

# Stellar Ages and Stellar Rotation

by

Ruth Angus

Submitted to the Subdepartment of Astrophysics  
in partial fulfillment of the requirements for the degree of

Dphil in Astrophysics

at the

UNIVERSITY OF OXFORD

June 2016

© University of Oxford 2016. All rights reserved.

Author .....  
Subdepartment of Astrophysics  
April 30, 2016

Certified by .....  
Suzanne Aigrain  
Professor  
Thesis Supervisor

Accepted by .....  
Prof. Garret Cotter  
Head of graduate studies



# Stellar Ages and Stellar Rotation

by

Ruth Angus

Submitted to the Subdepartment of Astrophysics  
on April 30, 2016, in partial fulfillment of the  
requirements for the degree of  
Dphil in Astrophysics

## Abstract

Stellar ages will play a big role in the next generation of astronomy. Useful to exoplaneteers and galactic archaeologists alike, this relatively under-exploited stellar property is currently limited by the precision of dating techniques. The work presented in this thesis contributes incrementally to a greater understanding of rotation period decay in Main Sequence (MS) stars as a proxy for stellar age. Inferring stellar ages from rotation periods, ‘gyrochronology’, is the *only* dating method with the potential to provide ages for stars on the hundreds-of-thousands scale. Unfortunately however, it suffers from being poorly calibrated as the sample of cool, MS stars with precise ages is extremely sparse. Using light curves of spotted, rotating, MS, FGK stars with asteroseismic ages from the *Kepler* spacecraft, I attempted to recalibrate the relation between rotation period, colour and age. I demonstrate that the simple, ‘straight line’ gyrochronology relations used in the past are unable to explain the new asteroseismic sample. Questions are raised about the power of gyrochronology—does it accurately predict ages for old stars? To fully answer this question, it will be necessary to exploit new data from the *K2* (the repurposed *Kepler* mission). *K2* has observed (and is still observing) several open clusters and asteroseismic field stars which may provide new insight into stellar rotational evolution. Unfortunately, systematic features in *K2* light curves produced by *Kepler*’s reduced pointing precision inhibit the detection of astrophysical signals in the data. These systematic features can be removed by modelling and subtracting them from the time series, ‘detrending’, but this process can remove some signals and can even inject noise. For this reason I developed a method for detecting periodic signals in *K2* light curves without detrending: the Systematics-Insensitive Periodogram (SIP). This method is particularly useful for red giant asteroseismology. Precise ages can be inferred for oscillating red giants using the SIP and will be useful for galactic archaeology and open cluster age inference. In the next chapter of this thesis I return to the problem of stellar rotation period inference. Current methods for rotation period inference can produce inaccurate, imprecise periods with poorly approximated uncertainties and often without uncertainties altogether. I present a new method for inferring precise, accurate, probabilistic rotation periods with accurate uncertainties using Gaussian processes.

Although expensive to compute, this method is ideal for applying to individual targets. I hope to continue to develop this method and apply it to a large ensemble light curves from *Kepler* and other photometric surveys in the future. Star spots and acoustic (p-mode) oscillations are not the only mechanisms that produce variability in dwarfs and giants. A combination of asteroseismic pulsations and granulation on the stellar surface produces variability on short timescales. It has been shown that the amplitude of this short-term variability, called ‘flicker’ is strongly correlated with both surface gravity and stellar density (Bastien et al., 2013, 2016; Kipping et al., 2014). However, there is substantial additional scatter in these relations that is not accounted for by the observational uncertainties. I provide a new calibration of these relations which models this level of additional, astrophysical variance using hierarchical probabilistic inference. In the final chapter of this thesis I explore rotation period recovery with the Large Synoptic Survey Telescope (*LSST*). With its ten year baseline, *LSST* light curves will be sensitive to long rotation periods which are characteristic of old and low-mass stars. If the rotation periods of such stars can be measured from *LSST* light curves, it may be possible to improve the gyrochronology relations. We find that *LSST* is most sensitive to rotation periods between 10 and 20 days. Its sensitivity falls at short periods due to the sparsity of its sampling and at longer periods due to the lower amplitudes of variability and smaller apparent magnitudes of slow rotators.

Thesis Supervisor: Suzanne Aigrain  
Title: Professor

## Acknowledgments

More than anyone I would like to thank my Dphil supervisor, Suzanne Aigrain, who has been infinitely supportive over the last four years. She has always had time to provide excellent advice and has encouraged independent development, building my confidence as an emerging scientist. Over four years of working with Suzanne my admiration of her expert knowledge has only grown and our working relationship has been unceasingly wonderful. This is a testament of Suzanne's superb advising style. Because of Suzanne's love of statistics and incredible foresight she suggested I attend a summer school on 'statistics with *Kepler* data' at the SAMSI institute in North Carolina, USA in the first year of my Dphil. At this summer school I met the collaborators of my future. I would like to thank David Hogg, Eric Ford, Rebekkah Dawson, Tom Barclay, Ben Montet, Angie Wolfgang and the whole SAMSI working group for providing wonderful advice and expert guidance throughout my Dphil. A special thanks goes to John Johnson for supporting my time at the Harvard-Smithsonian Center for Astrophysics where I learnt more than I ever expected to about exoplanets and observational astronomy. Thank you to John's entire group, particularly Ben Montet and Andrew Vanderburg for so many stimulating discussions at group meetings. To Dave Latham: thank you for being a wonderful mentor at the CFA, twice. I would like to thank all my collaborators and coauthors: Suzanne Aigrain, Amy McQuillan, Dan Foreman-Mackey, John Johnson, David Kipping, Suzanne Hawley, Jim Davenport, Derek Buzasi and Søren Meibom. Finally I would like to thank Dan Foreman-Mackey, a major influence on all my work, and my parents for being unwaveringly supportive and excellent proof readers.



# Contents

<b>1</b>	<b>Introduction</b>	<b>17</b>
1.1	A little motivation . . . . .	17
1.2	Stellar dating methods . . . . .	18
1.2.1	Isochrone fitting . . . . .	20
1.2.2	Asteroseismology . . . . .	21
1.2.3	Age-rotation relations . . . . .	28
1.3	Other age diagnostics . . . . .	35
1.3.1	Magnetic activity . . . . .	35
1.3.2	Lithium depletion boundary . . . . .	37
1.3.3	Dynamics . . . . .	37
1.4	Kepler . . . . .	38
1.5	Statistical methods . . . . .	46
1.5.1	Multi-dimensional and unknown uncertainties: hierarchical probabilistic modelling . . . . .	46
1.5.2	Methods for removing systematics . . . . .	49
1.5.3	Rotation period inference . . . . .	51
<b>2</b>	<b>Calibrating Gyrochronology using Kepler Asteroseismic targets</b>	<b>57</b>
2.0.1	Gyrochronology . . . . .	59
2.1	Observations . . . . .	61
2.2	Calibrating the Gyrochronology relation . . . . .	67
2.2.1	The model . . . . .	67
2.3	Results and Discussion . . . . .	74

2.4	Conclusions . . . . .	86
<b>3</b>	<b>Systematics-insensitive periodic signal search with K2</b>	<b>89</b>
3.1	Introduction . . . . .	89
3.1.1	Asteroseismology . . . . .	90
3.1.2	Stellar rotation . . . . .	93
3.2	Method . . . . .	94
3.3	Application to real light curves . . . . .	96
3.3.1	Tests and discussion . . . . .	104
3.4	Conclusions . . . . .	110
<b>4</b>	<b>Stellar rotation period inference using Gaussian processes</b>	<b>113</b>
4.1	Introduction . . . . .	114
4.2	Rotation period injection and recovery . . . . .	119
4.2.1	ACF . . . . .	121
4.2.2	The sine-fitting periodogram method . . . . .	122
4.2.3	The GP method . . . . .	125
4.2.4	Real <i>Kepler</i> data . . . . .	127
4.3	Discussion . . . . .	133
4.3.1	The ACF method . . . . .	133
4.3.2	Initialisation . . . . .	134
4.3.3	Kernel function choice and interpretation of hyper-parameters	135
4.3.4	Future work . . . . .	136
4.4	Conclusions . . . . .	138
<b>5</b>	<b>Probabilistic Inference of basic stellar parameters: application to flickering stars</b>	<b>139</b>
5.1	Introduction . . . . .	140
5.2	Probabilistic calibration . . . . .	144
5.2.1	Calibration Data . . . . .	144
5.2.2	Hierarchical Bayesian Model . . . . .	145

5.3	Discussion . . . . .	152
<b>6</b>	<b>LSST</b>	<b>155</b>
6.0.1	Introduction . . . . .	156
6.0.2	Simulations . . . . .	157
6.0.3	Results . . . . .	168
6.0.4	Conclusions . . . . .	172
<b>7</b>	<b>Conclusions</b>	<b>175</b>



# List of Figures

1-1	A CMD of the Hyades.	22
1-2	Solar power spectrum.	24
1-3	Skumanich (1972) results: early evidence for magnetic braking	32
1-4	Cluster rotation from Irwin and Bouvier (2009)	34
1-5	The <i>Kepler</i> spacecraft and its field of view.	43
1-6	An example <i>Kepler</i> light curve.	44
1-7	Current <i>K2</i> fields	45
1-8	Future <i>K2</i> fields	45
2-1	Period vs age for <i>Kepler</i> stars	64
2-2	Period vs colour for <i>Kepler</i> stars	65
2-3	$\log g_{\text{vs}} T_{\text{eff}}$ for <i>Kepler</i> stars	69
2-4	Fits to age-rotation curves for individual clusters	77
2-5	Marginal posteriors of gyrochronology parameters	79
2-6	A new gyrochronology relation: period vs age.	80
2-7	A new gyrochronology relation: period vs colour.	81
2-8	Period vs colour over a range of ages	82
3-1	Comparing the SIP to a detrended LS periodogram.	91
3-2	SIPs of oscillating red giants.	98
3-3	Example 1: an ACF and LS periodogram of a detrended light curve.	99
3-4	Example 1: A SIP with a conditioned light curve.	100
3-5	Example 2: an ACF and LS periodogram of a detrended light curve.	102
3-6	Example 2: A SIP with a conditioned light curve.	103

3-7	SIP detection efficiency. . . . .	105
3-8	How insensitive is the SIP? . . . . .	106
3-9	Thruster firing signals in $K2$ light curves. . . . .	107
3-10	Exoplanets with the SIP . . . . .	108
4-1	A light curve with a GP model. . . . .	118
4-2	A simulated light curve. . . . .	120
4-3	An ACF of a simulated light curve. . . . .	121
4-4	ACF results. . . . .	123
4-5	LS periodogram results. . . . .	124
4-6	GP results. . . . .	128
4-7	All results. . . . .	129
4-8	A simulated light curve with a rotation period of 14.4 days. The pink curve shows the GP model with the best-fit parameters. The marginal posteriors of the GP hyper-parameters used to produce this fit are shown in figure 4-9. . . . .	130
4-9	Marginal posterior PDFs of the QP GP model parameters. $\sigma$ is an additional white noise term added to the diagonal elements of the covariance matrix. It is the fraction by which the observational uncertainties have been underestimated. If the errorbars reported on the data are too small, this parameter will be non-zero. Since this light curve was simulated without noise, this parameter is very close to zero. The true rotation period of this star is 14.4 days. . . . .	131
5-1	Granulation on the Solar surface. . . . .	142
5-2	Flicker versus stellar surface gravity, from Bastien et al. (2013). . .	143
5-3	Stellar density versus flicker. . . . .	149
5-4	$\log g$ versus flicker. . . . .	150
5-5	Marginal posterior PDFs for flicker versus $\rho_*$ parameters. . . . .	151
6-1	An $LSST$ cadence histogram. . . . .	158

6-2	A histogram of the rotation periods of hot stars . . . . .	160
6-3	The effective temperature distribution for stars used in <i>LSST</i> simulations	161
6-4	The period distribution for stars used in <i>LSST</i> simulations . . . . .	162
6-5	The mean amplitudes of variability for stars in the McQuillan et al. (2014) catalogue, according to their temperatures. . . . .	165
6-6	Example simulated <i>LSST</i> light curve 1. . . . .	166
6-7	Example simulated <i>LSST</i> light curve 2. . . . .	167
6-8	<i>LSST</i> rotation period recovery results with temperature dependence .	169
6-9	<i>LSST</i> rotation period recovery results with magnitude dependence .	170
6-10	<i>LSST</i> rotation period recovery results for different spectral types . .	171



# List of Tables

2.1	Gyrochronology results.	60
2.2	Cluster data and references.	64
2.3	Field star data.	66
2.4	Individual cluster results.	74
2.5	Nuisance parameter results	76
2.6	Field star predictions	84
5.1	Flicker results.	148



# Chapter 1

## Introduction

### 1.1 A little motivation

The desire to infer precise and accurate stellar ages is motivated by a number of scientific questions. For example, without this fundamental stellar property, the evolution of stars and stellar populations would be a mystery. In addition, our understanding of the formation history of the Milky Way would be limited and with it, the formation process of all galaxies. Some day in the (hopefully) not-too-distant future, the first biomarker will be detected in the atmosphere of an Earth-like planet. When that time comes, knowing the age of the planetary system will be fundamental. These are all excellent reasons for studying stellar ages but the application that most piques my personal interest is the implication for understanding the dynamical evolution of exoplanetary systems.

Despite the rapidly accelerating interest in exoplanet population studies (e.g. Burke et al., 2015; Dressing and Charbonneau, 2015; Foreman-Mackey et al., 2014a; Petigura et al., 2013), we still know very little about how planetary systems *evolve* because stellar ages are difficult to infer. Planetary system architectures are not static in time: chaotic gravitational interactions can fling planets into interstellar space or plunge them into the surface of their star. If they cross orbits they can even collide or be excited to highly eccentric and inclined orbits. Simulations of planetary systems often demonstrate that planet losses are most common in the few millions of

years immediately after formation and continue at a gradually decreasing rate (e.g. Funk et al., 2010; Pu and Wu, 2015; Smith and Lissauer, 2009; Zhou et al., 2007). Planet-planet scattering on extended timescales may result in a decrease in exoplanet frequency with stellar host age in the enormous sample of over five thousand exoplanets discovered to date *if* their ages can be sufficiently constrained (Veras et al., 2015).

Determining the detectability of trends in the ages of *Kepler* systems is challenging as the outcomes of simulations depend strongly on input assumptions—different studies therefore produce different predictions (see figure 3 of Pu and Wu, 2015). However, based on the Smith and Lissauer (2009) simulations of systems with three Earth-mass planets, Veras et al. (2015) demonstrate that a decrease in planet occurrence rate will be detectable for K dwarf hosts, even if stellar age uncertainties are as large as 5 Gyr, and for G dwarfs with age uncertainties of 3.5 Gyr. Clearly, most systems do not consist of equal-mass planets, but the study of these simulations demonstrates that an age trend lies well within the realms of detectability.

This is just one of the tantalising scientific discoveries looming over the horizon, if only we could improve our methods for inferring precise stellar ages. As I explain in this thesis, my research has advanced our understanding of the stellar age-rotation, or ‘gyrochronology’ relations, contributing to our ability to infer stellar ages as a community. This research may ultimately lead to discovering variations in the architecture and frequency of exoplanets as a function of host star age.

## 1.2 Stellar dating methods

The hydrogen-burning era of a star’s life is extremely stable. This fact is unfortunate from a stellar chronologist’s point of view as, without an observable property that is a strong function of age, stellar age precision will always be limited. This is the main stumbling-block for stellar age inference: age precision is restricted by the fundamental evolutionary timescale of stars. Despite this limitation however, there is still plenty of room for progress to be made in understanding the evolution of the

observable properties from both empirical and theoretical standpoints, and in the precision with which we can measure these properties.

Isochronal ages are notoriously difficult to infer because stars vary little in brightness or temperature during their hydrogen-burning lifetimes. Fitting stellar evolutionary models to these two observables usually produces age estimates with uncertainties in the order of 50-150%. However, because cool stars spin down predictably over their main sequence lifetime due to magnetic braking, their current rotation periods depend, to first order, only on their masses and ages. Photometric measurements of these two parameters can therefore be used to infer an age. This is the concept behind ‘gyrochronology’—inferring an age from a rotation period—developed after observations of stars in clusters revealed a trend for decreasing rotation period with time, at a given mass (Barnes, 2003; Irwin and Bouvier, 2009; Skumanich, 1972; Weber and Davis, 1967). This method has never held as much promise as it does now, with a glut of stellar rotation periods available from a new generation of spacecraft designed for high precision, high cadence photometry. The main science goal of these missions is searching for exoplanets, but the exoplanet gold-rush has also lead to a new understanding of stars and stellar variability. The desire to develop a method of dating stars using only photometry is an understandable one with the current availability of such large photometric data sets. However, as with any phenomenological investigation, inferences about the data can only be made with models. These models can be physical or empirical but whatever the origin of their design, they must be calibrated using observations. No age model can be developed in isolation from the observations—our understanding of stars just isn’t good enough. So when a new dating method is developed, its efficacy relies on the existence (and precision) of previous dating methods. It is therefore important to understand the currently available dating methods, their advantages and their limitations. In what follows, I describe the main methods used today.

### 1.2.1 Isochrone fitting

As stars burn hydrogen they become hotter and more luminous, slowly moving towards the top right of the Hertzprung-Russel Diagram (HRD), or Colour-Magnitude Diagram (CMD). As the hydrogen-burning process progresses, helium ‘ash’ is produced in the core and energy production decreases. The core slowly contracts over time and its temperature therefore increases. The increased temperature enhances energy production, and the result is that stars burn hotter and more brightly over time. By inferring a star’s absolute magnitude and effective temperature, or measuring its colour, plus its composition and evolutionary stage, it is possible to place it on an HRD or CMD and infer its age.

In practice, it is not usually possible to feed luminosity, temperature, metallicity and  $\log g$  into a stellar evolution model, crank the handle and pull out an age. Evaluating these models is expensive, so pre-calculated model grids are used. In order to infer an age from the observations it is therefore necessary to interpolate between grid lines. The Dartmouth (Dotter et al., 2008), Yonsei-Yale ( $Y^2$ ) (Spada et al., 2013; Yi et al., 2001), Padova (Girardi et al., 2002) and PARSEC (Bressan et al., 2012) models are some of the most commonly used sets of models. These and other isochrone grids compute slightly different ages for the same star because the underlying stellar evolution models are different. For example, the four models listed above use different equations of state, different models for opacity within the stars and different atmospheric models (see Thompson et al., 2014, for a comparison of these four sets of isochrones).

An obvious limitation of the isochrone method is that the composition of the star in question will affect its placement on a HRD or CMD: metal-rich stars appear cooler and redder than metal-poor ones with the same mass and age. It is often difficult or impossible to obtain precise metallicities, helium abundances and alpha-element fractions (all needed for precise isochrone placement), as these measurements come from expensive, high-precision spectroscopy. However, for an ensemble of coeval stars with identical compositions this process becomes much easier: not only are

there more opportunities to measure stellar compositions, providing a  $\sqrt{N}$  reduction in measurement uncertainty, a group of stars with the same age and composition but different masses will reveal the shape of the best-fitting isochrone on the CMD. In addition, the position of the main sequence (MS) turn-off further improves age precision. Since MS lifetime is strongly dependent on mass, at any given age the most massive stars in a cluster will have turned off the MS. Inferring the mass at which this happens in a cluster provides a very precise age estimate. For these reasons, many open clusters have precise ages with typical age uncertainties of around 10%. Age discrepancies arising from differences in isochrone models can also be as high as 10%. Together with the Sun they are the most precisely dated of all astronomical objects and provide the benchmarks from which all other dating methods are calibrated. Figure 1-1 shows 69 field stars in the Hyades plotted on a CMD from Perryman et al. (1998). Isochrones at a range of ages are also shown, as is the Zero Age Main Sequence (ZAMS), the line upon which stars fall as they enter the main sequence. Perryman et al. (1998) infer an isochronal age of  $625 \pm 50$  Myr for the Hyades.

