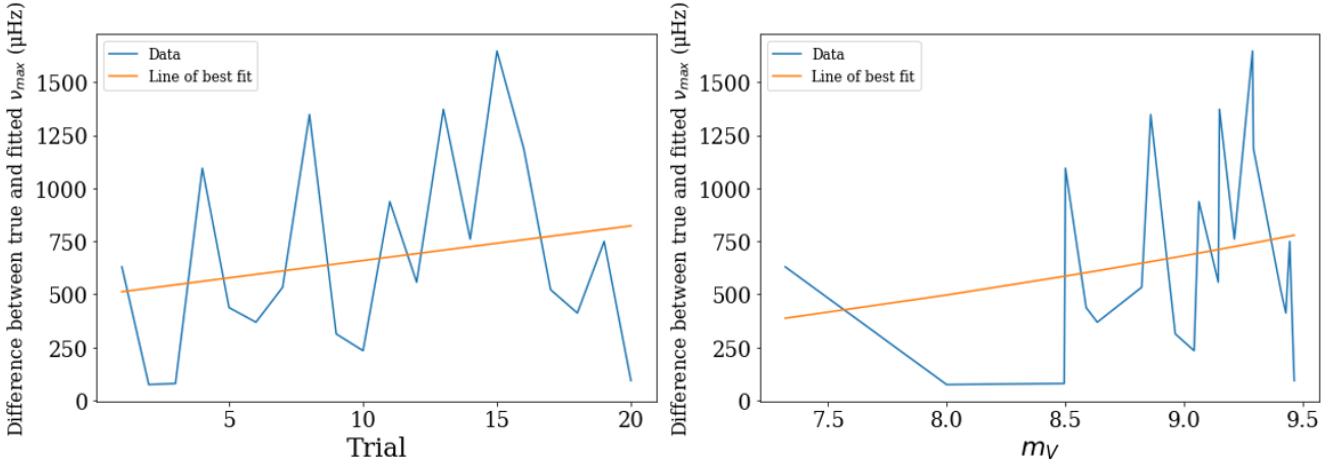


Figure 16.2: This looked at how the difference from the true  $\nu_{\max}$  varies with each trial and  $m_v$  for the 1st-20th K dwarf; it can be seen that the pattern is quite erratic.

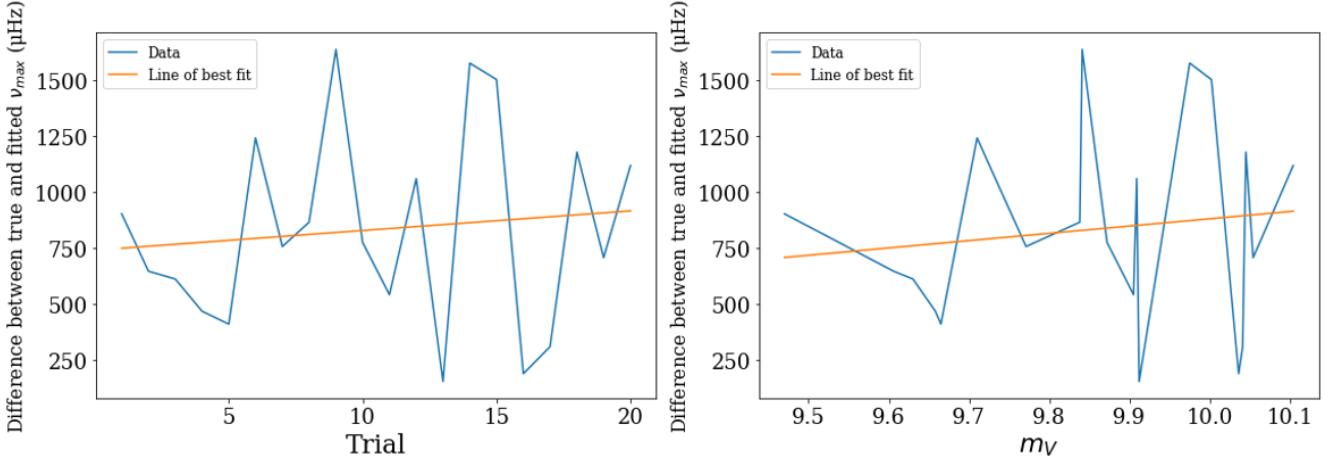


Figure 16.3: this looked at how the difference from the true  $\nu_{\max}$  varies with each trial and  $m_v$  for the 21st-40th K dwarf.

It was clear that at higher magnitudes, the percentage fit to  $1\sigma$  was lower as the Gaussian became harder to fit accurately. In the first 20 dwarfs, there were several stars with over a 40% rate for being within  $1\sigma$  and two that had over a 50% rate; for the next 20 stars, however, this rate plummeted with only 1 star having a greater than 40% rate. Generally speaking, the first 20 K dwarfs had a higher percentage fit to

$1\sigma$  compared to the next 20. It was quite evident just from examining the spectra by eye at around magnitude 8-8.5, that the p-modes became difficult to distinguish from the noise, so the fitted  $\nu_{\max}$  was less likely to be near to the true value. It was also interesting to note that for stars with very low percentage fits (e.g. star number 15, 19, 22 and a few others), the Gaussian was not fitted in the majority of the itera-

tions, contributing to this low percentage value. These stars were likely to be fainter so the effects of weaker p-modes and more shot noise were present, creating a more chaotic spectrum. This is also backed up by figures 16.2 & 16.3, where the difference between the true & fitted values did not show a consistent pattern and fluctuated; the randomness of the pattern gave more evidence that the Gaussian was being fitted to the noise at higher magnitudes.

### 16.2.2 M dwarfs (San)

The signals from the M dwarfs were much harder to distinguish from the noise, so this test did not work well for the M dwarf catalogues; the background interference and shot noise meant that the code could not find optimal parameters for a Gaussian fit on the majority of the M dwarfs, so this test only worked for a limited number of stars in the catalogue (only 8 M dwarfs gave any results as seen below in table 12).

Star no.	Percentage fit to $1\sigma$ (%)	Fitted $\nu_{\max}$ ( $\mu\text{Hz}$ )	True $\nu_{\max}$ ( $\mu\text{Hz}$ )	$m_v$
1	33.04	9829.08	9680.82	+8.018
2	0.80	9884.96	7073.96	+9.769
3	1.76	9919.98	7998.57	+10.169
4	1.12	9847.71	7330.14	+10.256
5	16.20	9906.29	9211.12	+10.291
6	9.40	9889.69	11041.10	+10.525
7	2.88	9812.56	8456.92	+10.760
8	4.60	9768.30	11372.67	+10.831

Table 12:  $1\sigma$  test done on the M dwarf catalogues showing the percentage of times from 2500 iterations, that the true  $\nu_{\max}$  was within  $1\sigma$  of the fitted  $\nu_{\max}$ . The code could not find optimal parameters for the majority of M dwarfs available due to faintness and higher background interference experienced.

As the M dwarfs in the catalogue covered a magnitude range of +8 to +12.5, this suggested that M dwarfs have an inherently more noisy

spectrum compared to K dwarfs of similar visual magnitude. Once again, a bin size of 250 and 2500 iterations were chosen, the initial guess for the centre, however, was changed to 10,000  $\mu\text{Hz}$  as this better reflected the  $\nu_{\max}$  values for M dwarfs. It was important to change the initial guesses as otherwise, the Gaussian fitting would always be occurring on the noise instead of the p-modes, especially for these faint, noisy M dwarfs.

The M dwarfs faced similar problems to the K dwarfs, but it was significantly amplified, thus realistically, none of the M dwarfs shows a very reliable fit to the  $1\sigma$  limit. With such a limited set of results, it was hard to reach a better understanding from just this method.

Reducing the number of iterations gave more success, as parameters were fitted for more stars (giving more results), but this wasn't a viable option as the low number of iterations introduced a larger error. The effect of initial guesses also became more significant as with a noisier spectrum, it was more likely that the Gaussian was fitted onto noise near the initial guess than the p-modes. It would have been more accurate to have a set of initial guesses for each star that was close to the true  $\nu_{\max}$  values for M dwarfs, but implementing such a method would take much more time, and could be problematic for larger catalogues. As stated before, there were issues with using the  $1\sigma$  method, especially for fainter stars as factors came into play limiting the accuracy of the signal, so a stricter implementation of this method would be required in the future. It was hoped that the other two detection thresholds could provide more tangible conclusions for the targets of interest (M dwarfs in particular).

### 16.3 Grid Stars (K-Ryan)

The procedure of categorising stars based on the Grid Method (Method 2) is detailed in full

above with the main result being that only one star from the catalogue of 80 M and K dwarfs were said to within the ‘possible detection’ category - ‘possible detection’ referring to possibly having detected oscillations. This star was from the K dwarf category and had  $m_v$  of 7.319 and can be seen in figure 12.11 within the possible detection range. The significance of this result can not be understated as it is an achievement to be able to detect the oscillations in any star of this size.