### 1.2.2 Asteroseismology

The NASA *Kepler* mission plays a fundamental role in this thesis. I describe the *Kepler* space craft and its unique data set in detail in §1.4. The main objective of the *Kepler* mission is to search for extra solar planets using the transiting method, however since extremely precise photometry is required for exoplanet search, a number of other scientific fields have benefited from its exquisite light curves. *Kepler*'s legacy is not limited to exoplanets—it has arguably advanced stellar astronomy more than any other single purpose astronomical instrument ever created. Among its several successes is the improvement of stellar dating methods, as I explain below.

After exoplanet search, arguably the second most important branch of *Kepler*'s legacy is asteroseismology; the study of stellar pulsations. Sound waves ripple through the deep interiors of stars. The frequencies of these waves can reveal internal structure, even localising the age-dependent hydrogen burning region. Asteroseismology is a powerful tool in the stellar dating arsenal and while it is not the central topic of my

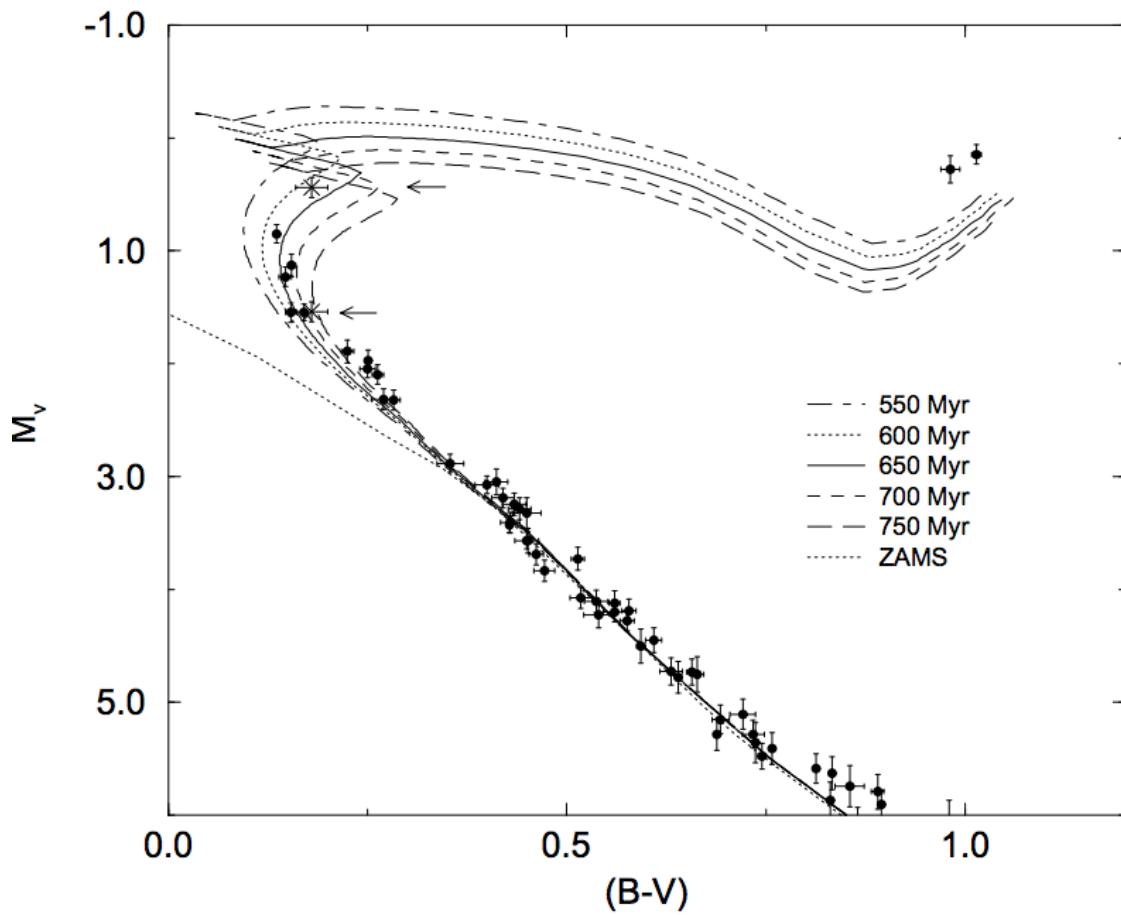


Figure 1-1: A CMD of the Hyades with Geneva isochrones (Schaller et al., 1992) and the ZAMS from Perryman et al. (1998). Perryman et al. (1998) infer an isochronal age of  $625 \pm 50$  Myr for the Hyades.

thesis, it features at some level in almost every chapter, so I cover the basic principles below.

Acoustic oscillations in the Sun are generated by turbulent convection near the surface. The movement of ionized gas stochastically excites the Sun's spherical oscillation modes. An analogy to this process is that of a bell in a room filled with air. Air particles colliding with the surface of the bell cause it to vibrate at all of its spherical harmonic frequencies. There is no coherent driving force, rather the stochastic collisions of air particles induce a continual ringing. Similarly, the Sun is continually oscillating at a range of discrete frequencies; those that correspond to its spherical harmonics. The restoring force of these acoustic oscillations is pressure, thus these waves are called ‘pressure’ or ‘p’-modes. A power spectrum of the Sun, taken from Brown (2000) is shown in figure 1-2. The comb-like, evenly spaced peaks in this power spectrum are the Solar p-mode oscillations. The peak amplitudes are modulated by a Gaussian envelope and the mean of that Gaussian,  $\nu_{max}$  corresponds to a frequency of around 3000  $\mu$ Hz, a period of around five minutes.

Many properties of the Sun have been measured using asteroseismic pulsations, for example the variation in radial pressure, the rate of differential rotation, the depth and fraction of helium in its core and even the structure of active regions below the Solar surface. The Sun is an exquisite example of a pulsating star, having tens of millions of detectable modes and mode lifetimes that are several thousand times longer than the periods of the oscillations. These long mode lifetimes result in narrow peaks in the Fourier transform of a Solar light curve or RV time series. Short-lived modes produce signals that are incoherent over long timescales and therefore have broadened peaks in frequency space. Its proximity, which provides both enormous S/N and allows us to resolve the surface, makes it a paragon of seismology.

Three types of oscillations manifest themselves in stars: pressure (p-modes), surface (s-modes) and gravity (g-modes). Pressure provides the restoring force for p-mode oscillations, and these are effectively sound waves. Surface modes are, as the name suggests, waves on the surface of the star. G-modes are excited by buoyant gas in the radiative zone and the restoring force is gravity. G-mode waves rapidly

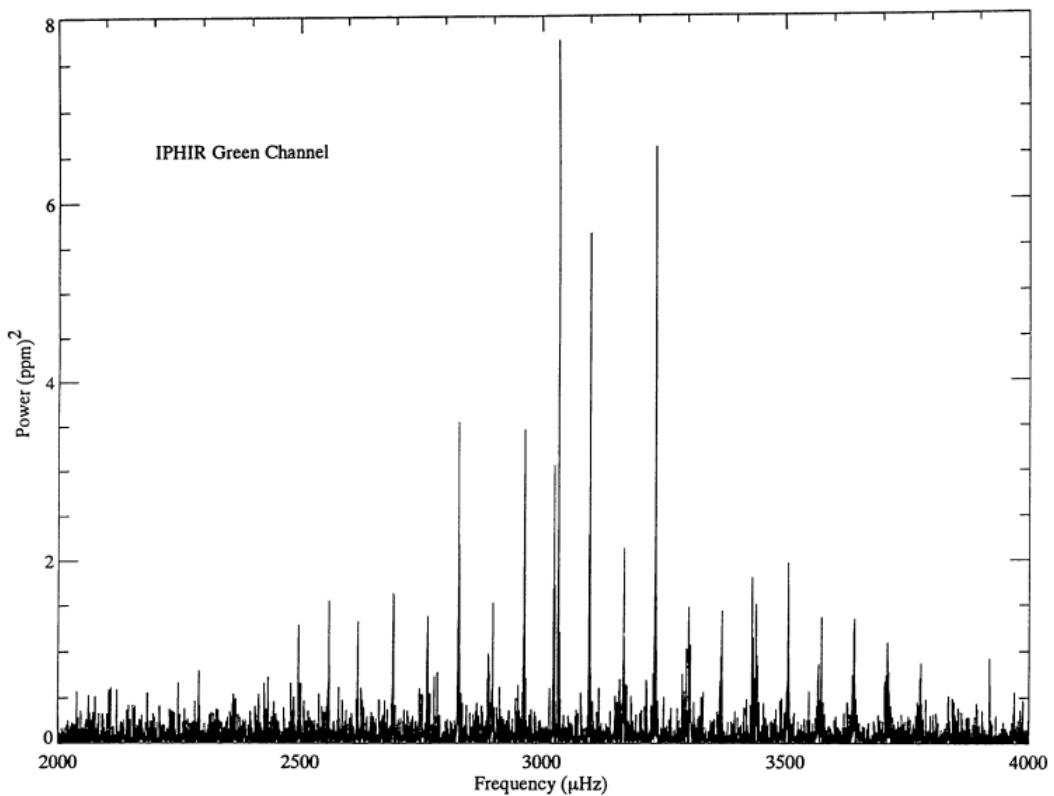


Figure 1-2: A power spectrum of one month of disc-integrated Solar photometry, taken from Toutain and Froehlich (1992). Solar p-modes are clearly visible in this figure.

evanesce in stellar convective zones and therefore appear at very low amplitudes at the surface of Sun-like stars. For the work in this thesis I am concerned only with p-mode observations as these waves reveal the internal structure of a star which varies as a function of mass, radius and age.

The frequency of a p-mode wave is proportional to the sound-speed of the gas along its path through the stellar interior. Time-dependent spatial perturbations to a star's equilibrium state can be written as a product of a term that depends on stellar radius and a spherical harmonic (assuming that a star can be approximated as a sphere) as follows (Brown and Gilliland, 1994),

$$\xi_{nlm}(r, \theta, \phi, t) = \xi_{nl}(r) Y_l^m(\theta, \phi) e^{-i\omega_{nlm}t}. \quad (1.1)$$

$\xi$  is a spatial perturbation, associated with a mode,  $r, \theta, \phi, \omega$  and  $t$  are the radial coordinate, colatitude, longitude, angular frequency<sup>1</sup> and time, respectively.  $n$  is the radial order, defined as the number of nodes between the star's centre and surface and  $l$  is the angular degree, the product of the stellar radius and the total horizontal wavenumber of the mode. For example, a star oscillating with an  $l = 10$  mode will have a standing wave with 10 nodes along a line connecting the poles. Finally,  $m$  is the azimuthal order; the projection of  $l$  onto the equator and must be less than or equal to  $l$ . A star with an  $m = 5$  mode will have a standing wave with 5 nodes along the equator. The relation between  $\omega$  and  $n$  and  $l$  is complicated and depends on the structure of the star. Oscillations with different values of  $l$  penetrate to different depths in the star.  $l = 0$  waves travel through the centre and waves with increasing  $l$  skirt the central region by a larger and larger distance. For this reason, modes with different  $l$  provide information about the sound speed gradient in the stellar interior. The 'small frequency separation',  $\delta_{n,l} = \nu_{n+1,l} - \nu_{n,l+2}$  is often used to parameterise the variation in frequency with stellar radius via (Brown and Gilliland, 1994)

$$\delta_{n,l} = \Delta\nu_0 \frac{(l+1)}{2\pi^2\nu_{nl}} \int_0^{R_\star} \frac{dc}{dr} \frac{dr}{r}. \quad (1.2)$$

---

<sup>1</sup>The convention in asteroseismology is to report circular frequencies,  $\nu_{nlm} = \omega_{nlm}/2\pi$ .

Since nuclear-burning material in the core changes its molecular weight as the star evolves, the sound-speed gradient is time-dependent and the small separation therefore contains information about the age of the star.

There is no simple harmonic relation between the frequencies of modes with adjacent mode number (Brown and Gilliland, 1994). However, in the limit where  $n \gg l$ , mode frequency can be approximated as

$$\nu_{nl} = \Delta\nu_0 \left( n + \frac{l}{2} + \epsilon \right) - \frac{AL^2 - \eta}{(n + l/2 + \eta)}, \quad (1.3)$$

where parameters  $\Delta\nu_0$ ,  $A$ ,  $\epsilon$  and  $\eta$  depend on the structure of the star and  $L^2 = l(l+1)$ .  $\Delta\nu_0$  is the ‘large frequency separation’ which, together with the peak frequency of the Gaussian envelope that modulates the amplitudes of the oscillation modes,  $\nu_{max}$ , makes up the two fundamental asteroseismic observables. It is related to the sound travel time through the centre of the star:

$$\Delta\nu_0 = \left( 2 \int_0^{R_\star} \frac{dr}{c} \right)^{-1}, \quad (1.4)$$

where  $c$  is the local sound speed and  $R_\star$  is the stellar radius. This travel time is related to the mean density of the star via,

$$\Delta\nu_0 \cong 135 \left( \frac{M_\star}{R_\star^3} \right)^{1/2} \mu Hz, \quad (1.5)$$

where  $M_\star$  and  $R_\star$  are the stellar mass and radius in Solar units (Brown, 2000; Cox, 1980).

Today, p-modes have been detected in hundreds of Sun-like stars, however it was not until the late 1990s that the first conclusive detection was made in a star other than our Sun. The reason for this is simply that p-mode perturbations are extremely small—these changes are around  $10 \text{ cms}^{-1}$  in velocity and  $3\mu\text{mag}$  in brightness for typical oscillation modes in the Sun (Brown, 2000). Asteroseismic pulsations are usually detected in two different ways: by searching for the subtle change in luminosity caused by the temperature fluctuations of the stellar surface, or by measuring the

changing radial velocity of the surface. A Fourier transform of these time series will reveal the presence of oscillation modes, allowing for the modelling of the star's interior structure. The earliest p-modes detections were made using radial velocity data (Kjeldsen and Bedding, 2001) for stars  **$\eta$  Boo** (Kjeldsen et al., 1995), **Procyon** (Barban et al., 1999; Martić et al., 1999),  **$\zeta$  Herculis** (Martić et al., 2001),  **$\alpha$  Cen A** (Kjeldsen et al., 1999) and  **$\beta$  Hyi** (Bedding et al., 2001). Today, we can take advantage of space-based missions *Kepler* and *CoRoT* whose photometric precision provides sufficient S/N to detect p-mode-induced luminosity variations. These missions have provided fundamental parameters for hundreds of oscillating giants, subgiants and Sun-like stars (e.g. Bruntt, 2009; Chaplin et al., 2014; Michel et al., 2008).

The frequencies of oscillations accessible in *Kepler* light curves are limited by two things: the overall duration of observations and the time sampling interval. The duration of observations, *i.e.* the length of the *Kepler* mission sets the minimum resolvable frequency and the time sampling sets the Nyquist frequency: the high frequency limit, which is equal to half the sampling rate. For FGK dwarfs, the Nyquist frequency of long cadence observations ( $\sim 283\mu\text{Hz}$ ) is too low as these stars oscillate at around  $3000\ \mu\text{Hz}$ . For this reason, *Kepler* operates in two observing modes: long and short cadence. Long cadence observations are taken approximately every 30 minutes and short cadence every minute ( $\nu_{Nyquist} \sim 8333\mu\text{Hz}$ ). Around 560 dwarfs and subgiants were observed in short cadence mode (Chaplin et al., 2014).

There are two different approaches to inferring asteroseismic ages from *Kepler* light curves, depending on the signal-to-noise ratio (S/N) of pulsations. In the high S/N cases it may be possible to identify oscillation frequencies of individual modes (e.g. Lebreton and Goupil, 2014; Metcalfe et al., 2010; Silva Aguirre et al., 2012a). However for the majority of short cadence targets, only the mean large frequency separation,  $\Delta\nu$  and frequency of maximum power,  $\nu_{max}$  can be measured. Bulk physical properties of stars can be inferred from these two parameters via the scaling relations (Brown et al., 1991; Kjeldsen et al., 1995):

$$\frac{M}{M_\odot} = \left( \frac{\nu_{max}}{\nu_{max,\odot}} \right)^3 \left( \frac{\Delta\nu}{\Delta\nu_\odot} \right)^{-4} \left( \frac{T_{\text{eff}}}{T_{\text{eff},\odot}} \right)^{3/2} \quad (1.6)$$

$$\frac{R}{R_\odot} = \left( \frac{\nu_{max}}{\nu_{max,\odot}} \right) \left( \frac{\Delta\nu}{\Delta\nu_\odot} \right)^{-2} \left( \frac{T_{eff}}{T_{eff,\odot}} \right)^{1/2}. \quad (1.7)$$

These scaling relations have been tested on Solar-like oscillators and give good agreement with the observations (*e.g.* Chaplin and Miglio, 2013; Coelho et al., 2015). Ages are then inferred by comparing these quantities to those predicted by model grids (spectroscopic estimates of  $T_{eff}$  and [Fe/H] are also required). Two of the most sophisticated codes used for age analysis are the BAyesian STellar Algorithm (BASTA Silva Aguirre et al., 2015) and the Bellaterra Stellar Properties Pipeline (Serenelli et al., 2013). Ages inferred using  $\Delta\nu$  and  $\nu_{max}$  typically have relatively large uncertainties, in the order of 15-40% (Silva Aguirre and Serenelli, 2015). In high S/N cases, ages are inferred by adjusting the parameters of stellar interior models (*e.g.* Kjeldsen et al., 2008) until the observed frequencies, or combinations of them are reproduced. Uncertainties on ages inferred from high S/N light curves using this ‘boutique’ method of modelling individual frequencies are typically around 10-15% (Silva Aguirre and Serenelli, 2015).

Asteroseismology is a powerful dating method with the potential to yield extremely precise stellar ages. However, the era of asteroseismology has only just arrived and this fledgling field is still only applicable to a small number of extremely bright stars observed by precise photometric space-missions. Given the quantity of up-and-coming photometric space missions, asteroseismology will continue to provide precise stellar parameters for decades to come. However, accurate and precise asteroseismic ages demand high-precision photometry of bright (brighter than around 12th magnitude) stars. Even with *Kepler* and *CoRoT*, this only amounts to around one hundred MS stars with asteroseismic ages available today. It is therefore essential that alternative dating methods, which can be applicable to a much larger sample of stars, are developed; age-rotation relations, for example.

### 1.2.3 Age-rotation relations

Asteroseismology is revolutionising stellar astronomy and, perhaps in particular, stellar ages (partly because the bar is so low to begin with). However, the quantity of

ultra-precise, short cadence light curves is limited and it is still a ‘by hand’ method applied to hand-picked, bright, text-book Solar-like oscillators. In order to produce a catalogue of stellar ages large enough to be useful for stellar population studies, we need a method that can be applied to thousands of stars. Such a method should require inexpensive observables, be computationally cheap and be easily automated. Gyrochronology has the potential to satisfy these criteria.

### Gyrochronology: the theory

Today, we know that angular momentum loss in main sequence stars is caused by a magnetised stellar wind (Mestel, 1984; Schatzman, 1962; Weber and Davis, 1967). The stellar wind is composed of charged ions that stream away from the stellar surface. Each of these ions carries mass, charge and angular momentum as it travels away from the star. The angular momentum carried by these particles is lost from the system, not at the photosphere because the wind particles travel along magnetic field lines which are radial close to the surface of the star (in the corotating frame), but further away from the photosphere, as thermal expansion slowly begins to dominate the motion of stellar wind particles. This mechanism is the main reason why the rotation periods of stars decay during their stint on the MS. In addition, the stellar wind carries mass away from the star and its internal structure changes due to core hydrogen burning. Both of these processes alter the star’s moment of inertia, however this has a small effect on rotation period evolution. Stellar magnetism is almost entirely responsible for rotational evolution on the main sequence.