As discussed previously, it becomes increasingly more difficult to detect these low mass stars due to the smaller and smaller change in flux caused by these oscillations. For this particular K dwarf, the variation in flux around the p-mode was on order  $\approx 1 \text{ ppm}^2/\mu\text{Hz}$  with the random noise being of the same order. Whereas, for the Sun-like p-mode oscillations although they were of the same order,  $\approx 1 \text{ ppm}^2/\mu\text{Hz}$  - the random noise in comparison then became negligible which is why the two spectra for the Sun and the K dwarf look notably different (figures 15.1 and 15.3). Moreover, the spectra of the majority of the M and K dwarfs look closer to figure 15.2 due to the spectra being dominated by noise. This stems from both the shot noise level and the vast difference in  $m_v$  which were  $-26$  and  $7.319$  for the Sun and K dwarf respectively. This drastically influenced the detectability resulting in only one K dwarf populating the ‘possible detection’ area. The black markers in the ‘no detection’ region are K dwarfs that produced similar graphs to that seen in 15.2 where a detection could not be made via any method; the remaining M and K dwarfs are not within the limits of this graph and are definite ‘no detections’. To increase the detectability of these M and K dwarfs, more substantial noise reduction should have been employed to reduce the noise levels as this was the main issue faced during the analysis of these smaller stars. Having the ability to see the raw spectra and then adding noise (the

process seen in the upper two panels in figure 16.1) was unique as the extent to which noise interferes with measurements was completely displayed. A similar response is seen when the noise is added to Sun-like data however, the p-modes were strong enough to not be drowned out by the noise. Binning the data was useful to amplify the signal but it also had the same effect on the noise; to amplify the p-modes alone, either the noise needed to be removed or reduced. This process was attempted, to some effect by adding the noise n times and averaging the noise over those n iterations but no noise reduction was completed.

The advantage of displaying the data in this grid format is that the region of ‘definite detections’ and ‘possible detections’ are defined meaning, with the detection methods employed, stars within this region would be detectable. The grid shows where the majority of the K dwarfs lie, if K dwarfs with  $m_v$  lower than  $\approx 7.5$  were observed, the likelihood of being able to confirm a detection would be higher - even if it fell into the ‘possible detection’ - as then they would be able to be verified via the power integration method. Unfortunately, only one of the stars fell within this limit and so this could not be tested. If this wasn’t the case and multiple stars had  $m_v < 7.5$ , the lower half of the grid would be more highly populated and a more detailed sense of the limits of detections would be obtained.

## 16.4 Temperature from fitted $\nu_{\max}$ values (K-Ryan and San)

Using the relationship found in figure 6.17 (to re-quote the result,  $\nu_{\max} = -1.4374T_{\text{eff}} + 12148$ ), the temperatures of all the K dwarfs were successfully calculated and these values can be seen in Table 13. As is explored above, Star 1 became the only star that was deemed a

‘possible detection’ and thus, it is the most significant star as it represents a detectable star; moreover, all parameters one can measure from this star are measurable parameters future missions could detect and catalogue which is an important aspect not to overlook. For Star 1, the Fitted  $\log_{10} T$  and Actual  $\log_{10} T$  differ by  $2.35 \times 10^{-2}$ (K) corresponding to a temperature difference of  $\sim 260$ (K). A similar relationship was found which allows for the surface gravity,  $g$ , of the star to be calculated from the fitted  $\nu_{\max}$  (see appendix A9); the surface gravity was calculated to be  $\sim 42000 cm/s^2$ . For comparison, the surface temperature of the Sun is 5777 K (from the Astropy constants library - [Robitaille et al. \(2013\)](#)) and the surface gravity is  $27400 cm/s^2$ . The relationship between surface gravity of an object and both mass and radius is shown in equation 16.1.

$$g \propto \frac{M}{R^2}. \quad (16.1)$$

For K dwarfs, it is known that their masses are much less than that of the Sun and so are their radii but due to the radius being raised to the power of 2, this would result in the surface gravity of lower mass objects to be greater than the surface gravity of more massive objects and this is seen. Being able to calculate multiple parameters for stars again is promising. Had more stars populated the ‘definite detection’ and ‘possible detection’ areas in figure 12.11, more results would be available, allowing a more detailed analysis to take place.

Star no.	Fitted $\nu_{\max}$ ( $\mu$ Hz)	Fitted $\log_{10} T$ (K)	Actual $\log_{10} T$ (K)
1	5314.31	3.6771	3.7006
2	6003.09	3.6313	3.6457
3	6186.97	3.6326	3.6188
4	5932.21	3.6192	3.7037
5	6037.29	3.6195	3.6512
6	6032.94	3.6288	3.6528
7	5997.71	3.6313	3.6705
8	5979.15	3.6326	3.7046
9	6166.74	3.6192	3.6331
10	6163.29	3.6195	3.6197
11	6009.48	3.6305	3.6890
12	6080.80	3.6254	3.6683
13	6090.71	3.6247	3.7077
14	6011.72	3.6303	3.6770
15	6128.48	3.6220	3.7149
16	6076.81	3.6257	3.6994
17	6092.07	3.6246	3.6725
18	6159.33	3.6198	3.6431
19	6055.31	3.6272	3.6739
20	6151.12	3.6203	3.6249
21	6129.26	3.6219	3.6732
22	6213.70	3.6158	3.6797
23	6110.96	3.6232	3.6584
24	6155.22	3.6201	3.6756
25	6204.39	3.6165	3.6518
26	6074.98	3.6258	3.6999
27	6153.35	3.6202	3.6716
28	6079.93	3.6255	3.6752
29	6198.97	3.6169	3.7104
30	6112.43	3.6231	3.6623
31	6232.07	3.6144	3.6658
32	6176.79	3.6185	3.6918
33	6246.09	3.6134	3.5905
34	6151.55	3.6203	3.7128
35	6207.25	3.6163	3.7077
36	6222.90	3.6151	3.5925
37	6235.57	3.6142	3.6226
38	6103.30	3.6238	3.7107
39	6183.44	3.6180	3.6607
40	6212.13	3.6159	3.6196

Table 13: This was the temperatures gained from the fitted  $\nu_{\max}$  values, using the relationship for temperature &  $\nu_{\max}$  previously seen in 6.17 for K dwarfs.

## 16.5 Power Integration Ratio (San)

From filtering via the two other detection thresholds, only 1 star was found to be a ‘possible detection’, when observed by eye so this star’s power was analysed. Through this analysis, the smallest ratio for the power integration was 0.019 in K dwarf 1; this suggested that for a star like this, signals which show a ratio higher than around 0.019, are more likely to be the p-modes being detected by observations. However, this integration ratio does not agree with the values seen from the automatic Gaussian fitting, as the ratio found in the automatic process is 0.0000128 (around 1000 times smaller, see figure 16.1). This discrepancy can also be seen in figures 13.5 & 16.1, where the heights of the Gaussian (and therefore, the area under the curves) are very different. Unfortunately, the differences in the manual & automatic Gaussian fitting meant that even though both processes fit near the correct location, the ratio test gave different shapes and ratios dependent on the process. This made applying observations from one process to another inaccurate, and so this method was ultimately not used.

If the integral in the automatic process took limits of FWHM/2 either side of the fitted  $\nu_{\max}$ , this would have been more accurate- this would have ignored the noisy effect of granulation acting at lower frequencies (preventing the ratio being offset by the granulation section’s area). Ensuring the limits of the integration in the automatic process are the same and using flattened data for this test, would ensure

consistency between the two Gaussian fitting processes. Consequently, the values gained in the manual process are more likely to be lower this way, so much more comparable to the automatic process. As mentioned before, an element of human error exists in the manual fitting as the p-modes were not always clear when trying to find the cutoff value - especially at higher shot noises. If these issues were rectified in these ways, this test would likely give usable results (this is also detailed in the Power Ratio Evaluation section above).

## 16.6 $\Delta\nu$ Results (K-Ryan)

As only one star was categorised as a ‘possible detection’, this was the only star to undergo further analysis and have its  $\Delta\nu$  measured. Figure 16.4 is the Frequency-Power Spectrum analysed to measure  $\Delta\nu$  with the results in Table 14.

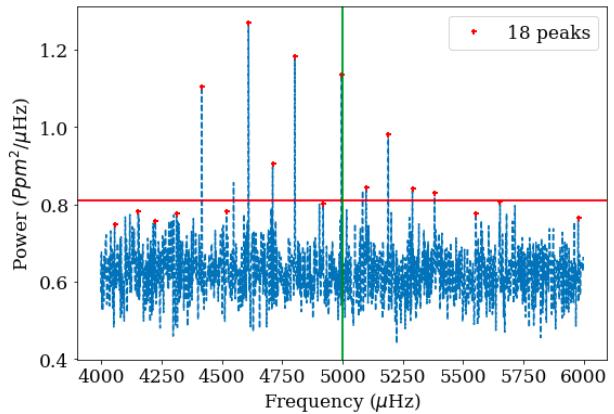


Figure 16.4: Final Power-Frequency spectra for Star 1 of the K dwarf catalogue with  $\Delta\nu = 176.38 \pm 9.12 \mu\text{Hz}$ .

Star	$m_v$	$\nu_{\max}(\mu Hz)$	$\Delta\nu(\mu Hz)$	$\pm\sigma_{\Delta\nu}(\mu Hz)$	#Peaks	$\pm\sigma_{\text{Peaks}}$
1	7.32	5423.60	176.38	9.12	18.00	2.13
Kepler-444	8.86	4538.00	179.64	0.76	—	—

Table 14: Large Separation for K dwarfs and Kepler-444. Noise was added 100 times and in each iteration, the number of peaks were found and averaged over using a set threshold=0.375 introducing an error on the mean number of peaks found (rounded up). Kepler-444 data available from [SIMBAD](#).

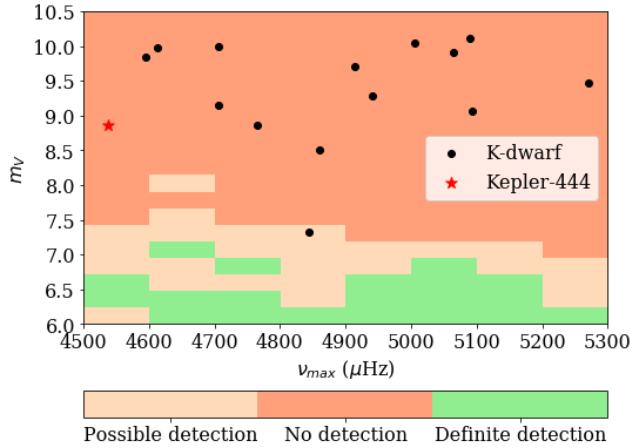


Figure 16.5: Background colour represents whether the star corresponding star to that location could be detected, the black dots represent K dwarfs and the red star represents Kepler-444.

In a recent paper studying an Extrasolar System [T. L. Campante et al. \(2015\)](#), measurements for a Sun-like star, Kepler-444, were made and these can also be seen in Table 14. Kepler-444 is a star of mass  $0.754 \pm 0.03 M_\odot$  [Bulgen. G. et al. \(2019\)](#) and is of interest as, to date, this is one of the only K dwarfs to have these parameters measured and has five exoplanets orbiting in its solar extrasolar system. Figure 16.5 is a variation on figure 12.11 but with Kepler-444 overlaid in red.

Upon initial comparison, the two stars differ somewhat drastically in terms of measured  $\nu_{\max}$  and this could be down to experimental error or due to the specific classification of star. Kepler-

444 is a confirmed KI star (the upper edge of this spectral class) and Star 1 has no specific classification and as such, could be any K type. Moreover, the experimental error could arise from the problem encountered earlier where the curve\_fit is susceptible to the initial guesses meaning each iteration with a new initial guess will differ. In terms of  $\Delta\nu$ , both stars appear to have similar values and again the difference in the values will come from the various steps taken to produce these values. The established relationship seen in equation 5.8 states that for stars less massive than the Sun, which both K dwarfs are,  $\Delta\nu \propto \frac{M}{M_\odot}^{1/2} \frac{R}{R_\odot}^{-3/2}$ ; thus, relative to the Sun,  $\Delta\nu$  should be higher and this is observed. This prediction can be seen in both stars as  $\Delta\nu$  for both stars are higher than the  $135 - 136 \mu Hz$  for the Sun. A point not to overlook is that Star 1 represents a star with detectable p-modes due to its apparent  $m_v$  being lower than Kepler-444 and because Star 1 was in the ‘possible detection’ category, K (and possibly M) dwarf stars with a  $m_v < 8.86$  will have detectable p-modes; this is purely since that Kepler-444 has already been detected and had both  $\nu_{\max}$  and  $\Delta\nu$  measured.

## 16.7 Overall Findings (K-Ryan and San)

Overall, it is seen that K & M dwarfs will be difficult to detect based on the simulated models used; each detection threshold returns a small proportion of stars compared to the 80 stars

used for the analysis. The  $1\sigma$  test returned two stars (K dwarfs 1 & 2) with potential for a few more dependent on further testing with the other thresholds, however, this method is not sensitive to the fainter K dwarfs and does not work well for the M dwarfs. Stronger background noise presents an increasing difficulty in fitting the Gaussian on the p-modes correctly and eventually prevents the ‘curve\_fit’ function from fitting any Gaussian for the fainter stars. If used on PLATO, this method should only be used for brighter K dwarfs with some knowledge of the real  $\nu_{\max}$  values. This value could be found via literature of similar stars, or if mass & radius values for the target star are known, equation 5.8 could be utilised for a reasonable estimate.

The grid method for determining detectability of a star was unique and produced results that can be verified through the other methods. Unfortunately, there happened to only be one star (Star 1) that was deemed a ‘possible detection’, nevertheless, this method was able to categorise stars in such a way that, if more stars were available for analysis, they could efficiently be placed on this grid and it would be instantly known whether they were detectable or not. To optimise this method, more stars of lower  $m_v$  (more luminous stars) were needed to both populate the grid as a whole and also extend the ‘definite detection’ area for a more complete picture to be obtained; if there were more luminous stars, the area the ‘definite detection’ region spanned would be larger and could extend the ‘possible detection’ region to Kepler-444.

This method did return Star 1 as being a ‘possible detection’ and was fed through for further analysis with the final results being  $\nu_{\max} = 5423.60 \mu\text{Hz}$  and  $\Delta\nu = 176.38 \pm 9.12 \mu\text{Hz}$ . As a result of relationships calculated using the simulated data (figures 6.17 and A8), the surface temperature and surface gravity

were calculated, from  $\nu_{\max}$ , to be 4754.54 K and  $42000\text{cm/s}^2$  respectively. Having the ability to determine these values stems from the scaling relations and established relationships, however, the vast difference in the calculated values of this K dwarf and the Sun highlight that basing the physics of the K spectral class off of the Sun is not sufficient. To gain a full appreciation of these spectral classes, more M and K dwarfs need to be observed and have their properties measured. Star 1 was on the border of ‘possible detection’ and ‘no detection’ thus it stands to reason stars with  $m_v < 7.319$  would have measurable p-modes.

The power integration was to serve as a final check of detectability for stars filtered as a ‘positive detection’ from the other two methods. The cutoff ratio, which shows when a star’s p-modes become undetectable in observations, was to provide information on the detectability limits for stars with particular characteristics; this detail would be useful for observations of similar stars. Unfortunately, this method did not work in the way desired due to the differences in the automatic and manual fitting process. It is suggested that employing a consistent set of integral limits and using flattened data for both processes would allow for usable results. But it has to be stressed that this method requires some list of possible detections to work with initially, as using it as the first method would be inefficient and human error is introduced.

## 17 Discussion of Results (Luke)

The results produced by the Data Analysis group show that only a single detection of p-mode oscillations could possibly be made out of the entire population of 80 simulated K and M dwarf stars, comprised of the brightest 40 of each type. It is likely that these results

are somewhat conservative, and are the lower bound to what might be expected from PLATO observation since the grid of detectability produced based on the plateau code (Figure 16.5) places Kepler-444 within the zone where detections cannot be made.

This statement is based on the fact that it is known from literature [T. L. Campante et al. \(2015\)](#) that photometric measurements of the K dwarf Kepler-444 revealed the presence of p-mode oscillations, meaning that oscillations are detectable in this star.

The methods used to detect the p-mode oscillations in Kepler-444 differ to that used to create figure 16.5 which places the star in a region of no detection, so the analysis may be accurate for the method used. In this case, it would be necessary to further examine both the accuracy of the Gaussian fitting and the subjective nature of the plateau detection code used to create the grid. This would show if either detection method is lacking a fundamental feature that might lead to the discrepancy, or that the bounds used within the detection criteria of the models are too restrictive.

Exploring these concepts further, specific attention shall be directed towards the limitations of the Gaussian fitting method of detections, since this was the primary method and that which formed the basis for both the 1 sigma detection test and the plateau detection code. This method also had an impact on the results obtained from the automated power-integration method used towards the end of the project, the results of which can be seen in the footnote of figure 16.1.