The motion of turbulent plasma in the outer layers of low-mass stars drives the production of a magnetic field (Schatzman, 1962). Dynamo theory predicts that this field is then amplified by rotation and differential rotation (*e.g.* Parker, 1970). Convective turbulence, combined with stellar rotation produces the large-scale magnetic fields that lock the stellar wind to the star (*e.g.* Charbonneau, 2010). Rotation period influences magnetic field strength: more rapid rotators have stronger magnetic fields. The relation between angular momentum loss rate and angular velocity is a power-law, where the exponent depends on the magnetic field geometry (Kawaler,

1988; Mestel, 1984). Magnetic field strength controls the rate of angular momentum loss: a stronger magnetic field leads to stellar wind particles corotating with the photosphere out to a greater radius which increases the angular momentum lost per unit time. Increasing the mass loss rate and temperature of the stellar wind also results in an increased angular momentum loss rate, although these have secondary effects. For this reason the rate of angular momentum loss is not constant: as a star loses angular momentum its rotation period decays, as does its magnetic field strength, so the rate of angular momentum loss decreases. The overall effect of this process is to force stellar rotation periods to converge. Rapid rotators experience a greater angular momentum loss rate than slower rotators. Their rotation periods therefore decay more rapidly than slow rotators and, after a few hundred million years all low-mass stars appear to rotate at a rate that depends, to first order, only on their age and mass.

This convergence happens more quickly for early-type stars than late-type. Whilst F, G and K-type stars may have converged by 500 million years, (Irwin and Bouvier, 2009; Radick et al., 1987) late M dwarfs may not converge within a Hubble time. Since the magnetic dynamo is believed to be generated in the convective zone, stars with a thin convective layer have a weaker magnetic field. For this reason they lose angular momentum more slowly. Angular momentum loss rate therefore varies as a function of stellar mass and fully radiative stars do not spin down appreciably over their MS lifetime (Noyes et al., 1984b). The rotational evolution of fully convective stars is still unknown, however it is an active field of interest (*e.g.* McQuillan et al., 2013a; Newton et al., 2015).

Weber and Davis (1967), Mestel and Spruit (1987) and Kawaler (1988) developed the first physical model of a star undergoing magnetic braking. Since this original work, several theoretically motivated gyrochronology models have been developed by different groups. For example, Collier Cameron and Jianke (1994); Epstein and Pinsonneault (2014); Reiners and Mohanty (2012); van Saders and Pinsonneault (2013) all draw on the principles laid down by Kawaler (1988), to create physical models of rotating stars with magnetic fields, where the rate of angular momentum loss is

related to the rotation period and field strength, and evolve those stars forward in time. Although these models are based on the physical processes at play, they must still be calibrated using observations. For example, the Kawaler (1988) expression, assuming a linear dynamo is

$$\frac{dJ}{dt} = -K_w \Omega^3 \left( \frac{R}{R_\odot} \right)^{0.5} \left( \frac{M}{M_\odot} \right)^{-0.5}, \quad (1.8)$$

where  $J$  is angular momentum,  $R$  and  $M$  are radius and mass respectively and  $K_w$  is a constant that must be calibrated so that the Sun has Solar rotation at Solar age. See Barnes and Kim (2010) for a discussion on the various adaptations of this braking law. van Saders and Pinsonneault (2013) use the following model to describe the rotation period distribution of cool field stars:

$$\frac{dJ}{dt} = \begin{cases} f_K K_M \omega \left( \frac{\omega_{crit}}{\omega_\odot} \right)^2, & \omega_{crit} \leq \omega \frac{\tau_c}{\tau_{c,\odot}} \\ f_K K_M \omega \left( \frac{\omega \tau_c}{\omega_\odot \tau_{c,\odot}} \right)^2, & \omega_{crit} > \omega \frac{\tau_c}{\tau_{c,\odot}} \end{cases} \quad (1.9)$$

With

$$\frac{K_M}{K_{M,\odot}} = \left( \frac{R}{R_\odot}^{3.1} \right) \left( \frac{M}{M_\odot} \right)^{-0.22} \left( \frac{L}{L_\odot} \right)^{0.56} \left( \frac{P_{phot}}{P_{phot,\odot}} \right)^{0.44}, \quad (1.10)$$

where  $f_K$  is a constant, scaled to produce solar rotation at solar age for the Sun and  $P_{phot}$  is the pressure at the photosphere.  $\tau_c$  is the convective overturn time at the base of the convective zone and  $P_{rot}/\tau_c$  is known as the Rossby number,  $Ro$ . It is dependent on spectral type and for this reason many people argue that Rossby number is directly related to the surface magnetic field strength and a more natural quantity to correlate with age and activity than rotation period (*e.g.* van Saders et al., 2016) (although this view is not universally held — see, *e.g.* Reiners et al., 2014). This is one of the more recently developed theoretical gyrochronology relations — however, as revealed in chapter 7, this equation has now been adapted to describe the most recent observations.

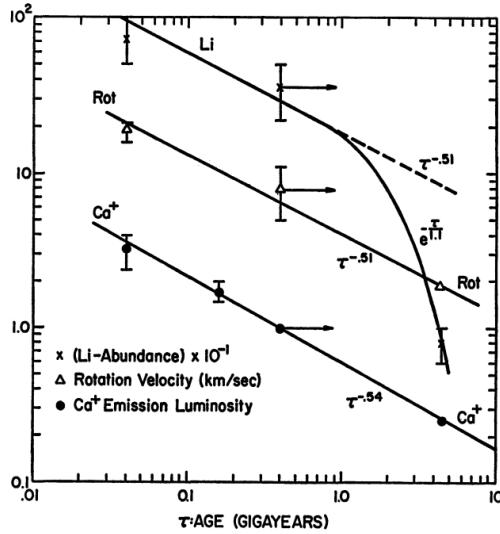


Figure 1-3: Figure from Skumanich (1972). Rotation periods for the G stars in the Hyades and Pleiades and the Sun are plotted.

### Gyrochronology: the observations

Skumanich (1972) provided one of the first indications that stellar rotation periods decay over time and demonstrated that rotation period, lithium abundance and chromospheric activity decay is proportional to the square-root of age. In figure 1-3, equatorial rotational velocity<sup>2</sup> versus time (triangles) is plotted on the same axes as lithium abundance (crosses) and Ca<sup>+</sup> emission (circles) for the Hyades and Pleiades clusters, as well as the Sun. The Sun is the far right point for each age indicator. Ursa Major stars are included on the Ca<sup>+</sup> emission scale. Lithium abundance and Ca<sup>+</sup> emission were previously known age indicators (see §1.3) and this work demonstrated that rotation period was also related to age.

Kawaler (1989) applied the Kawaler (1988) magnetic braking law to stars in the Hyades cluster and found good agreement with the model. Barnes (2003) compiled rotation period measurements from members of several young clusters, plus the Mt. Wilson stars. When the rotation periods of this sample were plotted against B-V colour, Barnes (2003) noticed that there were two morphological features of the data (see their figure 2). In each cluster there was a sequence of stars falling neatly on

---

<sup>2</sup> Only the equatorial velocities of the G stars in the two clusters are indicated.

the predicted relation between rotation period and mass (called interface or ‘I’ stars), but there was also a group of rapid rotators (called convective, ‘C’ stars). The C sequence was most obvious in the younger clusters and less so in the older. Barnes (2003) attributed this behaviour to an evolving magnetic dynamo produced by the still evolving internal structure of the young stars, and postulated that stars transition rapidly from C to I. In this work I will only address stars that have already (in theory) transitioned from the C sequence to the I sequence as they are older than the Hyades which is the oldest cluster to show this behaviour.

Irwin and Bouvier (2009) compiled rotation period measurements for stars in open clusters with masses  $< 1.2M_{\odot}$ , shown in figure 1-4. These data show the enormous spread in rotation periods for the youngest clusters (top left), contrasted with the extremely well-defined rotational sequence in the Hyades (second panel from the bottom on the right). The currently accepted view is that after the age of the Hyades, stellar rotation periods lie on this converged sequence. When applying gyrochronology to stars younger than this age, the I and C sequences must be modelled separately.

Barnes (2003) pioneered an alternative approach to gyrochronology: an entirely empirical one where simple functional forms are fit to the observations (*e.g.* Barnes, 2007; Mamajek and Hillenbrand, 2008). Barnes (2003) proposed the following functional form for the relation between period, colour and age,

$$P = A^n \times a(B - V - c)^b, \quad (1.11)$$

where  $P$  is rotation period (in days),  $A$  is age (in Myr),  $B$  and  $V$  are B and V band magnitudes respectively and  $a$ ,  $b$ ,  $c$  and  $n$  are dimensionless free parameters. Barnes (2007) used rotation periods for stars in open clusters ranging in age from 30-600 Myr as well as the Sun to calibrate this relation.

Barnes and Kim (2010) present a hybrid part theoretical, part empirical gyrochronology relation:

$$\frac{dP}{dt} = \frac{\tau_c}{k_I P_{rot}}, \quad (1.12)$$

where  $k_I$  is a dimensionless constant, calibrated with observations. The convective

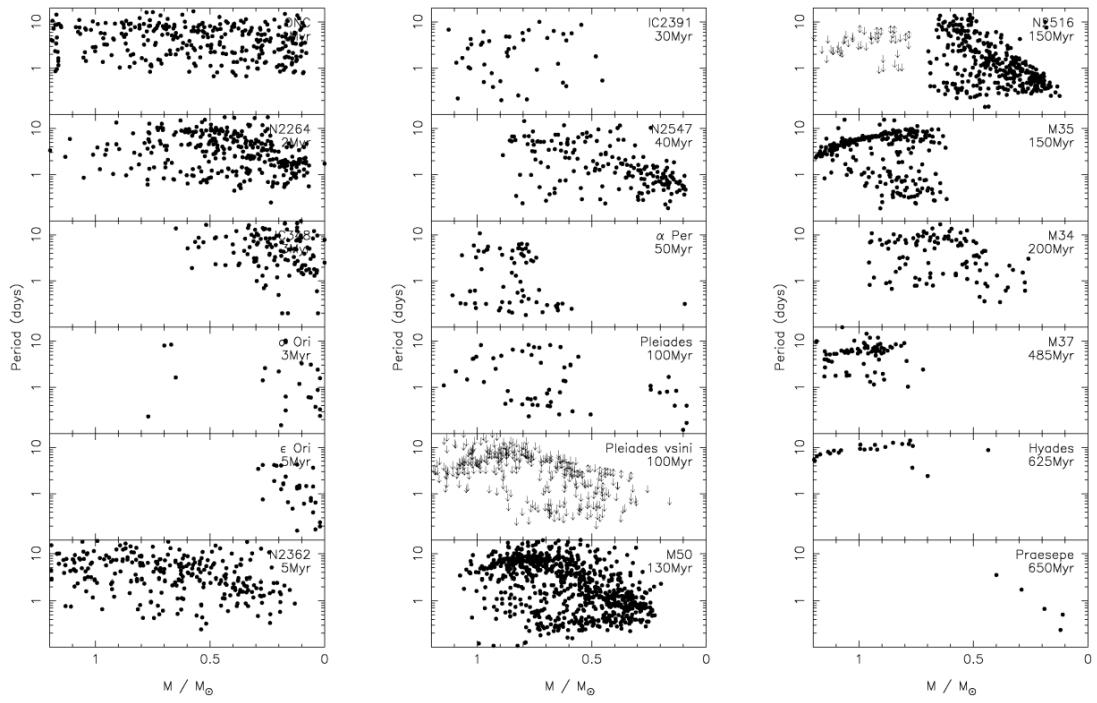


Figure 1-4: Figure from Irwin and Bouvier (2009). A compilation of most of the available rotation periods for stars in open clusters with  $M < 1.2M_{\odot}$  in 2009. Clusters increase in age from the top left to the bottom right.

overturn time,  $\tau_c$  encodes the dependence on stellar mass. In this thesis I use the Barnes (2003) functional form because the only available observables for the majority of *Kepler* stars are rotation period and colour. In the interest of using relatively model-independent parameters we avoid calculating  $\tau_c$ .

## 1.3 Other age diagnostics

Isochrone fitting, asteroseismology and gyrochronology are probably the most widely used dating methods today. There are a handful of others available however, which I briefly outline below.

### 1.3.1 Magnetic activity

Stars are born very active and become more inactive over time. As well as decaying over billions of years, activity also varies on several year timescales: the Sun, for example, has an 11 year activity cycle. The activity cycle period,  $P_{cyc}$  is related to rotation period,  $P_{rot}$  via  $P_{cyc} \approx (P_{rot}/\tau_c)^n$  where  $n \approx 1.5$  (Noyes et al., 1984a). Although stellar activity originates from one mechanism only, magnetism, it manifests in a variety of detectable ways, listed below.

- Star spots. Surface differential rotation, where the equator rotates at a different speed to the poles (in the Sun it rotates faster), twists magnetic flux into tubes which emerge from the stellar surface. The convection of material at the points where they emerge is inhibited and these areas become cooler and darker as a result. These cool, dark regions are called Sun spots on the Sun and star spots on the surfaces of other stars. Because these regions are darker, the integrated, optical flux emitted by a star when there are more spots on the surface is decreased. These decreases in brightness are detectable on the time scale of the stellar rotation period, where spots rotate into and out of view, and on longer timescales—the overall activity cycle.
- Chromospheric activity. Magnetic activity in stellar chromospheres produces

emission in the cores of singly ionised calcium H (3969.5) and K (3933.7) lines<sup>3</sup>. This emission reversal is produced by the excitation of  $\text{Ca}^+$  electrons to a higher energy level via magnetic heating. Just as with rotation period, chromospheric activity decreases with time and the two properties are related (*e.g.* Kraft, 1967; Noyes et al., 1984b). Relations between age and chromospheric activity have, just as with rotation period, been calibrated (*e.g.* Donahue, 1993; Lachaume et al., 1999; Mamajek and Hillenbrand, 2008; Soderblom et al., 1991). Chromospheric activity is usually quantified via the  $R/\nu_{HK}$  index, defined as the flux excess in the lines, normalised to the bolometric flux. The limitations of using chromospheric activity indices as age diagnostics are similar to the drawbacks of gyrochronology: the data are sparse and particularly so at late ages and low masses. Crucially though, measurements of Ca II H & K emission are difficult to obtain since high-resolution spectra are required ( $R \gtrsim 2000$ ).

- UV and X-ray flux. Magnetic heating in the chromosphere and corona excites photons to high energies, producing substantial UV and X-ray fluxes (*e.g.* Pallavicini et al., 1981). Since M dwarfs have large convective zones, some even being fully convective, they are highly active. Their UV flux is significant and this topic is currently attracting attention in the exoplanet community. M dwarfs with transiting exoplanets are premium targets for follow-up with the James Webb Space Telescope (JWST) since their petite size relative to their Earth-radius planets provides deep transits. The large S/N of these deep transits will allow JWST to search for biomarkers in the atmospheres of these planets. However, since M dwarfs are more active than G dwarfs, and their habitable zones are closer in, any potential life on an Earth twin may have been obliterated by the extreme UV flux.
- Flare rate. Occasionally magnetic flux tubes reconnect, causing acceleration of particles to enormous velocities, resulting in huge numbers of particles and amounts of energy being released into the Inter-Stellar Medium (ISM). These

---

<sup>3</sup>The H and K letters indicate the Fraunhofer designation. Fraunhofer assigned letters to the absorption lines he detected in the Solar spectrum.

events are called flares. The flare rate is related to the mean flare energy via a power law (Davenport, 2015; Hawley et al., 2014) and both of these properties are related to the strength of the stellar magnetic field. Flares occur extremely often in M dwarfs, the most magnetically active stars, with energies ranging from  $\sim e^{29} - e^{33}$  ergs and durations ranging from a few to a hundred minutes (Hawley et al., 2014).

### 1.3.2 Lithium depletion boundary

The Lithium Depletion Boundary (LDB) can be used as an age diagnostic for young, low mass ( $< 1M_{\odot}$ ) stars. As these young stars contract on the Pre-Main Sequence (PMS) their core temperature increases. When it reaches  $\sim 2.5 \times 10^6$  K, lithium is destroyed via  ${}^7\text{Li}(p, \alpha){}^4\text{He}$  and  ${}^6\text{Li}(p, \alpha){}^3\text{He}$  proton capture reactions (*e.g.* Bodenheimer, 1965). The time taken for a star to reach these temperatures depends on its mass. The lowest mass PMS stars are fully convective, so the mixing timescale is very short and lithium depletion takes place very rapidly. For young stellar groups therefore, the mass at which the LDB (the boundary between stars that do and do not show lithium in their atmospheres) is located is a very sensitive function of age (Basri et al., 1996). LDB dating can be very precise but it is only applicable to young groups of stars (20 Myr  $<$  age  $<$  200 Myr) since by this time Lithium has disappeared from the atmospheres of all members, regardless of their mass (Burke et al., 2004).

### 1.3.3 Dynamics

Stellar populations in the Milky Way galaxy can be localised into four groups: the thin disc (containing the Sun), the thick disc, the bulge, and the halo. Each of these stellar populations has a different distribution of stellar ages. The bulge is comprised of old stars and it is characteristically red because the only remaining stars on the MS are low-mass. The thin disc is comprised of young stars. This is the main star-forming part of the galaxy as the spiral arms, density waves of molecular gas that readily collapse into protostars, reside in the thin disc. The orbits of thin disc stars

get heated over time. Close encounters with neighbouring stars boost their orbital energies and many thin disc stars find themselves ‘kicked’ out of the galactic plane into the ‘thick disc’. The thick disc is statistically older than the thin disc since the majority of its residents have had time to experience one or more close encounters. The halo is even older still: these stars have been displaced even further from the thin disc in which they were born. Since the probability of experiencing a close encounter increases with time, stars with higher proper motions are likely to be older (see, *e.g.* Nissen and Schuster, 1991; Shevlev et al., 1989). However, although a rapidly moving star is likely to be older than a slowly moving star, stellar-orbit heating is a stochastic process and therefore cannot always be relied upon.

## 1.4 Kepler

The *Kepler* spacecraft plays a starring role in this thesis. *Kepler* was designed to survey over a hundred-thousand stars, searching for extrasolar planets via the transit method. Its ultimate goal was to answer the question "how common are Earth-like planets in our galaxy?". The first extra-solar planet orbiting a Sun-like star<sup>4</sup> was discovered using the radial velocity method, whereby the presence of a secondary body is inferred via the Doppler shift of the primary as it rotates about the system’s centre of mass (Mayor and Queloz, 1995). The first transiting exoplanet, HD 209458 b was discovered by Charbonneau et al. (2000) using ground-based photometry. Transits are produced when an exoplanet passes across the face of a star, blocking a small section of the photosphere and producing a small dip in the flux received at Earth. In 2009, the *Kepler* telescope was launched — a space mission dedicated to searching for exoplanets via the transit method (Borucki et al., 2010).

The radial velocity and transit methods remain the dominant methods of exoplanet detection. Together, they produce an impressive catalogue of confirmed exoplanets. Both methods, however, are subject to detection biases. The depth of an

---

<sup>4</sup>The first extra-solar planets were discovered orbiting a pulsar by observing timing variations (Wolszczan and Frail, 1992).

exoplanet transit is proportional to the ratio of stellar and planetary radii, squared. The larger the planet relative to its host, the larger the dip and the easier to detect. In addition, planets with short orbital periods are also more easy to detect due to the higher frequency of transit events. This makes large planets orbiting small stars with short orbital periods the easiest planets to detect. Several studies have attempted to characterise the completeness and detection efficiency of *Kepler* exoplanets in order to study their underlying distribution (*e.g.* Burke et al., 2015; Dressing and Charbonneau, 2015; Foreman-Mackey et al., 2014a; Petigura et al., 2013). The radial velocity (RV) method is more sensitive to planets with a large mass relative to their host star. It is also easier to detect planets around old stars via the RV method since magnetic activity can produce RV signals that can mask planet-induced Doppler shifting (*e.g.* Aigrain et al., 2012a; Rajpaul et al., 2015, 2016; Saar and Donahue, 1997; Wright, 2005).