The bias of the Gaussian fitting code (and of curve \_ fit as a function) to fit the immediate region around the initial guess parameters has been discussed in earlier sections. This formed the basis of the motivation for developing the

plateau code and the ‘Grid Stars’ matrix since it was thought that this would act to remove this bias by iterating over a number of initial guess centres. However, this was not the only factor limiting the effectiveness of this detection method.

It was noticed whilst running later iterations that all of the Gaussian curves fitted to the data had very small amplitudes of the order of  $\sim 0.05 \text{ ppm}^2 \mu\text{Hz}^{-1}$  or less, despite an initial guess amplitude of order  $1 \text{ ppm}^2 \mu\text{Hz}^{-1}$ . This fitted factor did not appear to change with any sort of observable significance even when changing the initial guess amplitude. This meant that the fitted Gaussians were not fitting directly to the overlying shape of the combined p-mode oscillation envelope as per the design of the code. This is evidenced in the lower two panels of figure 16.1 where the p-modes can be identified by-eye, and it can be seen that the red fitted Gaussian does not trace this outline. This amplitude factor also limited the capability to utilise power-integration in an automated fashion for the simulated stars. Since the fitted Gaussian failed to encompass the oscillations, the integrated power under it was, therefore, an incorrect representation, and as such, false limits of detection were obtained.

Since the method used to detect p-mode oscillations and extract the global parameters  $\Delta\nu$  and  $\nu_{\max}$  in Kepler-444 was “above average power spectrum” analysis ([T. L. Campante et al. \(2015\)](#)), it may be possible to use raw Kepler-444 data to troubleshoot the issues with both the Gaussian fitting and the related power-integration detection thresholds. This is one of many tests and further work that could be carried out to improve the detection methods for ultimate use in refining the detection limits of the PLATO telescope.

## 17.1 Implications for the ESA PLATO Mission

From the results of the plateau detection method on a series of fake ‘grid stars’ and the associated detection grid produced, it was concluded that making asteroseismic measurements of M and K dwarf stars with PLATO will unlikely be possible for the majority of targets. Analysis of the brightest simulated stars using by-eye and 1 sigma detection methods support this conclusion. Without further research and development into the detection techniques of the oscillations present in these stars, it is recommended that observation targets are revised by PLATO based on this conclusion.

One of the potential impacts this will have on the PLATO mission is on the ability to constrain the parameters of, and hence classify, any exoplanets discovered around host stars of these classes. Since PLATO will take photometric measurements of stars, any exoplanets discovered would be detected via periodic dimming of the star, caused by a planet crossing the line-of-sight between the observer and the star and blocking out a small fraction of the light. This is known as a planetary transit, and to date is the method that has yielded the largest number of confirmed exoplanet detections.

As the planet is not observed directly, its physical parameters, including its radius, orbital separation and later its density are largely calculated using known values of the stellar parameters. Therefore, more accurate knowledge of the physical stellar parameters, found from constraints obtained via asteroseismology, translates directly to a more accurate view of the planetary system.

A direct consequence of the inability to observe p-mode oscillations in the spectra of the host stars of these systems will fail to tightly con-

strain the value and associated uncertainties of these aforementioned parameters of newly discovered planets. This is further compounded by the expectation of detecting more planets around these dim stars, due to the intrinsic bias of the transit method, which produces a comparatively larger signal for the same size planet in the same location for a small dim dwarf star when compared to that of a larger, more luminous and massive star further up the main-sequence.

## 17.2 Other Considerations

When developing the models of the M and K dwarf stars and designing the detection methods that would be used to find these from noisy realistic data, some factors were not considered.

The first of these was rotation, a factor especially relevant in binary stars as the mutual gravitational attraction to one another can result in the “spin up” of one of the stars through the transferal of angular momentum. Since it is estimated that approximately 80% or more of stars are thought to be in binary or some other multiple-star system ([CSIRO \(2020\)](#) : [Space.com \(2018\)](#)), it is possible that this factor could affect the data collected by PLATO. However, such a transferal of angular momentum usually only occurs via the accretion of material from a smaller companion star onto the stellar remnant of the larger star, meaning that for the type of systems being studied, this effect is likely not of significant relevance. Further impacts on the visibility of oscillatory modes caused by binary star systems are discussed in Section 3.3.1 with the conclusion being that for the dwarf stars investigated by this project, the target being part of a binary system is unlikely to be detrimental.

As discussed briefly in Section 5.6, intrinsic stellar rotation can act to either enhance or reduce the visibility of the p-mode oscillations present

in a star, depending on the orientation of the rotation axis relative to the observer. Provided the splitting of individual modes could be also observed in the data, then there would exist the potential to reconstruct the rotation axis of the target via modelling. In turn, assuming favourable evolutionary conditions, this could provide an insight into the probability of making exoplanet detections using the transit method since the orbital plane any planets in the system should lie within would be found perpendicular to this stellar rotation axis.

One other factor not considered in the creation of the oscillation spectra or the subsequent analysis was the effect of a magnetic field in this type of star. This causes an increased level of activity in the star, which manifests in two ways. The first of these is an increased rate of granulation, leading to an increased number of features on the stellar surface such as star-spots. Although these features cannot be observed directly, they cause variations in the total flux received on the timescale of stellar rotation, since the features rotate with the stellar surface. As such modelling this effect would need to be done in conjunction with the modelling of stellar rotation. Since the period of rotation for stars is commonly of the order of days, with the exception of compact stellar remnants such as pulsars, this variation mostly occurs at the low frequency end of the spectrum. This elevates and changes the shape of the background curve, along with the total power contained within the spectrum. However, the effect has the same characteristic Lorentzian shape as that of the granulation modelled by the TSM group, meaning that it tails off and bleeds into the high frequency end of the spectrum. Therefore there would be some impact on the detectability of the p-modes, though successful background fitting such as that achieved by the DA group may act to largely mitigate this factor.

The other effect seen due to the presence of a

magnetic field is more frequent flaring events. Although the true effect of such events on a frequency-power spectrum is largely unknown, it is theorised that a change in the signal-to-noise level would occur at high frequencies where the oscillations in this class of star are known to lie. As such, during periods of flaring, detectability of such p-mode oscillations would be significantly impaired.

The obvious factor concerning overall detectability is therefore the frequency at which such flaring events occur. Whilst larger flares are still thought to be *kyr* events, M and K dwarfs stars typically show a higher rate of solar-energy flaring than that observed in Sun-like stars. Given that the Sun typically exhibits a few flares per day [Engvold \(2018\)](#) this effect would be common enough affect to have an impact on the photometric data collected by PLATO, which would likely translate to an increase in noise level. This would reduce the visibility of the oscillations and make detecting them in the data significantly more challenging.

Further compounding p-mode detectability is the one of the leading theories in asteroseismology and stellar modelling regarding the impacts magnetic fields have on the propagation of p-mode oscillations. In a paper titled "Evidence for the impact of stellar activity on the detectability of solar-like oscillations observed by Kepler" [Chaplin \(2011\)](#) a discussion is made that outlines how the amplitudes of the p-mode oscillations, which are driven by turbulence in near-surface convection layers are reduced due to a reduction in the velocity of the material within the convective cells. Additional attention is drawn to the affect magnetic structures in the solar photosphere have on the dynamics and propagation of the oscillations, acting to dampen them, with the link being made that between solar minimum and maximum there is a 12.5% reduction in the amplitude of the observable modes. This is then discussed in the

context of other stars based on Kepler data, where similar effects are observed.

## Summary (Luke)

It is recognised that the models created by the TSM group present best-case detection limits using the methods devised by the DA group. This showed that p-mode oscillations could be detected down to a visual magnitude of  $m_v \sim 7.25$  roughly across the full frequency range spanning  $4400 \mu\text{Hz}$  to  $5300 \mu\text{Hz}$ , with some zones of uncertainty present within these limits. Further investigation is needed to determine if these limits can be extended any further using the 'plateau' detection method chosen.

However, addition of the factors outlined in Section 17.2 above to the models to create more accurate simulations of real stars would act to both reduce the amplitude of the oscillations and increase the noise present in the frequency-power spectrum. This would reduce the visibility of the oscillation modes present in the stars, and would result in lower detection limits, particularly pertaining to the magnitude limit, for PLATO observations.

# Appendix

## A Derivation of convective instability (Sharif)

In the following section, all stellar parameters are functions of the stellar radius,  $r$ , i.e  $T = T(r)$ .