An increasing number of planets have now been detected using direct imaging (*e.g.* Kalas et al., 2008; Nielsen et al., 2013). Using a coronagraph to block light from the host star and adaptive optical instruments to cancel out disturbance from the atmosphere, it is possible to observe emission from the surface of some planets at infrared wavelengths where the planet is brightest. It is only possible to directly image large exoplanets that are young and still cooling from the formation process, therefore bright in infrared. They should also have a large orbital separation from their host in order that star light can be blocked without blocking radiation from the planet. Direct imaging surveys are therefore currently limited in their detection opportunities. Yet another planet detection method is microlensing (*e.g.* Abe et al., 2004; Cassan et al., 2012; Gaudi, 2012; Gould et al., 2010). Occasionally, given near geometric alignment, light from a background (source) star may be gravitationally lensed by a foreground star. This process is detectable using time-series photometry if that foreground star has high proper motion — the source star appears to slowly brighten, then fade over a matter of days as the lens passes in front of it. If the foreground star has a planet it will produce an additional, lower magnitude and much shorter increase in flux as it too passes in front of the source. Microlensing exoplanet

surveys are again subject to different detection biases than other methods. To first order, detectability depends only on the mass ratio of planet and star and the sky-projected separation between the two (Clanton and Gaudi, 2016).

Of all these methods, thanks to *Kepler*, the transiting method has been the most fruitful to date. *Kepler* has discovered more than 5700 planet candidates and over 1000 confirmed planets. These numbers are still growing at a breath-taking rate and planets are likely to continue being discovered in the data well after the *Kepler* funding has dried up, thanks both to the enormity of the data set and the continual improvement of planet search methods. *Kepler* produced high precision<sup>5</sup> broad-band (peaking between V and R bands) light curves for  $\sim 100,000$  stars. The majority of these targets were observed in long-cadence mode (once every half-hour), and a few hundred in short cadence mode (once every minute), continually for around four years. The *Kepler* field is centred at RA = 19h 22m 40s, Dec =  $+44^{\circ}30' 00'$ , in the Cygnus region along the Orion arm of the galaxy. This field was chosen to be neither over nor under crowded, far enough out of the ecliptic that the Sun never shines onto the detector and contamination from Solar system objects is reduced. *Kepler* has a limited amount of on-board storage and must point towards the Earth to down-link data every month. These pointings break the time series. Other gaps appear at three month (quarter) intervals. The spacecraft rotates every quarter to keep its Solar panels pointed at the Sun. The stars shift to new CCD modules every time this happens, eventually returning to the same module after one year. Both the down-link pointings and quarter rotations produce short gaps in the light curves and can also produce short-term temperature changes to the CCD which can affect the sensitivity of the detector. Temperature changes increase the gain of the CCD chip: more electrons are excited when the CCD is warmer so the flux appears to increase. So despite the unprecedented precision of *Kepler* lightcurves, they are not uninterrupted and contain systematic features. A typical *Kepler* light curve is shown in figure 1-6. Understanding the systematics in *Kepler* data is necessary for *all* of

---

<sup>5</sup>Precision ranging from a few hundred parts per million for bright ( $< 13$ th magnitude) targets to tens of thousands of ppm for faint ( $> 17$ th magnitude) targets.

*Kepler*'s science goals and is a key part of this thesis. This point is covered further in §1.5.2.

Unfortunately, that key question—"how common are Earth-like planets?"—may never be answered by *Kepler* (or at least not very precisely. Several inferences have been performed by extrapolation, *e.g.* Burke et al., 2015; Foreman-Mackey et al., 2014a; Petigura et al., 2013). Although *Kepler* launched with four functioning gyroscopic reaction wheels, one broke shortly thereafter. This was not a mission-ending scenario since the spacecraft only needed three to maintain its precise pointing: one for pitch, one for roll and one for yaw. Unfortunately however a second mission-critical reaction wheel broke in 2013. Without all three reaction wheels, precise pointing of the spacecraft was impossible to maintain. The extreme photometric precision of *Kepler* comes from its exquisite pointing position. With stars fixed in place, moving by less than a few milli-arcseconds per quarter, even without knowing the exact Point Spread Function (PSF) and Pixel Response Function (PRF), it is possible to perform precise relative photometry. With its pointing stability compromised, the community were asked for input for a repurposed *Kepler* mission (*e.g.* Aigrain et al., 2015a; Hogg et al., 2013).

*Kepler* was re-purposed as the *K2* mission in 2014. With just two reaction wheels controlling pitch and yaw, its third rotation axis had to be stabilised. This was achieved by balancing the spacecraft against Solar radiation, using its symmetric solar-panelled back. The spacecraft sits in an unstable equilibrium, drifting slowly about its roll axis, its new orientation reducing large forces from the Solar wind (Howell et al., 2014). The spacecraft's thrusters are used to correct the slow drift. Although the precision is not what it once was, stars drift across several pixels before the thrusters correct the motion, it is still good enough to obtain huge amounts of useful data. In order to keep its back to the Sun, *Kepler* can only point at fields in the ecliptic plane. This means that new stellar populations can be explored: from different galactic regions, the bulge, thin and thick discs and halo, to open and globular clusters. Crucially, from the point of view of stellar astronomy, *Kepler*'s new fields incorporate several open clusters. Open clusters are wonderful labs for stellar astronomy as they

are coeval populations of stars made from the same material. They are controlled environments in which we can study the observational properties of stars as a function of their mass. Figure 1-7 shows the current fields being observed by the spacecraft and figure 1-8 shows the future fields being proposed.

Not only is the spacecraft continuing its search for exoplanets (and has already discovered many, *e.g.* Becker et al., 2015; Crossfield et al., 2015; Foreman-Mackey et al., 2015; Montet et al., 2015; Vanderburg et al., 2016, 2015), it will provide time series of white dwarfs, active galactic nuclei, red giants, red dwarfs, binary stars and other interesting astronomical objects. Of course, with the reduced pointing position comes a new challenge for those astronomers willing to get their hands dirty: systematic features in *K2* light curves range from negligible for the brightest targets to overwhelming for the faintest. It is therefore necessary to employ heavy-duty ‘detrending’ methods to exploit *K2*’s rich data set. This subject is covered in §1.5.2

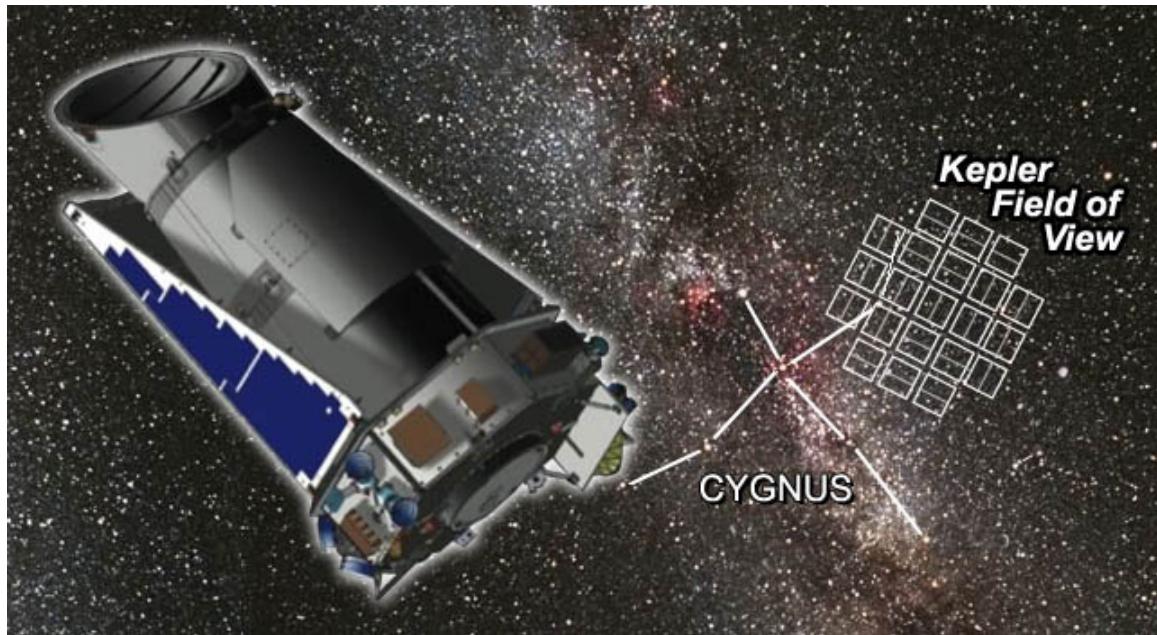


Figure 1-5: A depiction of the *Kepler* spacecraft and its field of view. The *Kepler* field is located near the constellation Cygnus. Image credit: [rst.gsfc.nasa.gov/Sect20/A11.html](http://rst.gsfc.nasa.gov/Sect20/A11.html)

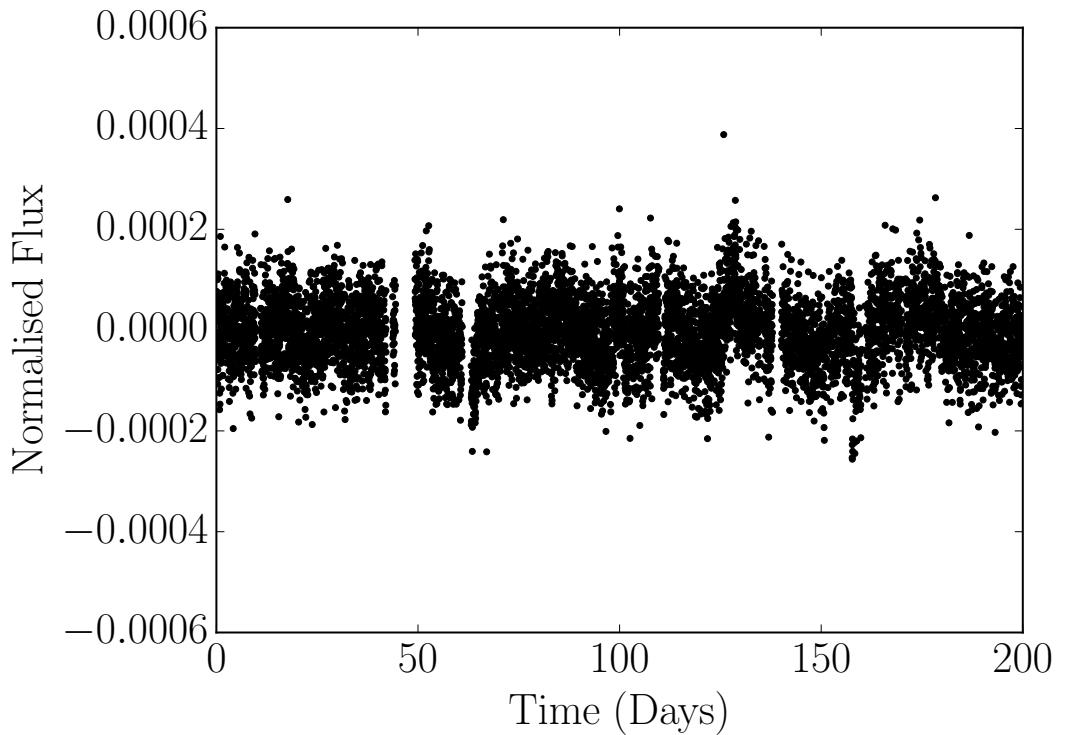


Figure 1-6: An example *Kepler* light curve. In this figure, normalised flux versus time is shown for the first 200 days of typical *Kepler* light curve, KID 2450729. This is a relatively quiet star, although stellar variability is still present at a low level. Gaps in the data occur during quarterly spacecraft repointings and monthly data downlinking.

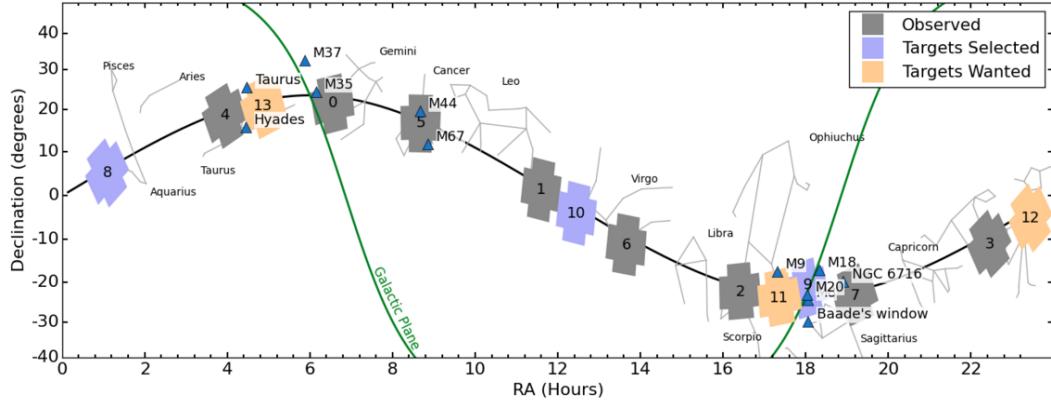


Figure 1-7: *K2*'s fields. Fields 1-9 have already been observed and field 9 is being observed at the time this thesis is handed in. Targets have been fixed for purple fields and proposals are currently being solicited for yellow fields.

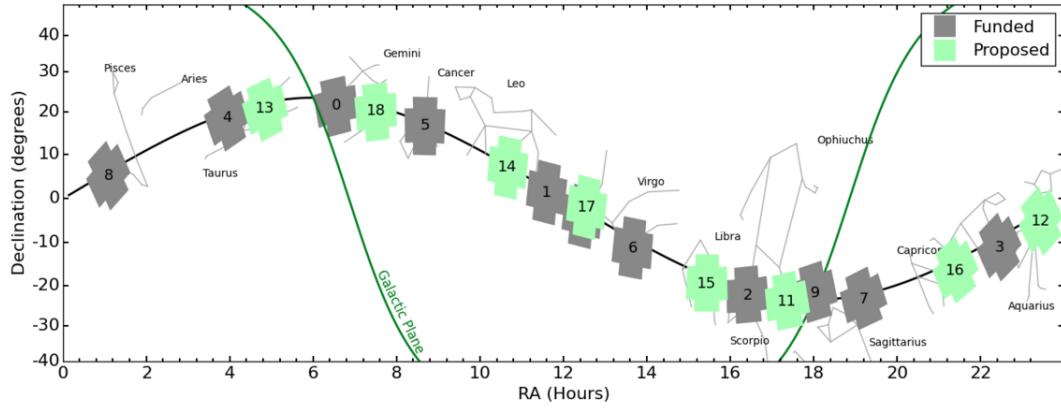


Figure 1-8: *K2*'s current and future fields. These green fields are those proposed if an extended *K2* mission is funded.

## 1.5 Statistical methods

A theme continued throughout this thesis is probabilistic statistical methods. None of the methods used here are new in themselves, but many have not been previously used on the sorts of problems presented here, or in an astronomical context.

### 1.5.1 Multi-dimensional and unknown uncertainties: hierarchical probabilistic modelling

In some astronomical data sets, only the dependent variable has significant uncertainties—for example in time-series or spectra—in which case it is reasonable to ignore independent variable uncertainties. When performing regression on data with only  $y$ -direction, Gaussian uncertainties where the measurements are independent and there is no correlated noise, it may be appropriate to use a simple Gaussian likelihood function. For example, the likelihood of  $y$ -values given some  $x$ -values, some model parameters,  $\theta$  and some  $y$ -direction uncertainties,  $\sigma$ , can be written

$$p(\{y\}_{i=1}^N | \{x\}_{i=1}^N, \{\sigma\}_{i=1}^N, \theta) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{[y_i - \mu(x_i, \theta)]^2}{2\sigma_i^2}\right), \quad (1.13)$$

where  $\mu$  is the mean model. It is often more practical to use the logarithm of this function,

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^N \left[ \frac{[y_i - \mu(x_i, \theta)]^2}{\sigma_i^2} + \ln(2\pi\sigma_i^2) \right]. \quad (1.14)$$

In many fitting-a-model-to-data problems in the astronomical literature, a frequentist approach is taken and  $\chi^2$  is minimised (or log-likelihood is maximised). The probabilistic (Bayesian) approach however, is to multiply this likelihood by a prior to produce a posterior Probability Density Function (PDF):

$$p(\theta | \{y\}_{i=1}^N, I) = \frac{p(\{y\}_{i=1}^N | \theta, I)p(\theta, I)}{p(\{y\}_{i=1}^N | I)}. \quad (1.15)$$

In this equation,  $I$  is used to represent all of one's knowledge and implicit assumptions

made about the data and the model (it is dropped hereafter). The prior,  $p(\theta, I)$  represents one's prior beliefs about the model parameters,  $\theta$ . There is a wide range of literature available on the science behind choosing priors which is beyond the scope of this thesis (*e.g.* Gelman, 2009; Kass and Wasserman, 1996; VanderPlas, 2014). The denominator in the above equation is the likelihood of the data, marginalised over all model parameters. It is sometimes known as the ‘fully marginalised likelihood’ and sometimes the ‘evidence’. It is a useful quantity to compute when performing model comparison but not is necessary for performing regression. I do not perform model comparison here so a discussion of the evidence integral is beyond the scope of this thesis.

Returning to the equation for the log-likelihood, equation 1.14, it is trivial to modify this function for the case where the  $y$ -direction uncertainties are unknown, under, or over-estimated. In this case one could write

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^N \left[ \frac{[y_i - \mu(x_i, \theta)]^2}{\sigma_i^2 + s^2} + \ln(2\pi(\sigma_i^2 + s^2)) \right], \quad (1.16)$$

where  $s^2$  is the additional variance needed to explain the scatter in the data.

In chapter 5 a modification to the above equation is used to fit a line to data where the uncertainties are underestimated. This is a hierarchical problem since there are two parameters that describe the model (a straight line),  $\alpha$  and  $\beta$ , where  $\mu = \alpha + \beta x$  and a parameter describes the standard deviation of a Gaussian spanning the mean model. In other words, the ‘true’ value of variable  $y$ —the value that would have been measured if there were no observational uncertainties— $\bar{y}$ , is drawn from a Gaussian with mean,  $\mu = \alpha + \beta \bar{x}$  and standard deviation,  $s$ :

$$\bar{y} \sim \mathcal{N}(\alpha + \beta \bar{x}, s^2), \quad (1.17)$$

and the *observations*,  $y_{obs}$  are in turn drawn from Gaussians with  $\mu = \bar{y}$  and standard

deviations described by the individual observational uncertainties,  $\sigma_{obs}$ :

$$y_{obs} \sim \mathcal{N}(\bar{y}, \sigma_{obs}^2). \quad (1.18)$$

Since it is the observations that are modelled, not the latent parameter,  $\bar{y}$ , this is a hierarchical process. In general, an inference problem can be described as hierarchical when there are multiple ‘layers’ of conditional dependencies. In this example there are two such ‘layers’: the dependence of the observations on the true values and the dependence of the true values on the model parameters.

Equation 1.14 can also be modified for the simple case of fitting a straight line to data, where the  $y$ -direction observational uncertainties are unknown *and* the  $x$ -direction uncertainties are non-negligible. Starting with equations defined in chapter 7 of Hogg et al. (2010a),  $\mathbf{S}_i$  is the covariance tensor,

$$\mathbf{S}_i \equiv \begin{pmatrix} \sigma_x^2 & \sigma_{x,y} \\ \sigma_{x,y} & \sigma_y^2 \end{pmatrix}, \quad (1.19)$$

$\hat{\nu}$  is a unit vector orthogonal to the line with slope  $\beta$ ,

$$\hat{\nu} = \frac{1}{\sqrt{1 + \beta^2}} \begin{pmatrix} -\beta \\ 1 \end{pmatrix}, \quad (1.20)$$

the variance of the data is given by,

$$\Sigma_i^2 = \hat{\nu}^T \mathbf{S}_i \hat{\nu} \quad (1.21)$$

and the log-likelihood is,

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{i=1}^N \left[ \frac{\Delta_i^2}{\Sigma_i^2} + \ln |\mathbf{S}_i| \right]. \quad (1.22)$$

By modifying equation 1.19 to be

$$\mathbf{S}_i = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 + s^2 \end{pmatrix}, \quad (1.23)$$

since we assume that the covariance between  $x$  and  $y$  is negligible, substituting 1.23, 1.20 and 1.21 into 1.22 gives

$$\ln(\mathcal{L}) = -\frac{1}{2} \sum_{i=1}^N \left( \frac{[y_i - (\alpha + \beta x_i)]^2}{\beta^2 \sigma_{x,i}^2 + \sigma_{y,i}^2 + s^2} + \ln[\sigma_{x,i}^2 (\sigma_{y,i}^2 + s^2)] \right). \quad (1.24)$$

This is the equation used in chapter 5 to analyse astrophysical data with intrinsic scatter.

In cases with more complicated mean models, it is not so easy to arrive at an analytic solution. For example, a similar problem is approached in chapter 2. Here, the extra variance parameter,  $s$ , is not used (although it should be included in future analyses!), however the uncertainties in the  $x$ ,  $y$ , and this time  $z$ , directions are large. The model is not a simple two-dimensional line as in the above case, it is the three-dimensional gyrochronology equation 2.1 where  $y$  is rotation period,  $x$  is age and  $z$  is B-V colour. In this case, an approximation must be made in order to take the uncertainties on all dimensions into account. This approximation is made via sampling and is described in chapter 2.

### 1.5.2 Methods for removing systematics

‘Detrending’ is a word that has been adopted by astronomers to mean removing systematic trends caused by non-physical (*i.e.* instrumental) variations in the experiment conditions. This process can be applied to any data set that contains correlated noise, be it a time series, image or spectrum. However ‘detrending’ usually pertains specifically to time series analysis, and light curves in particular. The detrending process involves generating some model of the systematic features which is then subtracted from the light curve. This implies that the contribution of systematics to the light

curve are known *a priori*. Of course in reality, the contributions coming from the signal and the noise can never be separated completely, so this is an approximate process. In some cases this approximation is close to the truth, or close enough—for example, if signal and noise look very different, rough detrending may still reveal useful signals in the data. However, in many cases the noise model cannot be adequately approximated. At best this leads to inaccurate inferences about the physical system being studied and at worse it leads to false positive detections and false negative non-detections.

In §3 I present a method for extracting periodic signals from *K2* light curves, without detrending. Here I summarise the detrending process for the original *Kepler* mission.

*Kepler* data is available to download at various stages of the reduction process. Target Pixel Files (TPFs), are arrays of the raw pixel fluxes, Simple Aperture Photometry (SAP) light curves are generated by combining pixel fluxes with simple masks and Presearch Data Conditioned, Maximum *A Posteriori* (PDC-MAP) light curves, are a detrended data product.

The detrending process concerns removing as much of the noise as possible whilst preserving the signal (Smith et al., 2012; Stumpe et al., 2012). In reality no detrending process does this perfectly and most favour a slightly more or less aggressive approach to removing systematics, depending on the science goal. Systematic features in *Kepler* light curves are generated by two main sources: pointing variations and temperature variations (although pointing variations are more serious for *K2* light curves than *Kepler* light curves). Regardless of the source, either of these variations will effect stars that lie close together on the CCD in similar ways. The *Kepler* pipeline uses this fact to isolate physical signals from instrumental: physical signals vary from star to star and instrumental signals are present in nearby groups of light curves. Intrinsically quiet stars that are nearby, but display similar patterns in variation, are used to generate a set of Cotrending Basis Vectors (CBVs) via Single Value Decomposition (SVD). A detrending model, constructed from a linear combination of CBVs is then fit to these stars and the distribution of the CBV weights are used to generate a prior.

A systematics model is then fit to each individual light curve, where a likelihood function is multiplied by the prior and the resulting maximum of the posterior PDF is taken to be the best fit. That systematics model is then subtracted from the data.

PDC-MAP light curves were designed to maximise exoplanet transit search capability. These signals are rarely longer than thirteen hours and have a characteristic upside-down top-hat shape. Stellar variability on the other hand is smoothly varying (similarly to systematic features) with timescales ranging from around one day to several years. PDC-MAP detrending does not preserve signals longer than  $\sim 30$  days.

An alternative approach to PDC-MAP detrending, designed to preserve signals from the stars as *well* as the planets is that of Roberts et al. (2013). They also fit a linear combination of CBVs to the data, but the main differences in their approach are to impose a maximum entropy criterion: trends in the most highly-weighted CBVs must be present in a large number of light curves, and to remove high frequency noise from the systematics model before detrending so as not to inject high frequency noise. A comparison of the two detrending methods is presented in Roberts et al. (2013).

Once *Kepler* lost its third reaction wheel, pointing variations became much more severe, systematic features rose in amplitude and more aggressive detrending techniques became necessary. In chapter 3, I introduce an alternative to detrending, specifically applicable to *K2* light curves.

### 1.5.3 Rotation period inference

The overall brightness of stars is periodically reduced by spots on their surfaces. A measurement of the time between successive dimmings provides an indication of the surface rotation period. Unfortunately, inferring accurate rotation periods is not as simple as finding the period of the best-fit sinusoid: the picture is complicated by several factors:

- The presence of *several* dark spots and plages (bright spots), with complicated geometries, on the surface at any given time creates non-sinusoidal brightness

variations.

- Spots are born, grow, then shrink and eventually fade away. The finite lifetimes of spots produces evolving patterns in light curves. Furthermore, spots are born at a range of longitudes, meaning that brightness variations are not perfectly periodic.
- Many stars exhibit signs of differential rotation, *i.e.* the poles and the equator rotate at different angular velocities. Star spots may be present at multiple latitudes simultaneously which will blur the periodicity of the signal—it will have a finite width in frequency space, rather than appearing as a delta function. If there are two dominant active regions at different latitudes, two distinct periodic signals may be present in the light curve. The famous butterfly diagram for the Sun demonstrates that solar spots are born at active latitudes, bands around  $30^\circ$  in width, between  $0$  and  $30^\circ$  north and south of the equator (*e.g.* Charbonneau, 2010). Spots are born at lower and lower latitudes over the course of one Solar cycle, therefore if the Sun were a *Kepler* star one would measure a different rotation period from one year to the next.

In this thesis I will take ‘rotation period’ to mean the equatorial period. For stars with significant differential rotation, this will be the minimum rotation period (assuming that equatorial rotation is faster than polar rotation, as in the Sun). In reality therefore, all photometric rotation periods are just upper limits. This is an interesting point for a future discussion but is beyond the scope of this thesis.

Aigrain et al. (2012b) and Dumusque et al. (2014) demonstrate the effects of spots on the flux variations of stars and provide realistic models for translating flux signals into Radial Velocity (RV) signals<sup>6</sup>. These models include parameters such as period, inclination, number of spots, spot lifetime, spot temperature, differential rotation amplitude (shear), and others. They are able to produce light curves that look very similar to real *Kepler* time-series, but fitting such models to data is not

---

<sup>6</sup>There is now a large body of (occasionally controversial) literature on this topic in the context of exoplanet search. Star spots signals can look like exoplanet signals in RV time series. Modelling star spots has now become a necessity for exoplanet RV search.

straightforward. It is impractical to have a set of parameters (*e.g.* latitude, longitude, size, temperature) for each spot, but it is difficult to simplify the model in a way that will allow enough flexibility to fit a real data set. It is therefore necessary to use alternative methods for inferring rotation periods. In chapter 4 of this thesis I present a new method for inferring rotation periods, and give a summary of the main existing methods below.

### Lomb-Scargle periodograms

Lomb-Scargle (LS) periodograms (Lomb, 1976; Scargle, 1982) were developed by astronomers to perform frequency analysis on unevenly spaced data. To produce a LS periodogram, a sinusoid is fit to a time series via least-squares over a grid of frequencies. The amplitudes of the sinusoid are plotted against frequency to produce a power spectrum. This is simple linear regression and LS periodograms are therefore inexpensive to compute, although more costly than Fast Fourier Transforms (FFTs). LS periodograms have been used to study periodic phenomena in astrophysical time series for decades and were the first tools used to measure photometric rotation periods (*e.g.* Mottola et al., 1995; Scott et al., 1992). They are often used to infer rotation periods from *Kepler* light curves (*e.g.* Reinhold and Reiners, 2013; Reinhold et al., 2013). The main drawback of the LS periodogram for rotation period measurement is that a sinusoid is not necessarily a good model for the kinds of signals produced by star spots and results in imprecise period measurement. This is not a problem however, for the ACF method.

### The ACF method

An AutoCorrelation Function (ACF) is often used to measure stellar rotation periods (*e.g.* Aigrain et al., 2008; García et al., 2014; McQuillan et al., 2013a,b, 2014). A method was developed specifically for *Kepler* data by McQuillan et al. (2013a). The ACF itself is defined as

$$r_k = \frac{\sum_{i=1}^N (x_i - \bar{x})(x_{i+k} - \bar{x})}{\sum_{i=1}^N (x_i - \bar{x})^2}, \quad (1.25)$$

where  $r_k$  is the autocorrelation coefficient at lag  $k$  for a time series with elements,  $x_i$  ( $i = 1, \dots, N$ ). An ACF is only applicable to time series with evenly spaced data (which is usually an acceptable approximation for *Kepler* light curves) and a Fourier transform of an ACF is the power spectral density.

Unlike the LS periodogram which uses a sinusoid to fit the data, because the ACF does not rely on a functional form—it simply looks for repeating patterns—it is better suited to measuring stellar rotation periods, particularly for relatively inactive stars where long time series are required to detect the weak rotational signal (see McQuillan et al., 2014, for an in-depth discussion). A series of uniformly spaced peaks will be visible in the ACF of a light curve with a rotational signal. The position of the first peak corresponds to the lag-time of greatest correlation. This peak is usually adopted as the rotation period. However, if the second peak is higher than the first, *that* will be taken as the rotation period. This is usually caused when there are two active regions on opposite hemispheres of the star,  $\sim 180^\circ$  apart, producing two dips in the light curve per rotation period. In cases where stars have very long-lived active regions, light curve dips produced by the two active regions will remain stable and the light curve will maintain a double-dip over long timescales. In these cases, the first peak in the ACF (corresponding to half the true period) may be higher than the second and half the true period would be measured. Unfortunately, we do not know the distribution of spot lifetimes and therefore cannot provide an estimate of harmonic contamination. The ACF method was used by McQuillan et al. (2014) to measure rotation periods of 34,030 *Kepler* targets.

The main drawback of the ACF method is that uncertainties are not well defined. ACF rotation periods with uncertainties reported in the literature are typically measured by fitting a Gaussian to the peaks in the ACF. However, the width of the ACF peak is related to the rate of correlation coefficient fall-off which depends on the shape of the signal and cannot be interpreted as a measurement uncertainty,  $\sigma$ .

## Wavelets

Wavelet transforms are used by García et al. (2014) to measure rotation periods for the *Kepler* asteroseismic targets used in chapter 2. Wavelet functions can be thought of as abscissa-dependent periodograms (Carter and Winn, 2009). Like the LS periodogram, a model is compared to the data over a range of frequencies (scales). However, unlike the LS periodogram, this is performed for a range of time-displacements. A wavelet transform therefore provides an indication of the location of periodic signals in a time-series. There are many functional forms, or ‘mother wavelets’ which can be used. García et al. (2014) used a Haar wavelet which is a sinusoid convolved with a Gaussian. They also used the ACF method to confirm their rotation period detections. The wavelet method has many of the same drawbacks as the LS periodogram method: stellar signals do not necessarily look like a Haar wave-form and their shapes evolve over time.

## Spectroscopy

Before the advent of high-precision space photometry, the majority of available rotation periods were measured spectroscopically. Doppler shifting of the rotating stellar surface causes broadening of spectral lines which can be used to infer the equatorial velocity, multiplied by the sine of the inclination angle,  $v \sin(i)$ . Not only does this method produce periods that are degenerate with inclination angle<sup>7</sup>, it also relies on the precision of the spectrograph, the magnitude of the target, the rotation speed and thus produces rotation periods with extremely large uncertainties. Photometric rotation periods are almost always preferable to spectroscopic ones, especially for slow rotators.

In this thesis both the ACF and LS periodogram methods are used for rotation period inference. However, as explained above, both these methods are flawed. In chapter 4 I introduce a new method for rotation periods that does not suffer from the aforementioned flaws. Specifically, it is capable of modelling quasi-periodic light

---

<sup>7</sup>This degeneracy can be resolved for populations of stars since one can assume that the population is isotropically oriented (e.g. Andrews et al., 2014).

curves of any shape, and provides accurate uncertainties. I explain the details in chapter 4. Before that however, in chapter 2, I present the results of my work on gyrochronology. I then present a method for detecting periodic signals in *K2* light curves in chapter 3, followed by the aforementioned method for probabilistic rotation period inference in chapter 4. In chapter 5 I present my work on the probabilistic calibration of the relation between short-term stellar brightness variability (known as flicker) and stellar density and surface gravity. Finally, I present an ongoing study on the potential for gyrochronology with the up-and-coming Large Synoptic Survey Telescope (*LSST*) and its rotation period yield in chapter 6.

## Chapter 2

# Calibrating Gyrochronology using Kepler Asteroseismic targets

### Abstract

Among the available methods for dating stars, gyrochronology is a powerful one because it requires knowledge of only the star's mass and rotation period. Gyrochronology relations have previously been calibrated using young clusters, with the Sun providing the only age dependence, and are therefore poorly calibrated at late ages. We used rotation period measurements of 310 *Kepler* stars with asteroseismic ages, 50 stars from the Hyades and Coma Berenices clusters and 6 field stars (including the Sun) with precise age measurements to calibrate the gyrochronology relation, whilst fully accounting for measurement uncertainties in all observable quantities. We calibrated a relation of the form  $P = A^n \times a(B - V - 0.4)^b$ , where  $P$  is rotation period in days,  $A$  is age in Myr,  $B$  and  $V$  are magnitudes and  $a$ ,  $b$  and  $n$  are the free parameters of our model. We found  $a = 0.40_{-0.05}^{+0.3}$ ,  $b = 0.31_{-0.02}^{+0.05}$  and  $n = 0.55_{-0.09}^{+0.02}$ . Markov Chain Monte Carlo methods were used to explore the posterior probability distribution functions of the gyrochronology parameters and we carefully checked the effects of leaving out parts of our sample, leading us to find that no single relation between rotation period, colour and age can adequately describe all the subsets of our data. Contrary to predictions, the *Kepler* asteroseismic stars, cluster stars and local field stars cannot all be described by the same gyrochronology relation. The *Kepler* asteroseismic stars may be subject to observational biases, however the clusters show unexpected deviations from the predicted behaviour, providing concerns for the overall reliability of gyrochronology as a dating method.

# Introduction

## Dating methods for field stars

Many fields of astronomy rely on precise age measurements of Main Sequence (MS) stars. Unfortunately, age is a notoriously difficult quantity to measure for these stars, as observable properties evolve slowly on the MS. Even with high precision spectroscopic measurements, ages often cannot be determined accurately to within 20% (Soderblom, 2010). Some of the most precise age measurements currently available are for stars in clusters where isochrones can be fitted to a coeval population with a range of masses, resulting in age measurements with uncertainties often as low as 10%. Isochronally derived *field* star ages, on the other hand, are much less precise than this, often having uncertainties of around 50% or more. Demand for age estimates of planet-hosting stars is high, but faint stars observed by *Kepler* are often expensive or impractical spectroscopic targets. Where high resolution spectra are unavailable, gyrochronology can be extremely useful. Gyrochronology is a dating method that utilises the potentially predictable rotation period evolution of low mass, MS stars. It requires only knowledge of the current rotation period—which is often easily extracted from *Kepler* light curves—and mass (or appropriate proxy) of a star. However the current gyrochronology relations are entirely empirically calibrated and still need refining at large stellar ages. *Kepler* provides the perfect opportunity to calibrate gyrochronology at late ages—it provides surface rotation periods of thousands of stars and new age estimates for hundreds of stars via asteroseismology<sup>1</sup>. This paper aims to use these new asteroseismic age measurements to improve the gyrochronology relations at late ages.

---

<sup>1</sup>Note that asteroseismic ages are only available for those *Kepler* stars which display high signal-to-noise Solar-like oscillations in their power spectra. The majority of stars that fall into this category are around the same temperature as, and slightly hotter than, the Sun

### 2.0.1 Gyrochronology

Mass loss via a magnetised stellar wind causes magnetic braking of MS stars (Weber and Davis, 1967). Observational evidence suggests that although stellar populations are born with a range of rotation periods, the rapid rotators rapidly lose angular momentum and rotation periods converge onto a well-defined sequence. Gyrochronology postulates that each star falls on a single three-dimensional plane described by mass, rotation period and age, i.e. given any two of these three properties, one can determine the third. The form of angular momentum evolution described above and calibrated in this chapter can only be applied to F, G and K MS stars. Fully convective M dwarfs have a different dynamo-driven magnetic field: their rotation periods evolve over extremely long timescales and they often do not converge onto the mass-period-age plane, even after several Gyrs. Hot stars with effective temperatures greater than  $\sim 6250$  K have shallow convective zones—they are almost fully radiative. Since bulk magnetic fields in MS stars are produced by the combination of convection and rotation, without a convection zone, these stars have weak fields (Kraft, 1967). Hot stars retain their initial rotation period throughout their brief MS lifetimes and are therefore not suitable gyrochronology targets.

The term ‘gyrochronology’ was coined by Barnes (2003) who proposed an empirically motivated functional form for the relation between period, colour and age,

$$P = A^n \times a(B - V - c)^b, \quad (2.1)$$

where  $P$  is rotation period (in days),  $A$  is age (in Myr),  $B$  and  $V$  are B and V band magnitudes respectively and  $a$ ,  $b$ ,  $c$  and  $n$  are dimensionless free parameters.

This gyrochronology relation was calibrated using open clusters, which are invaluable calibration tools since their ages are known relatively precisely and each cluster contains many stars of the same age which enables the period-mass relation to be tightly constrained. Unfortunately however, the majority of nearby clusters are young and until recently it was difficult to measure rotation periods for all but the youngest, most active stars (using ground-based observations). There is a significant

Table 2.1: Values of  $a$ ,  $b$ ,  $c$  and  $n$  in Barnes (2007) and Mamajek and Hillenbrand (2008) and Angus et al. (2014).

Parameter	Barnes (2007)	Mamajek and Hillenbrand (2008)	Angus et al. (2014)
$a$	$0.7725 \pm 0.011$	$0.407 \pm 0.021$	$0.40_{-0.05}^{+0.3}$
$b$	$0.601 \pm 0.024$	$0.325 \pm 0.024$	$0.31_{-0.02}^{+0.3}$
$c$	0.4	$0.495 \pm 0.010$	0.45
$n$	$0.5189 \pm 0.0070$	$0.566 \pm 0.008$	$0.55_{-0.09}^{+0.02}$

dearth of precisely measured ages for old stars and it is for this reason that the current gyrochronology relations are poorly calibrated at late ages. Barnes (2007) used 8 young open clusters aged between 30 and 650 Myrs to calibrate the dependence of rotation period on mass, and the Sun to calibrate the dependence on age. Best-fit values of  $n$ ,  $a$  and  $b$  ( $c$  was fixed at 0.4) reported in Barnes (2007) are presented in table 2.1. Equation 2.1 was further calibrated by Mamajek and Hillenbrand (2008) using updated rotation period and age measurements of stars in open clusters  $\alpha$  Per (Prosser et al., 1995), Pleiades (Krishnamurthi et al., 1998; Prosser et al., 1995), M34 (Meibom et al., 2011b), and Hyades (Henry, private comm., Paulson et al., 2004; Prosser et al., 1995; Radick et al., 1995, 1987). Once again, the Sun was used as an age anchor—a single data point specifying the shape of the period-age relation. Whereas Barnes (2007) fixed the position of  $c$ , the ‘colour discontinuity’ in equation 2.1, at 0.4, Mamajek and Hillenbrand (2008) allow it to be a free parameter in their model. The values of  $n$ ,  $a$  and  $b$ , resulting from their fit are shown in table 2.1. In both of these studies a maximum likelihood fitting approach was used. This method relies on the assumption that uncertainties are Gaussian, which may not always be the case, and only takes observational uncertainties on the dependent variable into account. As described in §2.2, we adopt a fitting method that properly accounts for uncertainties on all three observed variables: colour, period and age.

The data used in this chapter are described in §2.1, our calibration and model fitting process is outlined in §2.2 and the results are presented and discussed in §2.3.

## 2.1 Observations

The ages of 505 *Kepler* dwarfs and subgiants were published in Chaplin et al. (2014). They made use of two global asteroseismic parameters—the average large frequency separation and the frequency of maximum oscillations power—to estimate stellar properties, including the ages, with a grid-based approach that utilised several different search codes coupled to more than ten grids of stellar evolutionary models.