Consider a bubble of stellar plasma which is randomly perturbed such that it begins to rise. The buoyant forces per unit volume at any point along its path can be found by considering the difference between its density  $\rho_b$  and the density of the surrounding stellar material  $\rho_s$ .

$$f = -g(\rho_b - \rho_s) = -g\Delta\rho. \quad (.1)$$

If  $f < 0$ , the bubble will sink back to its equilibrium position and the region can be deemed convectively stable. However, if  $f > 0$ , the bubble will rise and if it continues to do so, the region is known as being convectively unstable. Whether it remains less dense than its surroundings as it rises can be determined by assuming that the plasma behaves adiabatically and as an ideal gas. Adiabatic assumptions are valid because the bubble rises on a timescale much shorter than the timescale for heat exchange [Pols \(2011\)](#). Considering the convectively unstable situation where  $f > 0$ , it must be true that  $\Delta\rho < 0$

$$\begin{aligned} \Delta\rho &= \left[ \frac{d\rho_b}{dr} - \frac{d\rho_s}{dr} \right] dr \\ &\simeq \left[ \frac{d\rho_b}{dr}_{ad} - \frac{d\rho_s}{dr} \right] \Delta r < 0 \end{aligned} \quad (.2)$$

$$\frac{d\rho}{dr}_{ad} < \frac{d\rho}{dr} \quad (.3)$$

where the subscripts  $b$  and  $s$  have been dropped and the derivative representing the change within the bubble is indicated with the subscript "ad" indicating its adiabatic behaviour. This is then linked to the temperature gradient by making the following substitutions.

The adiabatic term on the LHS of equation [.3](#) will be replaced with the use of the adiabatic exponent,  $\Gamma_1$

$$\Gamma_1 = \left( \frac{d \ln P_b}{d \ln \rho_b} \right) \rightarrow \frac{d\rho}{dr}_{ad} = \frac{\rho}{P} \frac{1}{\Gamma_1} \frac{dP}{dr} \quad (.4)$$

where  $P$  is the pressure. The term on the RHS of equation [.3](#) will be replaced by assuming that the plasma acts as an ideal gas and, therefore,  $\rho$  is a function of  $P$  and  $T$  only.

$$\rho(P(r), T(r)) = \frac{\mu m_u P}{k_b T} \quad (.5)$$

where  $\mu$  is the mean molecular weight,  $m_u$  is the atomic mass unit,  $k_b$  is the Boltzmann constant and  $T$  is the temperature. This can be expressed as a sum of derivatives with respect to radius.

$$\frac{1}{\rho} \frac{d\rho}{dr} = \frac{1}{P} \frac{dP}{dr} - \frac{1}{T} \frac{dT}{dr}. \quad (.6)$$

Substituting these equations into equation [.3](#) gives

$$\frac{\rho}{P} \frac{1}{\Gamma_1} \frac{dP}{dr} < \frac{\rho}{P} \frac{dP}{dr} - \frac{\rho}{T} \frac{dT}{dr} \quad (.7)$$

which rearranges to

$$\frac{\Gamma_1 - 1}{\Gamma_1} \frac{\rho}{P} \frac{dP}{dr} < \frac{\rho}{T} \frac{dT}{dr}. \quad (.8)$$

Due to the partial ionisation and subsequent departure from ideal conditions,  $\Gamma_1$  must be replaced by  $\Gamma_2$  [Christensen-Dalsgaard \(2008\)](#). This represents the fact that the bubble is not truly adiabatic with respect to its surroundings, however, it still loses energy at a relatively slow rate, see figure [A1](#) for a graphical representation of this situation.

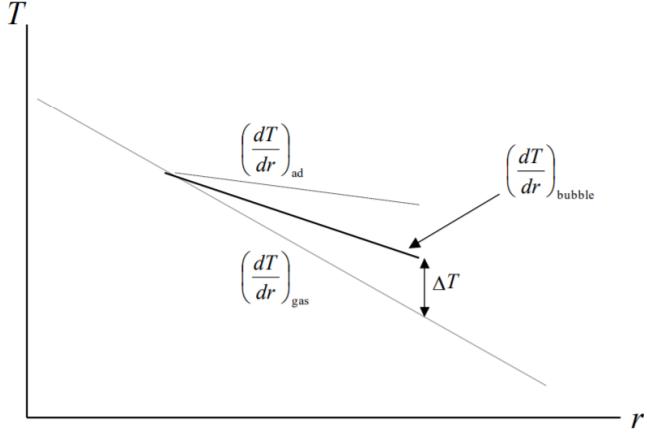


Figure A1: Temperature gradients showing how a bubble of gas cools at a slower rate than the passing stellar material. Consequently, the bubble has a temperature excess  $\Delta T$  causing it to continue to rise. [Chaplin \(2019\)](#)

$\Gamma_2$  is given by

$$\frac{\Gamma_2}{\Gamma_2 - 1} = \frac{d \ln P}{d \ln T_{\text{ad}}} = \nabla_{\text{ad}} \quad (9)$$

which, when substituting into equation .8 leads to

$$\frac{d \ln T}{d \ln P_{\text{ad}}} \frac{\rho}{P} \frac{dP}{dr} < \frac{\rho}{T} \frac{dT}{dr}. \quad (10)$$

Rearranging and taking the modulus of both sides gives

$$\left| \frac{d \ln T}{d \ln P} \right| > \left| \frac{d \ln T}{d \ln P_{\text{ad}}} \right| \quad (11)$$

or, more compactly,

$$\nabla > \nabla_{\text{ad}}. \quad (12)$$

## B Additional figures

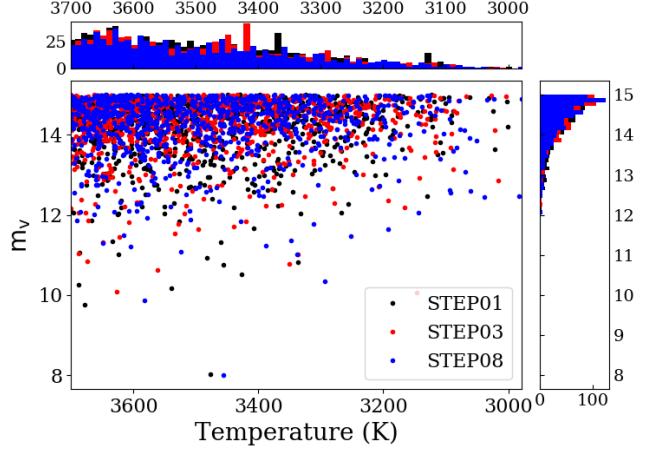


Figure A2: Effective Temperature of the M dwarfs in STEP01, 03, & 08 against their apparent visual magnitudes. Histograms showing the distribution of apparent magnitudes and Temperatures show a sharp exponential decrease in the number of targets at decreasing magnitudes and a modal effective temperature of  $\sim 3630$  K.

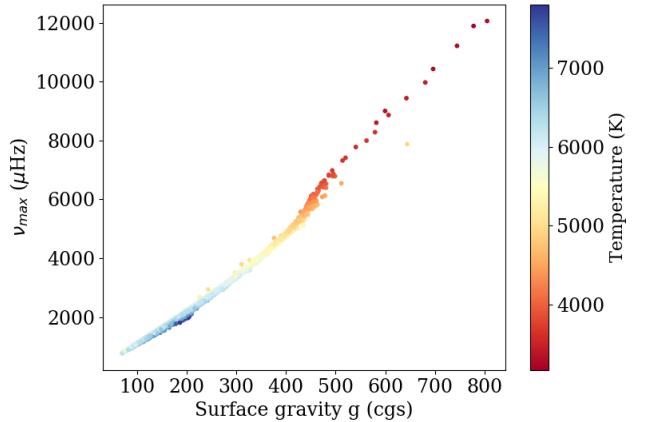


Figure A3: The frequency of maximum oscillations power as a surface gravity and temperature (TRILEGAL data). This spans across a wider range of spectral classes.

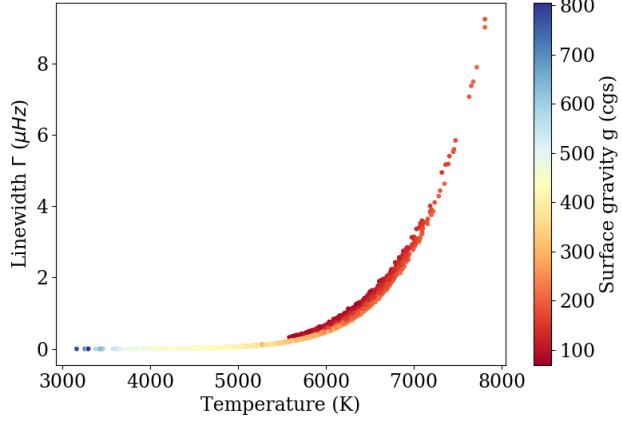


Figure A4: The linewidth as a function of temperature and surface gravity, following equation 5.29 (TRILEGAL data).