The ages quoted in Chaplin et al. (2014) come from one of the grid-code combinations, with uncertainties reflecting the scatter between the different sets of results. Chaplin et al. (2014) used two different sets of effective temperatures: one was derived using an Infra-Red Flux Method (IRFM) calibration (Casagrande et al., 2010; Silva Aguirre et al., 2012b) and the other from a recalibration of the SDSS griz filter KIC photometry by Pinsonneault et al. (2012) using Yale Rotating Stellar Evolution Code (YREC) models (Demarque et al., 2004). We use the IRFM temperatures since they are less dependent on metallicity, which is not well constrained for the asteroseismic sample, and their uncertainties are more conservative, but our analysis is relatively insensitive to this choice. 87 stars in the asteroseismic catalogue have spectroscopic measurements of  $T_{\text{eff}}$ , and [Fe/H]. These precisely measured spectroscopic properties allowed more tightly constrained ages to be calculated for these 87 stars, which were incorporated where available. In order to produce a relation that predicts the age of a star using only observable properties, we chose to convert  $T_{\text{eff}}$  to  $B - V$  for the asteroseismic sample using the relation of Sekiguchi and Fukugita (2000). This conversion added an extra element of systematic uncertainty to our data since the metallicities provided for most of the asteroseismic stars are simply an average value for the field:  $-0.2 \pm 0.3$  dex (see e.g. Silva Aguirre et al., 2011). However, since the age uncertainties dominate this analysis, we do not expect this to have a significant impact on our results.

The ages quoted in Chaplin et al. (2014) have typical uncertainties of 35%. These large uncertainties are the result of the fact that only approximate inferences can be made on the ages using the global asteroseismic parameters, neither of which has an

explicit dependence on age. It will be possible, however, to derive more precise ages for a subset of these stars. By measuring the frequency of each oscillation mode individually, not just the global asteroseismic parameters, one can provide much tighter constraints on ages. Ages derived from individual oscillation mode measurements can have uncertainties as small as 10% (Brown and Gilliland, 1994; Silva Aguirre et al., 2013). However, measuring frequencies for individual oscillation modes is a manual process and can only be applied in the highest signal-to-noise cases. Chaplin et al. (2014) predict that around 150 of the 505 stars will be suitable for this individual oscillation mode treatment. We obtained precise ages for 42 stars from Metcalfe et al. (2014), modelled with the Asteroseismic Modeling Portal (AMP, Metcalfe et al., 2009; Woitaszek et al., 2010), with effective temperatures and metallicities from Bruntt et al. (2012). Of the 42 stars in Metcalfe et al. (2014), we only incorporate the ‘simple stars’ (cool dwarfs) into our sample, ignoring the hotter F stars and more evolved subgiant stars as these are not expected to follow the simple gyrochronology relation.

The *Kepler* light curves of the 505 asteroseismic targets display quasi-periodic variations on timescales corresponding to the rotational periods of the stars—flux variations are produced by active regions on the stellar surface that rotate in and out of view. Rotation periods for 310 of these stars are published in García et al. (2014) who used a combination of an ACF and wavelet transform method to measure surface rotation. From the 505 targets in the original sample, 310 rotation periods were reliably measured, 14 of which have precise asteroseismic ages from AMP modelling. All stars in the asteroseismic sample with rotation periods published by McQuillan et al. (2014), also appear in the rotation period catalogue of García et al. (2014). There is excellent agreement between rotation period measurements where the two catalogues overlap. Of the 114 stars which appear in both catalogues, only the rotation periods of 4 were not consistent at the  $1\sigma$  level and of these only 1, KIC 4931390, was inconsistent at greater than  $2\sigma$ . The similarities between the two catalogues is further described in García et al. (2014).

The asteroseismic sample covers a large range of ages, but it does not provide good mass coverage across the entire range (see figures 2-1 and 2-2). Few stars have

temperatures below 6000 K ( $B - V \sim 0.55$ ) and of the low mass stars, most of them are old (note that massive stars evolve rapidly and so we do not expect many in the sample). The exclusion of young, low-mass stars from the asteroseismic sample is due to the fact that these stars are more active and their power-spectra do not display high signal-to-noise acoustic oscillations. The omission of these as well as other types of stars that are not ideal asteroseismic targets from our sample should not bias our results. The mere lack of data in some regions of parameter space will not skew the best fitting model, but it is important to note that the resulting gyrochronology relation will not necessarily be descriptive of those stars not represented in this sample. We filled in some of the missing parameter space by adding 50 stars to our sample from young clusters Coma Berenices (0.5 Gyr), and the Hyades (0.625 Gyr) (see table 2.2). Clusters younger than 0.5 Gyr often have large populations of rapid rotators that have not yet converged onto the gyrochronology plane, so no clusters younger than Coma Ber were included. Uncertainties on  $B - V$  colours associated with each cluster star were not provided in the catalogues from which rotation periods and ages were obtained. Since the uncertainty associated with each measurement plays such a key role in our analysis (see §2.2), we assigned an uncertainty of  $\pm 0.01$  mag to each colour measurement, based on a realistic estimate of the typical uncertainties expected. The 1.1 and 0.588 Gyr open clusters, NGC 6811 (Meibom et al., 2011a) and Praesepe, (Delorme et al., 2011a,b) were originally included in our analysis. However we discovered that their period-colour relations were different to those of the Hyades and Coma Ber, as well as to each other's, and we therefore did not include them in our final analysis. A further 6 field stars with precise age measurements were added to the sample: 16 Cyg A and B, Alpha Cen A and B, 18 Sco and, of course, the Sun (see table 2.3). The entire set of 365 stars is shown in figures 2-1 and 2-2. Asteroseismic targets are shown in grey, with cluster and field stars in blue and the Sun in red.

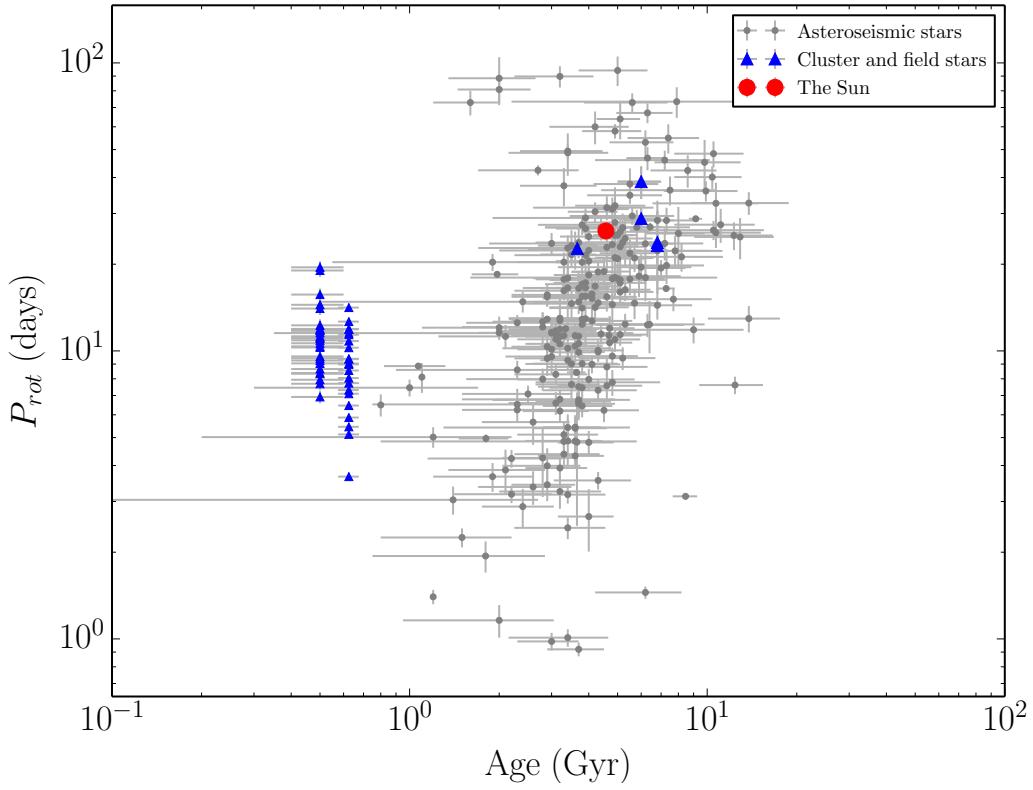


Figure 2-1: Photometric rotation period vs age for 310 *Kepler* targets (grey circles) plus cluster and field stars (blue triangles). The Sun is shown as a red circle.

Table 2.2: Clusters and References: (1) Dobbie et al. (2009), (2) Collier Cameron et al. (2009), (3) Perryman et al. (1998), (4) Radick et al. (1987).

Cluster	Age (Gyr)	Number of stars	Age ref	Rotation period ref
Coma Ber	$0.5 \pm 0.1$	28	1	2
Hyades	$0.625 \pm 0.05$	22	3	4

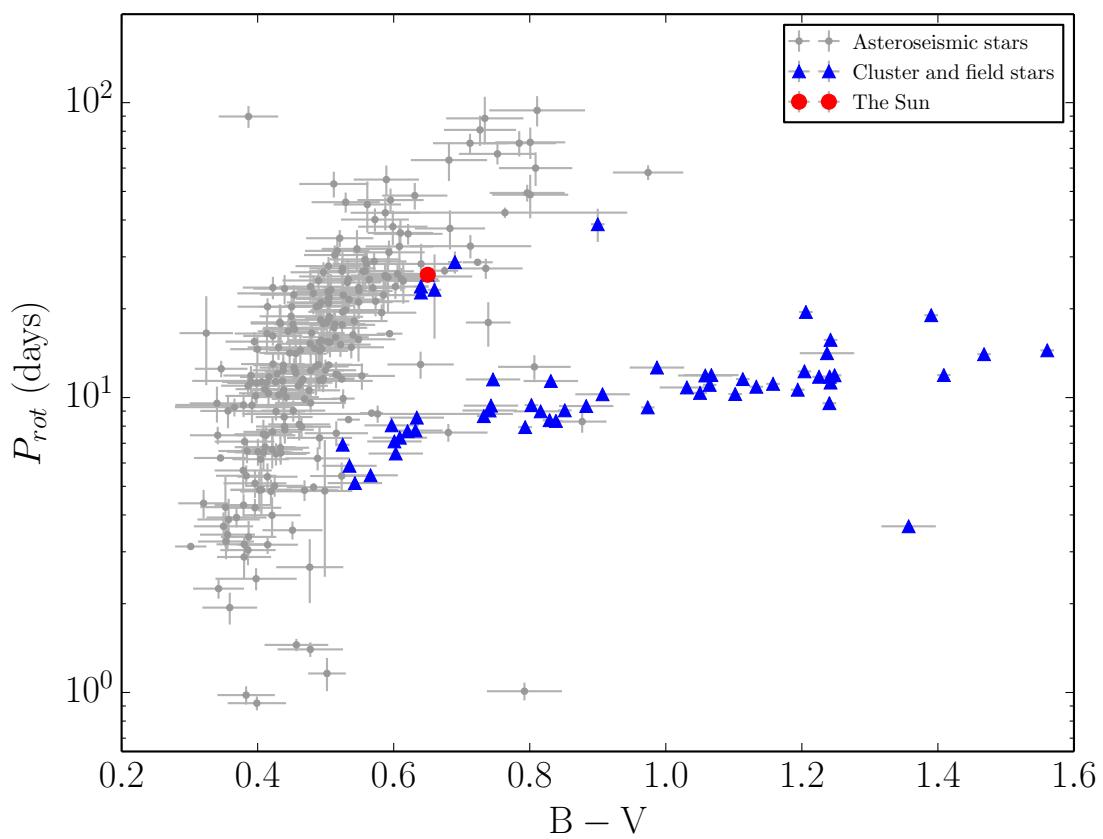


Figure 2-2: Photometric rotation period vs  $B - V$  colour for the data described in figure 2-1.

Table 2.3: Rotation periods and  $B - V$  colours for field stars with precise ages. Notes: Davies et al. (2014) measured internal rotation periods for 16 Cyg A and B using asteroseismology. However, this is likely to be close to the surface rotation value. Rotation periods for and  $\alpha$  Cen A and B were measured from variation in chromospheric emission lines. High-resolution spectropolarimetric observations were used to measure a rotation period for 18 Sco. The age of 16 Cyg AB was measured with asteroseismology. 18 Sco's age measurement was based chiefly on an asteroseismic analysis, however its rotation period was used as an additional constraint, so the age estimate is not entirely independent of rotation period for this star. An age for the  $\alpha$  Cen system was estimated from spectroscopic observations with additional seismic constraints. References: a = Metcalfe et al. (2012), b = Davies et al. (2014), c = Moffett and Barnes (1979), d = Li et al. (2012), e = Petit et al. (2008), f = Mermilliod (1986), g = Bouvier and Wadhwa (2010), h = Donahue et al. (1996), i = Cox and Pilachowski (2000), j = Bazot et al. (2012), k = Yıldız (2007), l = Hallam et al. (1991), m = Dumusque et al. (2012)..

ID	age	$P_{rot}$	$B - V$	References
16 Cyg A	$6.8 \pm 0.4$	$23.8^{+1.5}_{-1.8}$	$0.66 \pm 0.01$	a, b, c
16 Cyg B	$6.8 \pm 0.4$	$23.2^{+11.5}_{-3.2}$	$0.66 \pm 0.01$	a, b, c
18 Sco	$3.7 \pm 0.2$	$22.7 \pm 0.5$	$0.64 \pm 0.01$	d, e, f
The Sun	$4.568 \pm 0.001$	$26.09 \pm 0.1$	$0.65 \pm 0.001$	g, h, i
$\alpha$ Cen A	$6 \pm 1$	$28.8 \pm 2.5$	$0.69 \pm 0.01$	j, k, l, f
$\alpha$ Cen B	$6 \pm 1$	$38.7 \pm 5.0$	$0.90 \pm 0.01$	j, k, m, f

## 2.2 Calibrating the Gyrochronology relation

### 2.2.1 The model

The 310 asteroseismic stars in our sample have  $B - V$  colours converted from effective temperatures, photometric rotation periods, asteroseismic ages, and asteroseismic surface gravities. Each measurement of these properties is assumed to be independent with an associated Gaussian uncertainty. Not all of the cluster stars added to our sample have  $\log g$  values; however, since we only use  $\log g$  to separate the populations of subgiants and dwarfs (and we assume that the cluster stars are dwarfs) this should not affect our results. Following the treatment of Barnes (2007) and Mamajek and Hillenbrand (2008), rotation period was treated as the independent variable throughout the modelling process.

Hot stars and subgiants follow a different gyrochronology relation to MS dwarfs. Stars with effective temperatures above the Kraft-break,  $T_{\text{eff}} \sim 6250$  K, (Kraft, 1967) do not have a thick convective envelope and cannot support a strong magnetic dynamo, so do not spin down appreciably during their MS lifetimes. Subgiants spin down rapidly as they expand due to angular momentum conservation and thus diverge from the gyrochronological mass-period-age plane. The point in their evolution at which they depart, the ‘gyrochronological MS turn off’, is difficult to define. Classically, MS turnoff is defined as the hottest point on a star’s path on the HR diagram (before it ascends the giant branch) but theory predicts that evolving stars begin the process of spinning down relatively slowly after leaving the classically defined MS (van Saders and Pinsonneault, 2013). For this reason we chose a very simple definition of MS turnoff: we defined a  $\log g$  boundary of 4.2 to mark the transition between dwarfs and giants. We tried a range of boundary values and found that 4.2 minimised subgiant contamination whilst maximising the cool dwarf sample. It was also necessary to use a mixture model to account for misclassified subgiants—without it, subgiant contamination significantly biased the resulting fit. We did not exclude hot stars and subgiants from our sample during the modelling process, we modelled all three populations simultaneously. This allowed for the fact that stars have some probability

mass lying in all three regimes due to their large observational uncertainties.

Hot MS stars were defined as those with  $B - V < 0.45$ , corresponding to  $T_{\text{eff}} \approx 6250$  K for solar metallicity and  $\log g$ . Since there is no dependence of rotation period on age for massive MS stars, their rotation periods were modelled as a normal distribution with mean and standard deviation,  $Y$  and  $V$ , as free parameters. Subgiant rotation periods *do* depend on age and  $T_{\text{eff}}$ . However, since the rotational properties of these stars are not interesting for the purposes of gyrochronology calibration, we also modelled them with a normal distribution with mean and standard deviation,  $Z$  and  $U$ , as free parameters. We used a mixture model for the remaining population of stars, consisting of cool dwarfs and misclassified, contaminating subgiants. The subgiants were treated as if their rotation periods had been drawn from a background normal distribution with mean and standard deviation,  $X$  and  $U$ , again inferred from the data and another parameter,  $Q$ , the probability of each star belonging to that background population. The results of this analysis were not particularly sensitive to the choice of distribution for the hot stars and subgiants. These models were put in place so that we did not have to throw any data away and we could model everything at once. This was important because stars were classified according to their observed temperatures and  $\log g$ s, which are noisy. Due to the large uncertainties on  $T_{\text{eff}}$  and  $\log g$ , each star therefore has some probability of being a subgiant, some of being a cool dwarf and some of being a hot star. By throwing away data, one could accidentally throw away a misclassified star. We avoid this problem by modelling all stars simultaneously and taking a probabilistic approach to classification. Inferences made about the parameters of the normal distributions used to model subgiants and hot stars were not of particular interest for the purposes of gyrochronology calibration. We decided to use simple normal distributions rather than more physically motivated models in order to remain as model-independent as possible. Figure 2-3 shows  $\log g$  vs  $T_{\text{eff}}$  for the asteroseismic stars. Stars classified as cool dwarfs are shown in black, hot dwarfs in red and subgiants in blue.

Ideally both the hot star ( $B - V < 0.45$ ) and subgiant ( $\log g < 4.2$ ) boundaries would be free parameters in our model. However, since these two populations were

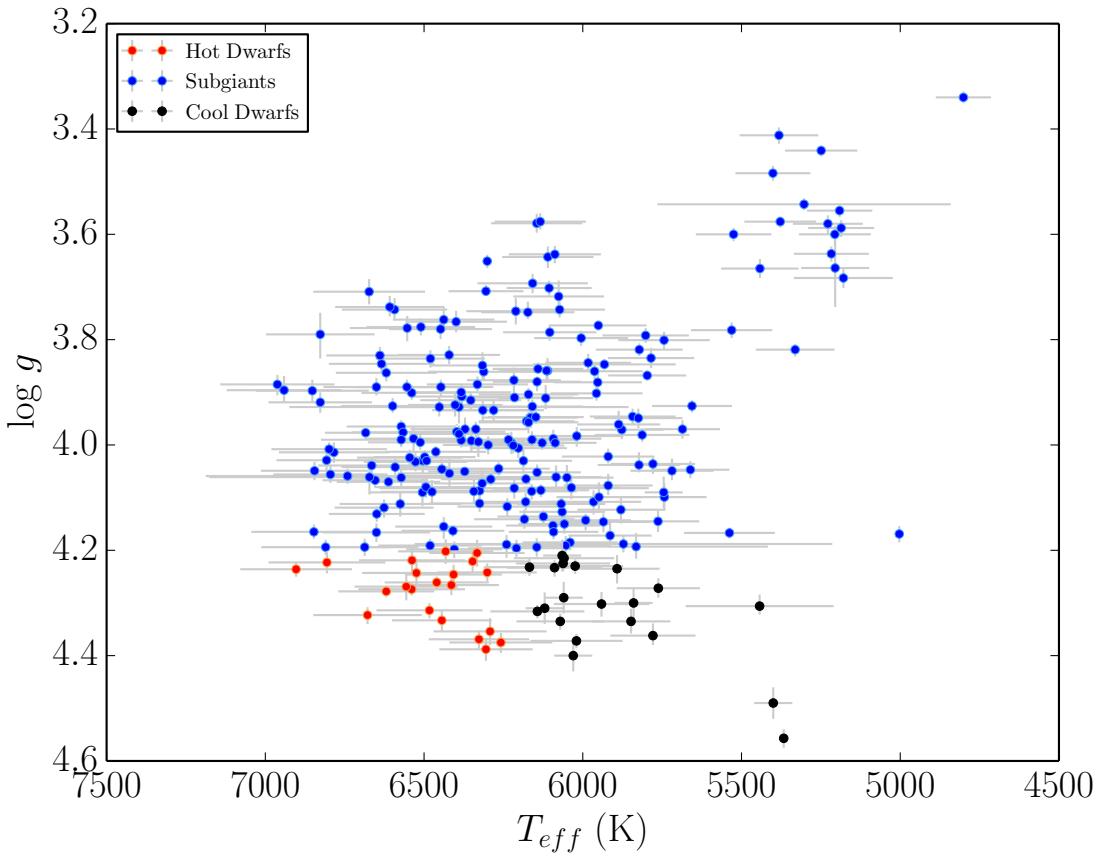


Figure 2-3:  $\log g$  vs  $T_{\text{eff}}$  for the 310 asteroseismic stars. Hot dwarfs with  $T_{\text{eff}} > 6250$  K and  $\log g > 4.2$  are red, subgiants with  $\log g < 4.2$  are blue. Only the black cool dwarfs with  $T_{\text{eff}} < 6250$  K and  $\log g < 4.2$  are expected to follow the gyrochronology relation in equation 2.1.

modelled with a relatively unconstraining normal distributions, these boundary parameters would not be well behaved. Both would be pushed to higher and higher values until all stars were modelled with a normal distribution. In order to avoid this problem, we fixed these two boundaries. A future analysis could avoid the assumption that the gyrochronology relation is infinitely narrow and assign it some intrinsic width, which would also be a free parameter.

We postulate that there is a deterministic relationship between the ‘true’ rotation period of a cool MS star and its ‘true’ age and colour, described by equation 2.1 (by ‘true’ we mean the value an observable property would take, given extremely high signal-to-noise measurements). Rotation period also depends on  $\log g$  since this property determines whether a star falls in the dwarf or subgiant regime.

In what follows we use  $P$  to denote rotation period and define  $\mathbf{w} = (\text{age}, B - V, \log g)$  as the vector of additional observational properties. Observations are denoted as  $\hat{P}_n$  and  $\hat{\mathbf{w}}_n$  and the unobserved (latent), ‘true’ parameters as  $P_n$  and  $\mathbf{w}_n$  for stars  $1, \dots, N$ .

In order to explore the posterior Probability Density Functions (PDFs) of the model parameters,  $\theta$ , conditioned on a set of noisy observations,  $\{\hat{P}_n, \hat{\mathbf{w}}_n\}$ , it was necessary to marginalise over the latent parameters,  $\{P_n, \mathbf{w}_n\}$ . This is because the model parameters,  $\theta$  are conditionally dependent on the ‘true’ values of rotation period and colour, *not* the observed values themselves. The observed values are only conditionally dependent on the ‘true’ values. In order therefore to infer the values of the model parameters using the observed values of rotation period and colour, it was necessary to marginalise over the latent parameters. This is another example of hierarchical inference. Assuming all measurements are independent, the marginalised likelihood can be written

$$p(\{\hat{P}_n, \hat{\mathbf{w}}_n\}|\theta) = \prod_{n=1}^N \int p(\hat{P}_n, \hat{\mathbf{w}}_n, P_n, \mathbf{w}_n|\theta) dP_n d\mathbf{w}_n. \quad (2.2)$$

The joint probability, on the right hand side of this equation, can be factorised as

$$p(\hat{P}_n, \hat{\mathbf{w}}_n, P_n, \mathbf{w}_n | \theta) = \quad (2.3)$$

$$p(P_n | \mathbf{w}_n, \theta)p(\hat{P}_n | P_n)p(\hat{\mathbf{w}}_n | \mathbf{w}_n)p(\mathbf{w}_n),$$

where we have utilised the fact that the observations,  $\hat{P}_n$  and  $\hat{\mathbf{w}}_n$  are *conditionally independent* of the model parameters,  $\theta$ : they only depend on  $\theta$  through the latent parameters,  $P_n$  and  $\mathbf{w}_n$ . The above integral can be written

$$p(\{\hat{P}_n, \hat{\mathbf{w}}_n\} | \theta) \propto \prod_{n=1}^N \int p(\mathbf{w}_n | \hat{\mathbf{w}}_n) d\mathbf{w}_n \quad (2.4)$$

$$\int p(P_n | \mathbf{w}_n, \theta)p(\hat{P}_n | P_n)dP_n,$$

where we have used Bayes' theorem:  $p(\mathbf{w}_n | \hat{\mathbf{w}}_n) \propto p(\hat{\mathbf{w}}_n | \mathbf{w}_n)p(\mathbf{w}_n)$ . The outer integral is the same for hot dwarfs, cool dwarfs and subgiants alike. In our model, the likelihood function—the probability of the ‘true’ rotation period given the ‘true’ observed parameters and the model parameters,  $p(P_n | \mathbf{w}_n, \theta)$ , is different in each regime because a different generative process is responsible for producing rotation periods. For cool dwarfs ( $B - V > 0.45$  and  $\log g > 4.2$ ) the likelihood function can be written

$$p(P_n | \mathbf{w}_n, \theta) = (1 - Q) \delta(P_n - f_\theta(\mathbf{w}_n)) \quad (2.5)$$

$$+ Q \left( \sqrt{2\pi U^2} \right)^{-1/2} \exp \left( -\frac{(P_n - X)^2}{2U^2} \right),$$

where  $Q$  is the probability that a star is drawn from the population of misclassified subgiants and

$$f_\theta(\mathbf{w}_n) = A^n \times a(B - V - c)^b \quad (2.6)$$

is the gyrochronology relation. For hot dwarfs ( $B - V < 0.45$  and  $\log g > 4.2$ ) the

likelihood function is:

$$p(P_n | \mathbf{w}_n, \theta) = \left( \sqrt{2\pi V^2} \right)^{-1/2} \exp \left( -\frac{(P_n - Y)^2}{2V^2} \right), \quad (2.7)$$

and for subgiants,

$$p(P_n | \mathbf{w}_n, \theta) = \left( \sqrt{2\pi W^2} \right)^{-1/2} \exp \left( -\frac{(P_n - Z)^2}{2W^2} \right). \quad (2.8)$$

We used hierarchical inference to account for observational uncertainties, following the method of Hogg et al. (2010b), also used by Foreman-Mackey et al. (2014a), Rogers (2014), Morton and Winn (2014) and Demory (2014). We computed the likelihood for each star (equation 2.4), up to an unimportant constant using a sampling approximation. The values of  $\hat{P}_n$  and  $\hat{\mathbf{w}}_n$  with uncertainties,  $\sigma_P$  and  $\sigma_{\mathbf{w}}$ , reported in catalogues provide constraints on the posterior probability of those variables, under a choice of prior PDF,  $p_0(\hat{\mathbf{w}}_n)$ . Ideally, these catalogues would provide posterior PDF samples, not just point estimates, which we could use directly. i.e. samples from

$$p(\mathbf{w}_n | \hat{\mathbf{D}}_n) = \frac{p(\hat{\mathbf{D}}_n | \mathbf{w}_n) p_0(\mathbf{w}_n)}{p_0(\hat{\mathbf{D}}_n)}, \quad (2.9)$$

where  $p(\hat{\mathbf{D}}_n | \mathbf{w}_n)$  is the likelihood of the data,  $\hat{\mathbf{D}}_n$  (in this case, the set of *Kepler* lightcurves plus spectroscopic  $T_{\text{eff}}$  and [Fe/H] measurements), given the model parameters,  $\mathbf{w}_n$ .  $p_0(\mathbf{w}_n)$  is an uninformative prior PDF, chosen by the fitter (Chaplin et al., 2014, used a flat prior PDF in age and  $\log g$ ). In the absence of posterior PDF samples<sup>2</sup> we generated our own from Gaussian distributions with means,  $\hat{\mathbf{w}}_n$  and standard deviations,  $\sigma_{\mathbf{w}}$ .  $J$  posterior samples were generated for each star (we used  $J = 500$ ):

$$\mathbf{w}_n^{(j)} \sim p(\mathbf{w}_n | \hat{\mathbf{w}}_n), \quad (2.10)$$

and were used to evaluate  $p(\mathbf{w}_n | \hat{\mathbf{w}}_n)$  up to a normalisation constant. Using these

---

<sup>2</sup>Posterior PDF samples for asteroseismic parameters are now beginning to be published and will be made available in future publications.

samples we computed the marginalised likelihood for a single star as follows,

$$p(\hat{P}_n, \hat{\mathbf{w}}_n | \theta) \approx \frac{1}{J_n} \sum_{j=1}^{J_n} p(\hat{P}_n | \mathbf{w}_n^{(j)}, \theta) . \quad (2.11)$$

The argument inside this sum is given by the integral

$$p(\hat{P}_n | \mathbf{w}_n^{(j)}, \theta) = \int p(P_n | \mathbf{w}_n, \theta) p(\hat{P}_n | P_n) dP_n . \quad (2.12)$$

Assuming that the period uncertainties are Gaussian with mean  $\hat{P}_n$  and variance  $\sigma_n^2$ , this integral can be evaluated analytically for each population. For example, starting from equation 2.5 the result for the cool dwarfs is

$$\begin{aligned} p(\hat{P}_n | \mathbf{w}_n^{(j)}, \theta) &= \frac{1 - Q}{\sqrt{2\pi\sigma_n^2}} \exp\left(-\frac{[\hat{P}_n - f_\theta(\mathbf{w}_n^{(j)})]^2}{2\sigma_n^2}\right) \\ &\quad + \frac{Q}{\sqrt{2\pi(U^2 + \sigma_n^2)}} \exp\left(-\frac{[\hat{P}_n - X]^2}{2[U^2 + \sigma_n^2]}\right) . \end{aligned} \quad (2.13)$$

A similar result can be derived for the other populations.

Finally, using these analytic results for the inner integral, the marginalised log-likelihood from equation 2.4 becomes

$$\begin{aligned} \log p(\{\hat{P}_n, \hat{\mathbf{w}}_n\} | \theta) &\approx \quad (2.14) \\ &\log \mathcal{Z} + \sum_{n=1}^N \log \left[ \sum_{j=1}^{J_n} p(\hat{P}_n | \mathbf{w}_n^{(j)}, \theta) \right] \end{aligned}$$

where  $\mathcal{Z}$  is an irrelevant normalisation constant. We used `emcee` (Foreman-Mackey et al., 2013), an affine invariant, ensemble sampler Markov Chain Monte Carlo (MCMC) algorithm, to explore the posterior PDFs of the model parameters,  $\theta$ . Flat prior PDFs were used for each parameter. Following the above method, a likelihood was computed as follows:

- For each star,  $J$  samples were drawn from three normal distributions: one in colour, one in age and one in  $\log g$ , where the means and standard deviations of

those distributions were the observed values and uncertainties. This step was performed just once and the following steps were performed for each likelihood evaluation.

- For those samples that fell in the cool dwarf regime ( $B - V > 0.45$  and  $\log g > 4.2$ ), model rotation periods were both calculated using equation 2.1 and assigned the value of parameter  $X$ . Likelihoods for the two model rotation periods were then evaluated using a Gaussian mixture model likelihood function (equation 2.13).
- For the samples that fell in the hot dwarf ( $B - V < 0.45$  and  $\log g > 4.2$ ) and subgiant ( $\log g < 4.2$ ) regimes, likelihoods were calculated by comparing observed rotation periods with the model rotation periods for the two populations:  $Y$  and  $Z$ .
- The total log-likelihood for each star was calculated as the sum of the log-likelihoods of each of the  $J$  samples.
- Finally, the sum of individual star log-likelihoods provided the total log-likelihood.

## 2.3 Results and Discussion

A gyrochronology relation was initially fit to the asteroseismic stars, the field stars and four clusters (Hyades, Coma Berenices, Praesepe and NGC 6811) all together. However, this resulted in extremely multi-modal posteriors PDFs for  $a$ ,  $b$  and  $n$ . After fitting a separate relation to various subsets of the data, it became evident that a

Table 2.4: Median values of  $a$ ,  $b$  and  $n$  for individual clusters (see equation 2.1).

Parameter	Coma Berenices	Hyades
a	$0.417^{+0.08}_{-0.07}$	$0.312^{+0.04}_{-0.06}$
b	$0.271^{+0.05}_{-0.06}$	$0.410^{+0.05}_{-0.04}$
n	$0.542 \pm 0.03$	$0.599^{+0.03}_{-0.02}$

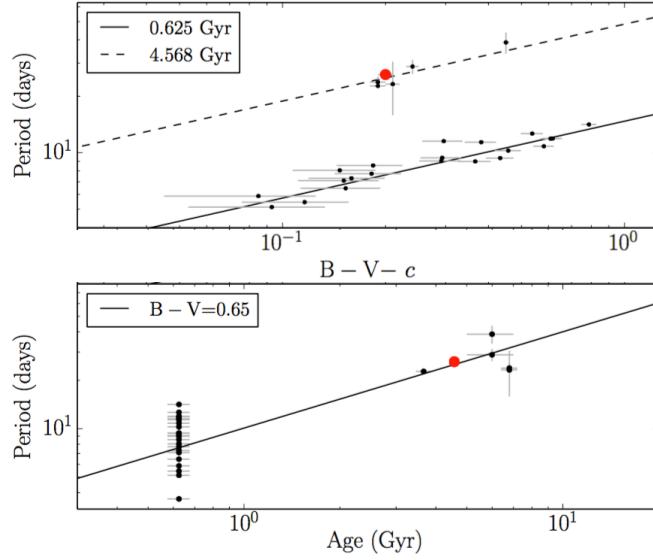
lower value of  $b$  was preferred by NGC 6811, i.e. the slope of the  $\log(\text{period}) - \log(B - V - c)$  relation was shallower for NGC 6811 than for the Hyades, Coma Berenices and Praesepe. The reason for NGC 6811's shallower slope is currently unknown, however the slope increased slightly after more quarters of *Kepler* data were included in the rotation period measurement process (Meibom, private communication), so this may be due to systematic uncertainties on the rotation periods. We therefore excluded NGC 6811 from our sample and attempted to fit a relation to the remaining data. Multi-modal posterior PDFs were still produced however, until Praesepe was also removed from the sample. The reason for this multi-modality is unclear, but we tentatively attribute it to Praesepe preferring a different value for the colour discontinuity,  $c$ , to the Hyades and Coma Berenices. We calculated the likelihood for Praesepe, plus the field stars (to provide the age dependence) with two different values of  $c$ : 0.45 and 0.5, finding a higher likelihood for  $c = 0.5$ . Since we do not fully understand the cause of this variation in  $c$ , and in order to keep our model simple, we chose to also exclude Praesepe from our final data set and fit a gyrochronology relation with  $c = 0.45$  to the remaining data (asteroseismic stars, field stars, Hyades and Coma Berenices). We also fitted relations with  $c$  values ranging from 0.4 to 0.55 to this final data set, finding that the results were relatively insensitive to variations in this parameter (solar ages predicted from each best-fitting model were consistent within uncertainties). Individual fits to Hyades and Coma Berenices, plus the field stars are shown in figures 2-4(a) and 2-4(b). Median values of  $a$ ,  $b$  and  $n$  for the two clusters with their 16th and 84th percentile uncertainties are provided in table 2.4. Note that none of these parameters are fully consistent between the two clusters. The fact that each cluster seems to prefer a different value of  $a$ ,  $b$ ,  $n$  and  $c$  paints a concerning picture for this form of a gyrochronology relation which assumes one set of parameters can be used to describe all stars. It is likely that the inability of equation 2.1 to fully describe the observed properties of our sample of stars is due to the simplifying assumptions that go into this relationship. The fact that there is no dependence on metallicity, for example, may contribute to the deficiencies of this model. One might expect metallicity to have an effect on the rotational evolution of a star since it impacts

Table 2.5: Median values of the parameters describing the populations of non-gyrochronological stars.

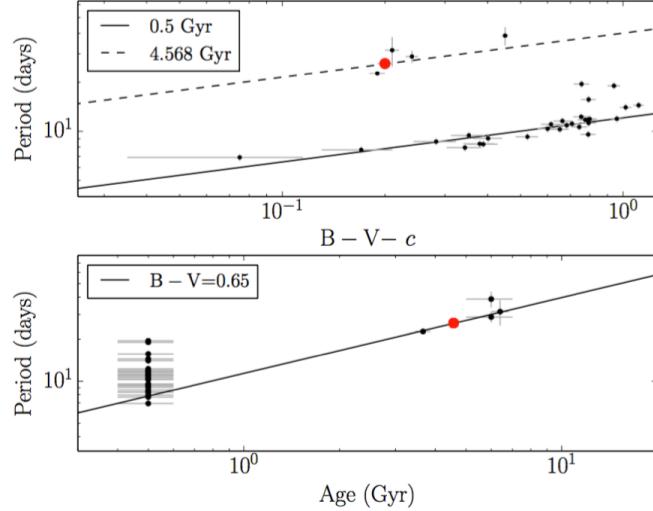
Parameter	Symbol	Median value
$\sigma_P(B - V > 0.45)$	$U$	$8_{-2}^{+4}$ days
$\sigma_P(B - V < 0.45)$	$V$	$2.1_{-0.6}^{+1}$ days
$\sigma_P(\log g < 4.2)$	$W$	$9.9_{-0.5}^{+0.7}$ days
$\mu_P(B - V > 0.45)$	$X$	$14 \pm 4$ days
$\mu_P(B - V < 0.45)$	$Y$	$5.0_{-0.8}^{+1}$ days
$\mu_P(\log g < 4.2)$	$Z$	$16.1_{-0.8}^{+0.7}$ days
Giant fraction	$Q$	$0.14_{-0.05}^{+0.06}$

its internal structure. We have not attempted to calibrate a metallicity-dependent gyrochronology relation, because there are no precise metallicity measurements for the majority of the *Kepler* stars, on which this study is chiefly based. In addition, we anticipate that the majority of stars for which this new gyrochronology relation will be most useful are *Kepler* stars, or stars targeted by missions like *Kepler*, again, the majority of which are unlikely to have precise metallicities.

The resulting highest probability values of  $a$ ,  $b$  and  $n$  from our fit to the final data set, with 16th and 84th percentile uncertainties are presented in table 2.1. The additional parameters of our model, describing the distributions of hot star and subgiant rotation periods, are presented in table 2.5. The posterior PDFs of these parameters were all unimodal. Note the value of  $Q$ , the parameter describing the fraction of misclassified subgiants is  $0.14_{-0.05}^{+0.06}$ , i.e., based on our simple ‘ $\log g = 4.2$ ’ definition of MS turn-off, which left only 21 stars classified as cool dwarfs, one or two of these are likely to be misclassified subgiants. Marginalised posterior PDFs for the three gyrochronology parameters are shown in figure 2-5 and the resulting relation between period and age for stars of Solar-like colour is shown in figure 2-6. The relation between period and colour for Solar-age stars is shown in figure 2-7 and for a range of ages in figures 2-8(a)-2-8(e). Note that we plot rotation period vs ‘ $B - V - c$ ’, producing a straight line, in order to give a more intuitive understanding of the quality of the fit to the data. In figures 2-6 to 2-8(e) we plotted 100 draws from the posterior PDFs of the gyrochronology parameters as faint grey lines, in order to demonstrate the widths



(a) Rotation period vs age and  $B - V - c$  for the Hyades.



(b) Rotation period vs age and  $B - V - c$  for Coma Berenices.

Figure 2-4: Individual fits to the clusters and field stars. The Sun is the red point. The top figure shows rotation period vs ' $B - V - c$ ' for the Hyades (top) and Coma Berenices (bottom) with Solar and cluster age isochrones. The bottom figure shows rotation period vs age for the Hyades (top) and Coma Berenices (bottom) with the period- age relation for a constant  $B - V$  value of 0.65 (Solar  $B - V$ ).

and bimodal natures of these distributions.