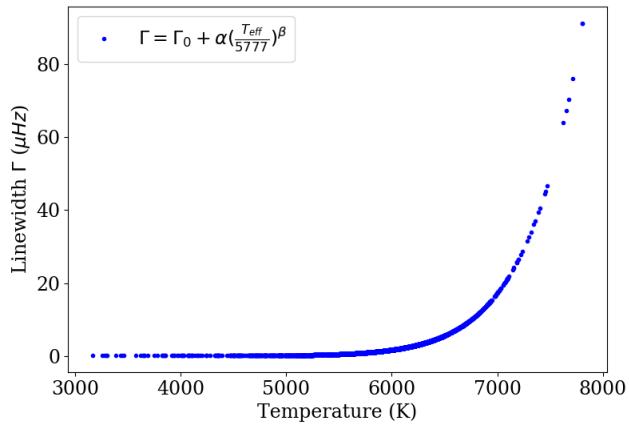


Figure A5: The linewidth as a function of temperature following equation 5.32 (TRILEGAL data).

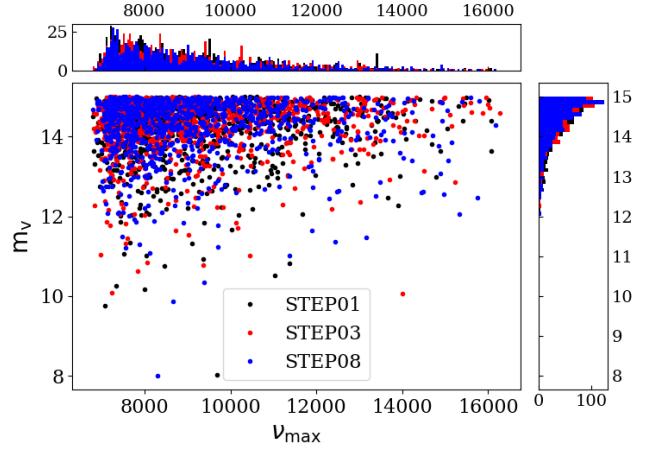


Figure A6: Graph showing the distribution of  $\nu_{\max}$  of M dwarfs in STEP01, 03, & 08 against their apparent visual magnitudes. Histograms showing the distribution of apparent magnitudes and  $\nu_{\max}$  show a sharp exponential decrease in the number of targets at decreasing magnitudes and a modal  $\nu_{\max}$  of  $\sim 7200 \mu\text{Hz}$ .

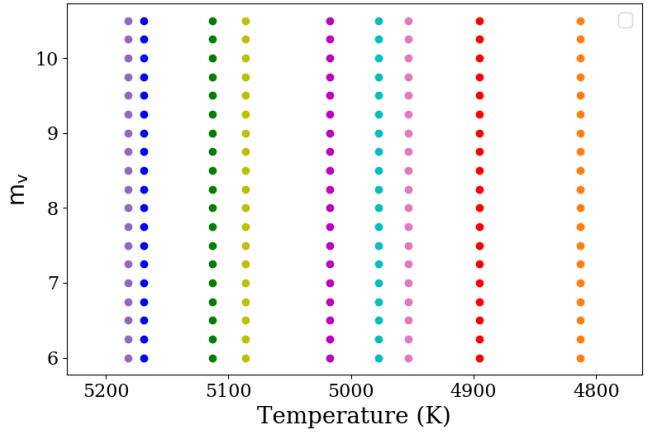


Figure A7: Temperature-Apparent Magnitude Graph showing the targets artificially created in Figure 6.16.

Figure A7 shows the temperature distribution of the targets chosen in Figure 6.16. Although the (pink) target with  $\nu_{\max} \approx 5200$  appears to be an anomaly with a greater temperature than expected, the rest of the grid appears in the same order, albeit at a varying distance apart. This can be compared to the "actual" distribu-

tion of temperatures and apparent magnitudes in Figure 3.3

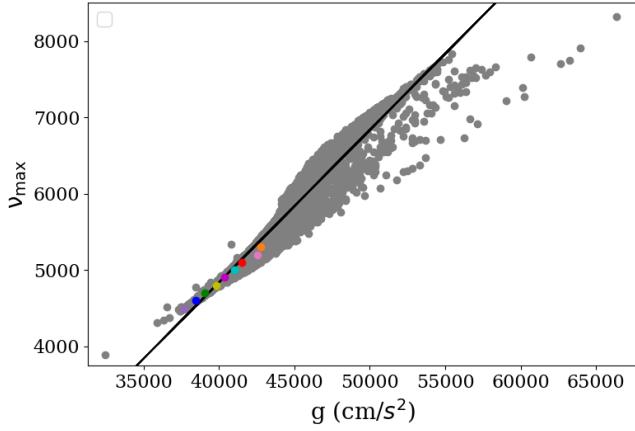


Figure A8: The surface gravities,  $g$ , and  $\nu_{\text{max}}$  of the stars chosen for the final artificial grid of K dwarf targets, compared to that of the K dwarfs identified in the STEP01 simulation from TRILEGAL. The line of best fit ( $\nu_{\text{max}} = 0.19979 g - 3154.4$  (to 5 s.f.)) is shown in black, and can be used to estimate the limits of detection in surface gravity from the limit on  $\nu_{\text{max}}$

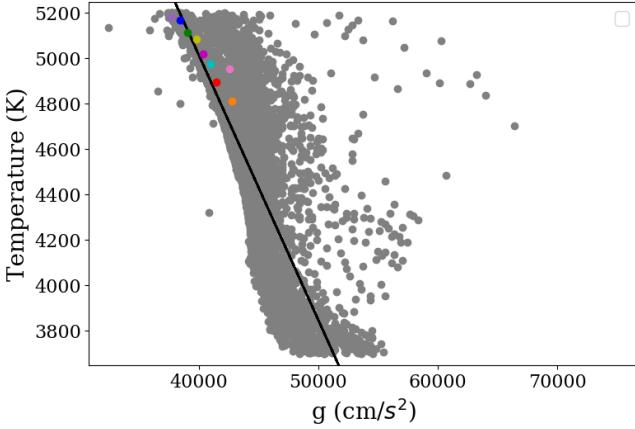


Figure A9: The surface gravities,  $g$ , and effective temperatures of the stars chosen for the final artificial grid of K dwarf targets, compared to that of the K dwarfs identified in the STEP01 simulation from TRILEGAL. The line of best fit ( $T_{\text{eff}} = -0.11677 g + 9686.4$  (to 5 s.f.)) is shown in black, displaying the inverse relationship between the two properties.

Star Number	$\nu_{\text{max}} (\mu\text{Hz})$
6	4034
7	4034
25	4119
26	4193
40	4193
42	4193
43	4193
44	4193
45	4304
58	4304
61	4304
62	4304
63	4304
64	4387
77	4387
82	4387
83	4500
101	4500
102	4500
103	4500
104	4600
115	4600
118	4600
119	4600
120	4700
136	4700
137	4700
139	4700
140	4700
142	4800
155	4800
157	4800
158	4900
173	4900
174	4900
175	4900
176	4900
177	5000
194	5000
195	5100
214	5200
232	5200
233	5300
249	5300
250	5300

Table 15: A table showing the list of star numbers in the ‘maybe’ category and their corresponding  $\nu_{\text{max}}$  values.

# References

Conny Aerts, Jørgen Christensen-Dalsgaard, and Donald W Kurtz. *Asteroseismology*. Springer Science & Business Media, 2010.

Víctor Silva Aguirre, Mikkel N Lund, HM Antia, Warrick H Ball, Sarbani Basu, Jørgen Christensen-Dalsgaard, Yveline Lebreton, Daniel R Reese, Kuldeep Verma, Luca Casagrande, et al. Standing on the shoulders of dwarfs: The kepler asteroseismic legacy sample. ii. radii, masses, and ages. *The Astrophysical Journal*, 835(2):173, 2017.

Edwin R. Anderson, Jr. Duvall, Thomas L., and Stuart M. Jefferies. Modeling of Solar Oscillation Power Spectra. *The Astrophysical Journal*, 364:699, December 1990. doi: 10.1086/169452.

T Appourchaux, Othman Benomar, Michael Gruenbergauer, WJ Chaplin, RA García, R Handberg, GA Verner, HM Antia, TL Campante, GR Davies, et al. Oscillation mode linewidths of main-sequence and subgiant stars observed by kepler. *Astronomy & Astrophysics*, 537:A134, 2012.

T. Appourchaux, H. M. Antia, O. Benomar, T. L. Campante, G. R. Davies, R. Handberg, R. Howe, C. Régulo, K. Belkacem, G. Houdek, R. A. García, and W. J. Chaplin. Oscillation mode linewidths and heights of 23 main-sequence stars observed by Kepler. *Astronomy and Astrophysics*, 566, June 2014.

M Auvergne, P Bodin, L Boisnard, J-T Buey, S Chaintréuil, G Epstein, M Jouret, T Lam-Trong, P Levacher, A Magnan, et al. The corot satellite in flight: description and performance. *Astronomy & Astrophysics*, 500(1):411–424, 2009.

Warrick H. Ball, William J. Chaplin, Mathew Schofield, Andrea Miglio, Diego Bossini, Guy R. Davies, and Léo Girardi. A Synthetic Sample of Short-cadence Solar-like Oscillators for TESS. *The Astrophysical Journal Supplement*, 239(2):34, December 2018. doi: 10.3847/1538-4365/aaedbc.

Warrick H Ball, William J Chaplin, Mathew Schofield, Andrea Miglio, Diego Bossini, Guy R Davies, and Léo Girardi. A synthetic sample of short-cadence solar-like oscillators for tess. *The Astrophysical Journal Supplement Series*, 239(2):34, 2018.

Isabelle Baraffe and Gilles Chabrier. Mass-spectral class relationship for m dwarfs. *The Astrophysical Journal Letters*, 461(1):L51, 1996.

Thomas Barclay. February 2020. URL <https://heasarc.gsfc.nasa.gov/docs/tess/operations.html>.

Sarbani Basu and William J Chaplin. *Asteroseismic Data Analysis: Foundations and Techniques*, volume 4. Princeton University Press, 2017.

George Keith Batchelor. *The theory of homogeneous turbulence*. Cambridge University press, 1953.

Timothy R Bedding and Hans Kjeldsen. Scaled oscillation frequencies and échelle diagrams as a tool for comparative asteroseismology. *arXiv preprint arXiv:1001.5038*, 2010.

CWJ Beenakker and Henning Schomerus. Counting statistics of photons produced by electronic shot noise. *Physical review letters*, 86(4):700, 2001.

K Belkacem, Marc-Antoine Dupret, F Baudin, T Appourchaux, JP Marques, and R Samadi. Damping rates of solar-like oscillations across the hr diagram-theoretical calculations confronted to corot and kepler observations. *Astronomy & Astrophysics*, 540:L7, 2012.

Earl Patrick Bellinger. A seismic scaling relation for stellar age. *Monthly Notices of the Royal Astronomical Society*, 486(4):4612–4621, 2019.

William J Borucki, David Koch, Gibor Basri, Natalie Batalha, Timothy Brown, Douglas Caldwell, John Caldwell, Jørgen Christensen-Dalsgaard, William D Cochran, Edna DeVore, et al. Kepler planet-detection mission: introduction and first results. *Science*, 327(5968):977–980, 2010.

Brainmap.org. *Statistics & Formulas*. <http://brainmap.org/meta/stats.html>, Last accessed: Mar 2020.

Timothy M Brown, Ronald L Gilliland, Robert W Noyes, and Lawrence W Ramsey. Detection of possible p-mode oscillations on procyon. *The Astrophysical Journal*, 368:599–609, 1991.

Buldgen. G. et al. Revisiting kepler-444 - i. seismic modeling and inversions of stellar structure. *A&A*, 630:126, 2019. doi: 10.1051/0004-6361/201936126. URL <https://doi.org/10.1051/0004-6361/201936126>. Remaining Authors : Farnir. M. and Pezzotti. C. and Eggenberger. P. and Salmon. S. J. A. J. and Montalban. J. and Ferguson. J. W. and Khan. S. and Bourrier. V. and Rendle. B. M. and Meynet. G. and Miglio. A. and Noels. A.

C Catala. Plato definition study report. *ESA/SRE*, 13:2011, 2011.

Bedding T.R.-et al Chaplin, W.J. Evidence for the impact of stellar activity on the detectability of solar-like oscillations observed by kepler. *The Astrophysical Journal Letters*, 2011.

William Chaplin. 3. observational signatures of convection, 2019. URL <https://canvas.bham.ac.uk/courses/34222/pages/AEXPart1:Asteroseismology>.

William J Chaplin and Andrea Miglio. Asteroseismology of solar-type and red-giant stars. *Annual Review of Astronomy and Astrophysics*, 51:353–392, 2013.

William J Chaplin, Hans Kjeldsen, Timothy R Bedding, Jørgen Christensen-Dalsgaard, Ronald L Gilliland, Stephen D Kawaler, Thierry Appourchaux, Yvonne Elsworth, RA García, Günter Houdek, et al. Predicting the detectability of oscillations in solar-type stars observed by kepler. *The Astrophysical Journal*, 732(1):54, 2011.

WJ Chaplin, Y Elsworth, GR Isaak, CP McLeod, BA Miller, and R New. Solar p-mode linewidths from recent bison helioseismological data. *Monthly Notices of the Royal Astronomical Society*, 288(3):623–626, 1997.

WJ Chaplin, G Houdek, Christoffer Karoff, Y Elsworth, and R New. Mode lifetimes of stellar oscillations—implications for asteroseismology. *Astronomy & Astrophysics*, 500(2):L21–L24, 2009.

WJ Chaplin, R Sanchis-Ojeda, TL Campante, R Handberg, D Stello, Joshua Nathan Winn, S Basu, Jørgen Christensen-Dalsgaard, GR Davies, TS Metcalfe, et al. Asteroseismic determination of obliquities of the exoplanet systems kepler-50 and kepler-65. *The Astrophysical Journal*, 766(2):101, 2013.

Michael Chrisp, Kristin Clark, Brian Primeau, Michael Dalpiaz, and Joseph Lennon. Optical design of the camera for transiting exoplanet survey satellite (tess). In *UV/Optical/IR Space Telescopes and Instruments: Innovative Technologies and Concepts VII*, volume 9602, page 96020C. International Society for Optics and Photonics, 2015.

J. Christensen-Dalsgaard. Solar models, 1996. URL [https://users-phys.au.dk/~jcd/solar\\_models/](https://users-phys.au.dk/~jcd/solar_models/).

Jørgen Christensen-Dalsgaard. Lecture notes on stellar oscillations. 1997.

Jørgen Christensen-Dalsgaard, Timothy R Bedding, and Hans Kjeldsen. Modeling solar-like oscillations in eta bootis. *The Astrophysical Journal*, 443:L29–L32, 1995.

Jørgen Christensen-Dalsgaard, Hans Kjeldsen, TM Brown, Ronald Lynn Gilliland, Torben Armentoft, Søren Frandsen, P-O Quirion, WJ Borucki, D Koch, and JM Jenkins. Asteroseismic investigation of known planet hosts in the kepler field. *The Astrophysical Journal Letters*, 713(2):L164, 2010.

Jørgen Christensen-Dalsgaard. Stellar structure and evolution, Mar 2008. URL [file:///C:/Users/shari/OneDrive/Documents/Work/3rdYr/LifeandDeathofStars/LN\\_stellar\\_structure.pdf](file:///C:/Users/shari/OneDrive/Documents/Work/3rdYr/LifeandDeathofStars/LN_stellar_structure.pdf).

CSIRO. Introduction to binary stars, 2020. URL [https://www.atnf.csiro.au/outreach/education/senior/astrophysics/binary\\_intro.html](https://www.atnf.csiro.au/outreach/education/senior/astrophysics/binary_intro.html).

M Cuntz and EF Guinan. About exobiology: the case for dwarf k stars. *The Astrophysical Journal*, 827(1):79, 2016.

Vial J.-C. & Skumanich A. Engvold, O. *The Sun as a Guide to Stellar Physics, Chpt 11*. Elsevier, 2018.

ESA. Plato definition study report (esa-sci(2017)1). 2017. URL <https://sci.esa.int/s/8rPyPew>.

Bernd Freytag. Convection dans les étoiles géantes, 2004.

Ronald L Gilliland, Jon M Jenkins, William J Borucki, Stephen T Bryson, Douglas A Caldwell, Bruce D Clarke, Jessie L Dotson, Michael R Haas, Jennifer Hall, Todd Klaus, et al. Initial characteristics of kepler short cadence data. *The Astrophysical Journal Letters*, 713(2):L160, 2010.

L Girardi, MAT Groenewegen, E Hatziminaoglou, and L Da Costa. Star counts in the galaxy-simulating from very deep to very shallow photometric surveys with the trilegal code. *Astronomy & Astrophysics*, 436(3):895–915, 2005.

Douglas Gough. On the principal asteroseismic diagnostic signatures. In *Asteroseismology Across the HR Diagram*, pages 165–185. Springer, 2003.

GMHJ Habets and JRW Heintze. Empirical bolometric corrections for the main-sequence. *Astronomy and Astrophysics Supplement Series*, 46:193–237, 1981.

Ana M Heras, Heike Rauer, and Don Pollacco. The plato space mission: Revealing habitable worlds. In *42nd COSPAR Scientific Assembly*, volume 42, 2018.

- Michele Johnson. August 2017. URL [https://www.nasa.gov/mission\\_pages/kepler/spacecraft/index.html](https://www.nasa.gov/mission_pages/kepler/spacecraft/index.html).
- Caitlin Dawn Jones. *Stellar variance for asteroseismic parameter estimation and inferences on the evolutionary state and binary population of red giant stars*. PhD thesis, University of Birmingham, 2018.
- Christoffer Karoff. Observational asteroseismology, May 2008. URL [https://phys.au.dk/fileadmin/site\\_files/publikationer/phd/Christoffer\\_Karoff.pdf](https://phys.au.dk/fileadmin/site_files/publikationer/phd/Christoffer_Karoff.pdf).
- Christoffer Karoff and Hans Kjeldsen. Evidence that solar flares drive global oscillations in the sun. *The Astrophysical Journal Letters*, 678(1):L73, 2008.
- Hans Kjeldsen, Timothy R Bedding, Torben Arentoft, R Paul Butler, Thomas H Dall, Christoffer Karoff, László L Kiss, CG Tinney, and William J Chaplin. The amplitude of solar oscillations using stellar techniques. *The Astrophysical Journal*, 682(2):1370, 2008a.
- Hans Kjeldsen, Timothy R Bedding, and Jørgen Christensen-Dalsgaard. Measurements of stellar properties through asteroseismology: A tool for planet transit studies. *Proceedings of the International Astronomical Union*, 4(S253):309–317, 2008b.
- Pawan Kumar, Joel Franklin, and Peter Goldreich. Distribution functions for the time-averaged energies of stochastically excited solar p-modes. *Astrophysical Journal*, 328(2):879–887, 1988.
- Johnson N L. Continuous univariate distributions. 1970. URL <http://www.statsref.com/HTML/index.html?erlang.html>.
- H Lamb. Hydrodynamics, art 235, 1932.
- Glenn Ledrew. The real starry sky. *Journal of the Royal Astronomical Society of Canada*, 95:32, 2001.
- Lightkurve Collaboration, J. V. d. M. Cardoso, C. Hedges, M. Gully-Santiago, N. Saunders, A. M. Cody, T. Barclay, O. Hall, S. Sagear, E. Turtelboom, J. Zhang, A. Tzanidakis, K. Mighell, J. Coughlin, K. Bell, Z. Berta-Thompson, P. Williams, J. Dotson, and G. Barentsen. Lightkurve: Kepler and TESS time series analysis in Python. *Astrophysics Source Code Library*, December 2018.
- Mikkel N Lund, Víctor Silva Aguirre, Guy R Davies, William J Chaplin, Jørgen Christensen-Dalsgaard, Günter Houdek, Timothy R White, Timothy R Bedding, Warrick H Ball, Daniel Huber, et al. Standing on the shoulders of dwarfs: the kepler asteroseismic legacy sample. i. oscillation mode parameters. *The Astrophysical Journal*, 835(2):172, 2017.
- Marcella Marconi and Francesco Palla. The instability strip for pre-main-sequence stars. *The Astrophysical Journal Letters*, 507(2):L141, 1998.
- P Marcos-Arenal, W Zima, Joris De Ridder, Conny Aerts, Rik Huygen, R Samadi, J Green, G Piotto, Sébastien Salmon, C Catala, et al. The plato simulator: modelling of high-precision high-cadence space-based imaging. *Astronomy & Astrophysics*, 566:A92, 2014a.
- P Marcos-Arenal, W Zima, Joris De Ridder, Conny Aerts, Rik Huygen, R Samadi, J Green, G Piotto, Sébastien Salmon, C Catala, et al. The plato simulator: modelling of high-precision high-cadence space-based imaging. *Astronomy & Astrophysics*, 566:A92, 2014b.
- J.M. Carrasco & E. Masana. Simulation of gaia catalogue in plato fields, 2019. URL [https://indico.ict.inaf.it/event/806/contributions/4954/attachments/2762/5352/20190525\\_Carrasco.pdf](https://indico.ict.inaf.it/event/806/contributions/4954/attachments/2762/5352/20190525_Carrasco.pdf).
- PFL Maxted. Detached eclipsing binary stars as benchmark targets for plato. In *20th Cambridge Workshop on Cool Stars, Stellar Systems and the Sun*, 2018.
- A Miglio, WJ 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.
- A Miglio, C Chiappini, B Mosser, GR Davies, K Freeman, L Girardi, P Jofre, D Kawata, BM Rendle, Milena Valentini, et al. Plato as it is: a legacy mission for galactic archaeology. *Astronomische Nachrichten*, 338 (6):644–661, 2017.
- Marco Montalto, the WP130 (PSM), and WP340 (PDC) teams. P1, p2 and p5 samples: target selection criteria, September 2019. URL [https://indico.ict.inaf.it/event/806/contributions/4127/attachments/2804/5482/20190924\\_Montalto.pdf](https://indico.ict.inaf.it/event/806/contributions/4127/attachments/2804/5482/20190924_Montalto.pdf).
- C Mulet-Marquis, I Baraffe, Suzanne Aigrain, and F Pont. Accuracy of stellar parameters of exoplanet-host stars determined from asteroseismology. *Astronomy & Astrophysics*, 506(1):153–158, 2009.

Dermott J. Mullan. *Physics of the Sun: A First Course*. Chapman and Hall, 2013.

Martin Bo Nielsen, Laurent Gizon, H Schunker, and J Schou. Rotational splitting as a function of mode frequency for six sun-like stars. *Astronomy & Astrophysics*, 568:L12, 2014.

Onno Pols. Stellar structure and evolution, Sep 2011. URL [https://www.astro.ru.nl/~onnop/education/stev\\_utrecht\\_notes/](https://www.astro.ru.nl/~onnop/education/stev_utrecht_notes/).

I. Ramírez. Granulation in k-type dwarf stars, ii. hydrodynamic simulations and 3d spectrum synthesis, May 2009. URL <https://www.aanda.org/articles/aa/pdf/2009/27/aa11741-09.pdf>.

Heike Rauer, Claude Catala, et al. The plato mission. *Proceedings of the International Astronomical Union*, 6 (S276):354–358, 2010.

Heike Rauer, Conny Aerts, J Cabrera, PLATO Team, et al. The plato mission. *Astronomische Nachrichten*, 337(8-9):961–963, 2016.

Thomas P Robitaille, Erik J Tollerud, Perry Greenfield, Michael Droettboom, Erik Bray, Tom Aldcroft, Matt Davis, Adam Ginsburg, Adrian M Price-Whelan, Wolfgang E Kerzendorf, et al. Astropy: A community python package for astronomy. *Astronomy & Astrophysics*, 558:A33, 2013.

SIMBAD. The simbad astronomical database: Kepler-444 data. URL <http://simbad.u-strasbg.fr/simbad/sim-basic?Ident=kepler+444&submit=SIMBAD+search>. This research has made use of the

SIMBAD database, operated at CDS, Strasbourg, France, Wenger et al., Mar. 2020.

Space.com. Binarystarsystems: Classification and evolution, January 2018. URL <https://www.space.com/22509-binary-stars.html>.

D Stello, WJ Chaplin, S Basu, Y Elsworth, and TR Bedding. The relation between  $\delta v$  and  $v_{\max}$  for solar-like oscillations. *Monthly Notices of the Royal Astronomical Society: Letters*, 400(1):L80–L84, 2009.

T. L. Campante et al. An ancient extrasolar system with five sub-earth-size planets. *The Astrophysical Journal*, 799(2):170, Jan 2015. doi: 10.1088/0004-637x/799/2/170. URL <https://doi.org/10.1088/0004-637x/799/2/170>.

Th Toutain and C Fröhlich. Characteristics of solar p-modes-results from the iphir experiment. *Astronomy and Astrophysics*, 257:287–297, 1992.

W. R. Leo. 2.3 The Gaussian or Normal Distribution. [https://ned.ipac.caltech.edu/level5/Leo/Stats\\_contents.html](https://ned.ipac.caltech.edu/level5/Leo/Stats_contents.html), Last accessed: Mar 2020.

Weisstein. Gaussian function. URL <https://mathworld.wolfram.com/GaussianFunction.html>.

Timothy R White, Timothy R Bedding, Michael Gruberbauer, Othman Benomar, Dennis Stello, Thierry Appourchaux, William J Chaplin, Jørgen Christensen-Dalsgaard, Yvonne P Elsworth, Rafael A García, et al. Solving the mode identification problem in asteroseismology of f stars observed with kepler. *The Astrophysical Journal Letters*, 751(2):L36, 2012.