Figure 2-5 shows that parameters  $a$  and  $n$  are highly correlated and their posterior PDFs are bimodal. The position of the second peak falls around  $a = 0.8$ ,  $b = 0.34$  and  $n = 0.44$ . The cause of this bimodality is clear when looking at the faint grey lines representing draws from the parameter posterior PDFs in figure 2-6. The majority of these draws fall close to the best-fitting model, which passes neatly through the Sun and field stars, but a significant fraction fall below the line of best-fit, passing through the *Kepler* asteroseismic stars which also mostly fall below the line. This result is reflected in García et al. (2014) who model the AMP *Kepler* asteroseismic stars from Mathur et al. (2012) and Metcalfe et al. (2014), without anchoring their relation to the Sun. They find that the model that best describes the *Kepler* asteroseismic stars underpredicts the rotation period of the Sun. The grey lines that fall beneath our best-fitting model in figure 2-6 are drawn from the smaller peak in the posterior PDFs of  $a$  and  $n$  and seem to describe the relation between rotation period and age for the *Kepler* asteroseismic stars. In other words, the bimodal posterior PDFs of  $a$  and  $n$  are produced by the disagreement between the *Kepler* asteroseismic stars and the Sun and field stars. One set of gyrochronology parameters is not capable of describing the Sun, plus field stars, and the *Kepler* asteroseismic stars simultaneously. There is more than one possible explanation for this result. Firstly, the asteroseismic ages could be systematically biased high. Secondly, the rotation periods could be systematically underestimated. This could occur if, for example, these stars were rotating differentially and the dominant spotted regions on their surfaces were not equatorial, as is the case for the Sun. Thirdly, this could be a result of an observational bias produced by incomplete detection. If there were a large spread in rotation periods for a given stellar mass and age and, due to the detection bias brought about because shorter periods are easier to detect than longer periods, this broad range of rotation periods might be truncated at some upper cut-off. It would therefore appear as though Solar-colour stars were rotating too slowly for their age simply because only rapidly rotating *Kepler* stars appeared in our sample. *Kepler* systematics hinder our ability to measure longer rotation periods, but in addition, more slowly rotating stars

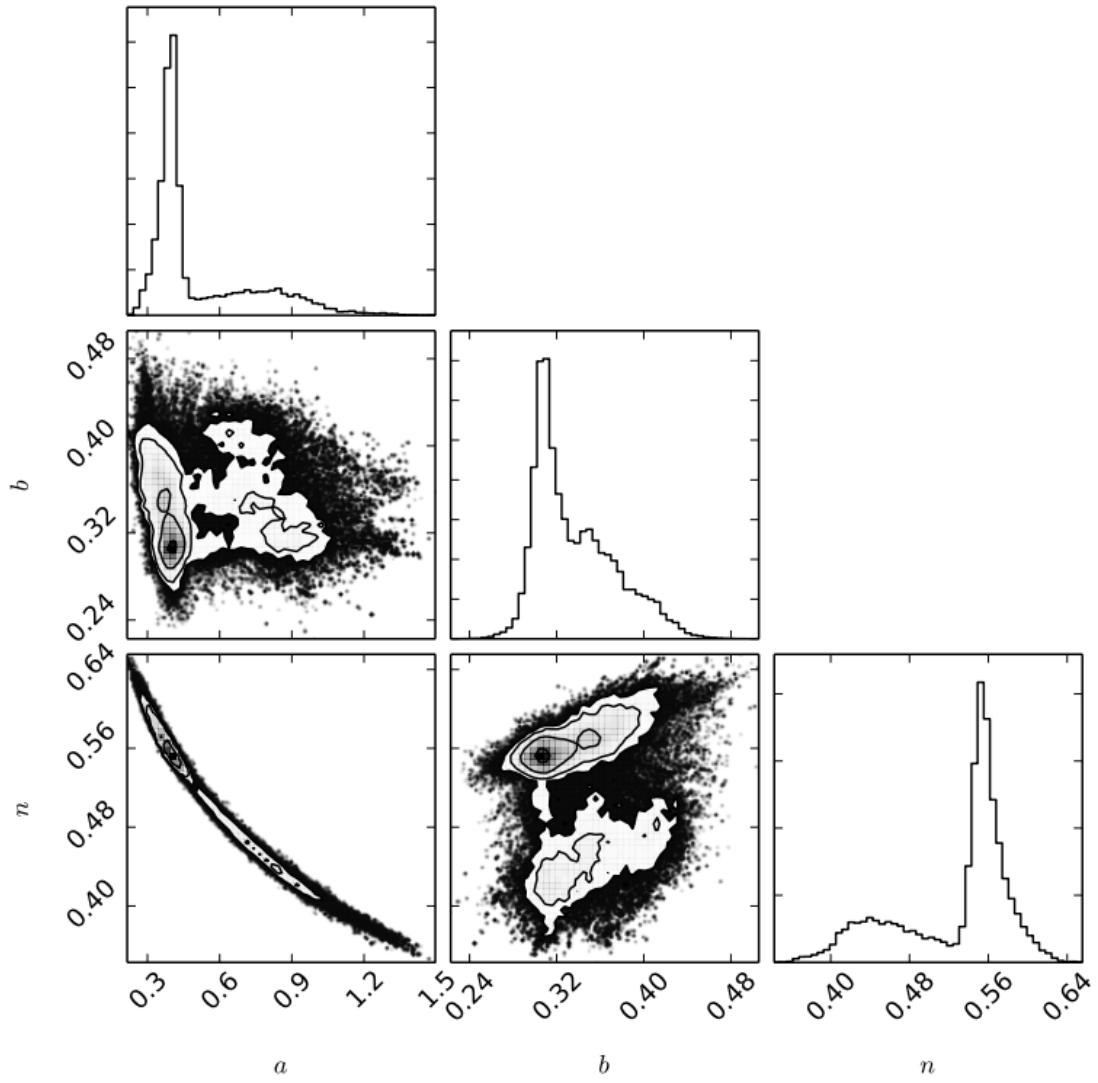


Figure 2-5: Marginalised posterior PDFs for the three gyrochronology parameters,  $a$ ,  $b$  and  $n$ . Parameters  $a$  and  $n$  are highly correlated and their posterior PDFs are bimodal. The main peaks in the posterior PDFs of  $a$  and  $n$  correspond to a fit to the Sun and field stars. The smaller peak corresponds to a fit to the *Kepler* asteroseismic stars. This plot was made using triangle.py (Foreman-Mackey et al., 2014).

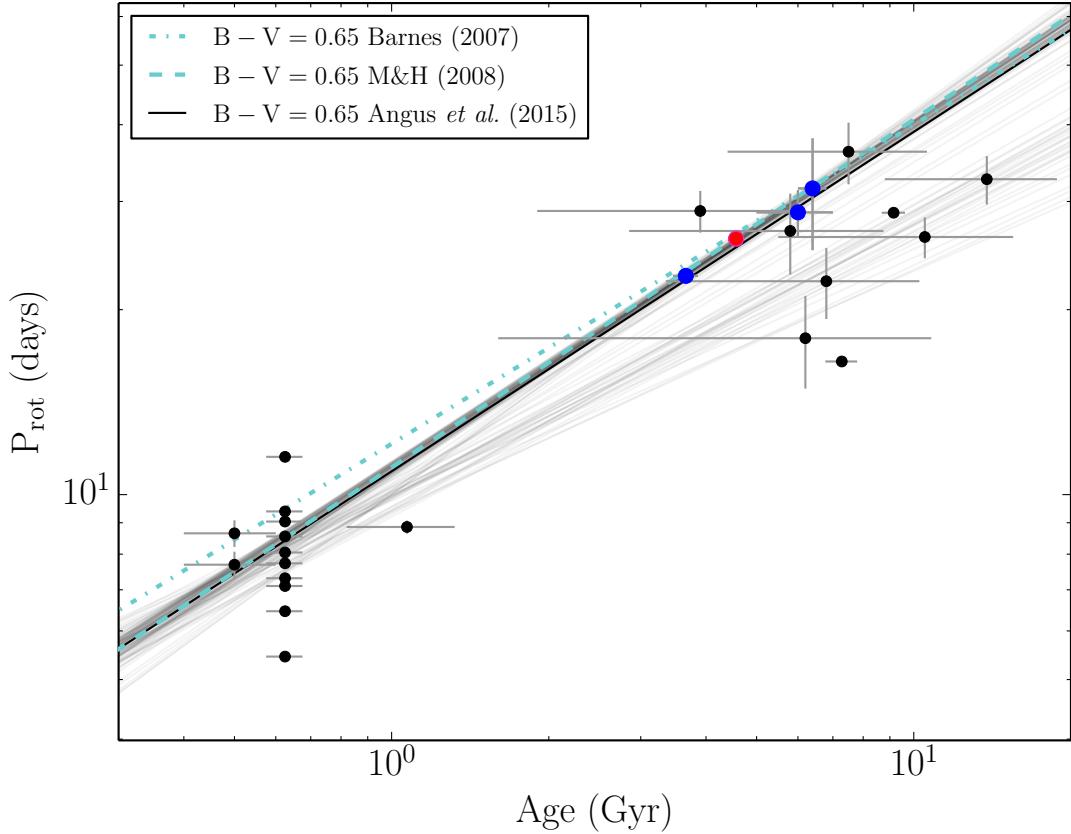


Figure 2-6: Rotation period vs age for cool dwarfs with colour within 0.1 of the Sun's: 0.65, with gyrochronology relations of Barnes (2007), Mamajek and Hillenbrand (2008) and this work. The Sun is shown in red and the field stars,  $\alpha$  Cen A, 18 Sco and 16 Cyg B from left to right, are shown in blue. The black points towards the lower left are cluster stars and those towards the upper right are *Kepler* asteroseismic stars. Each of the faint grey lines represents a sample drawn from the posterior probability distributions of  $a$ ,  $b$  and  $n$ . Whilst most of these draws come from the large peak in the posterior PDF and fall through the Sun and field stars, some describe the period-age relation of the *Kepler* asteroseismic stars. These lines are drawn from the smaller peak in the posterior PDF.

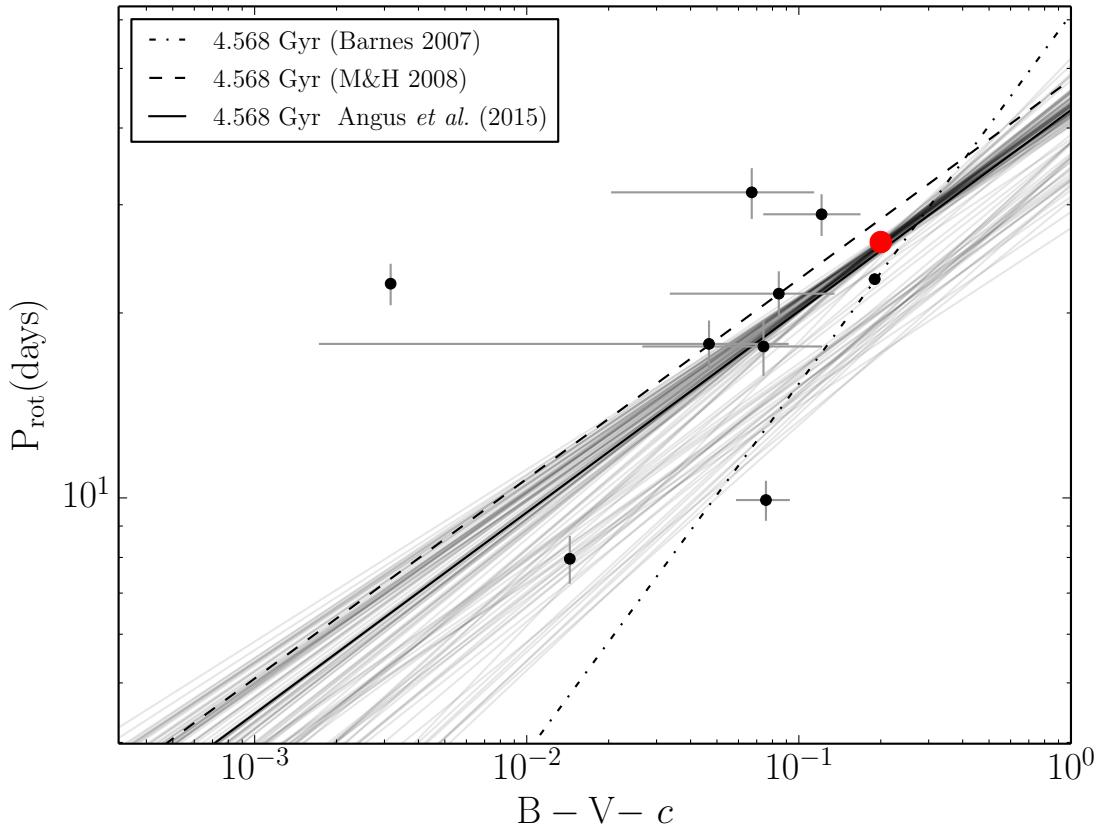


Figure 2-7: Rotation period vs ' $B - V - c$ ' for dwarfs with age within  $1\sigma$  of the Sun's age, 4.568 Gyr. The Sun is the red point. Each of the faint grey lines represents a sample drawn from the posterior probability distributions of  $a$ ,  $b$  and  $n$ . Many samples fall below the solid black line marking the highest probability parameter values due to the bimodal posterior.

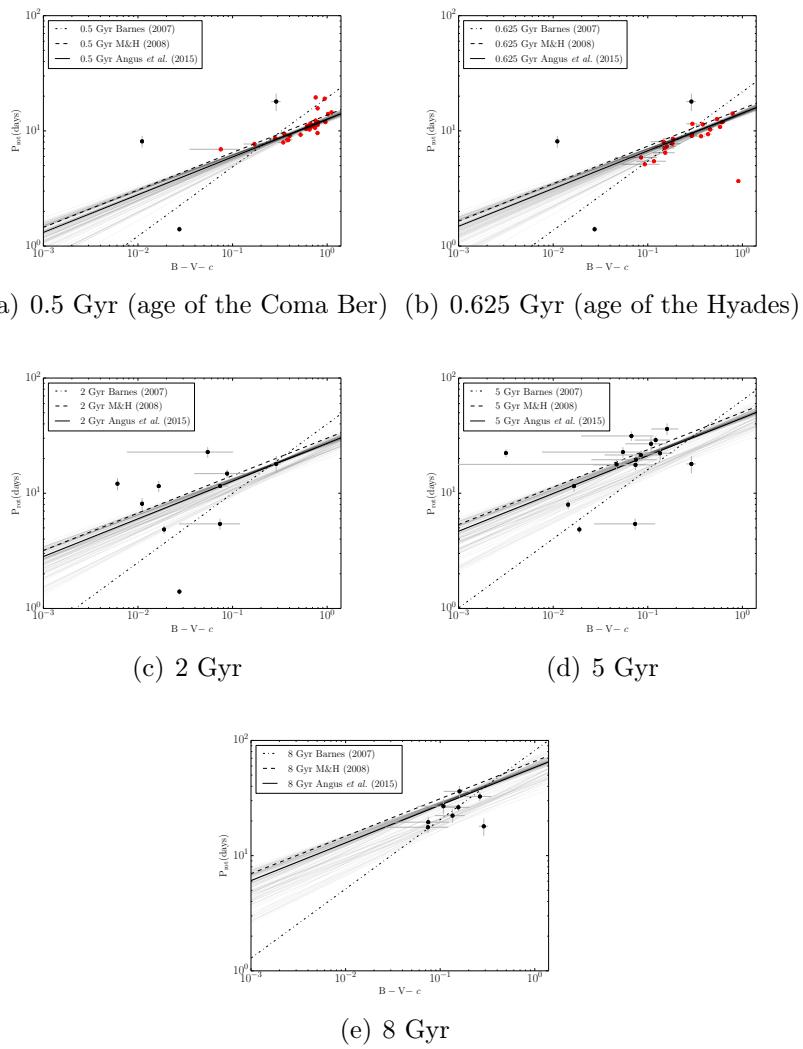


Figure 2-8:  $P_{rot}$  vs  $B - V - c$  for dwarfs within  $1\sigma$  of the reference age with the new gyrochronology relation and Barnes (2007), and Mamajek and Hillenbrand (2008) for comparison. Asteroseismic targets are black and cluster and field stars are red. Each of the faint grey lines represents a sample drawn from the posterior probability PDFs of  $a$ ,  $b$  and  $n$ . Many samples fall below the best fit model due to the bimodal posterior PDFs.

tend to be less active, with fewer surface features and are therefore more likely to be missing from our sample. Finally, it is possible that the *Kepler* asteroseismic stars follow a *different* spin-down relation to the Sun and field stars, perhaps due to having different metallicities. Unfortunately it is not currently possible to identify the cause of the observed mismatch between the *Kepler* stars and the field stars and we leave this question for a future investigation. The tight correlation between parameters  $a$  and  $n$  demonstrates that there is some redundancy in the fiducial gyrochronology relation. In the future, an alternative functional form with less correlation between parameters (for example a Gaussian process) should be explored.

The fact that our fit is so heavily dominated by the Sun with its small uncertainties is, perhaps, a cause for concern. We know the rotation period and age of the Sun very precisely, however, if the Sun is not *exactly* representative of a typical star, the resulting best-fit gyrochronology relation will also not represent typical stars. The field stars in our sample *are* well represented by the best-fit gyrochronology relation, which provides reassurance that the Sun is a typical star amongst this set. We did not attempt to tackle the problem of how to appropriately treat the Sun as a single, highly-precise data point in a sample that also contains imprecise data. Instead, we leave this problem for future consideration. Despite the fact that a lot of weight is attributed to the Sun because of its small uncertainties, this new gyrochronology relation is still the most representative, empirically calibrated relation between colour, rotation period and age for MS F, G and K stars to date.

Our final, newly calibrated gyrochronology relation can be written in full as

$$P = A^{0.55_{-0.09}^{+0.02}} \times 0.40_{-0.05}^{+0.3} (B - V - 0.45)^{0.31_{-0.02}^{+0.05}}, \quad (2.15)$$

with rotation period,  $P$  in days and age,  $A$  in Myr. An age can be calculated for a star with a rotation period and colour by inverting this relationship. Covariances between the gyrochronology parameters should be taken into account *whenever* the above relation is used to calculate uncertainties on an age or rotation period prediction. In order to do this properly, posterior PDF samples should be incorporated into Monte-

Table 2.6: Field star ages taken from the literature, compared with predictions from this work (1), Mamajek and Hillenbrand (2008) (2) and Barnes (2007) (3).

Star	Literature age (Gyr)	Age 1 (Gyr)	Age 2 (Gyr)	Age 3 (Gyr)
18 Sco	$3.7 \pm 0.2$	$3.7^{+2.5}_{-0.3}$	$3.5^{+0.6}_{-0.5}$	$3.7^{+0.8}_{-0.6}$
The Sun	$4.568 \pm 0.001$	$4.6^{+3.5}_{-0.3}$	$4.7^{+0.7}_{-0.6}$	$4.8^{+1}_{-0.8}$
Alpha Cen A	$6.0 \pm 1$	$5.0^{+3.3}_{-1.0}$	$4.5^{+1}_{-0.9}$	$5 \pm 1$
Alpha Cen B	$6.0 \pm 1$	$6.0^{+3.8}_{-1.7}$	$4^{+1}_{-0.9}$	$5^{+2}_{-1}$
16 Cyg A	$6.8 \pm 0.4$	$4.0^{+2.6}_{-0.7}$	$3 \pm 2$	$4 \pm 2$
16 Cyg B	$6.8 \pm 0.4$	$3.7^{+3.4}_{-2.1}$	$3 \pm 2$	$4 \pm 2$

Carlo uncertainty calculations<sup>3</sup>.

In order to test the predictive power of the new gyrochronology relation, we inverted equation 2.15 to compare previously measured ages with new age predictions for the 6 field stars (see table 2.6). For comparison, gyrochronological ages for the field stars were also computed using the relations of Barnes (2007) and Mamajek and Hillenbrand (2008). Uncertainties on ages predicted with the new relation were calculated using posterior PDF samples of the three parameters,  $a$ ,  $b$  and  $n$ .

The ages predicted by the three different relations are consistent within uncertainties, with the exception of Alpha Cen B. All three relations underpredict the age of the 16 Cyg system. The rotation periods of both 16 Cyg A and B used in this paper are asteroseismic measurements of the internal (not surface) rotation periods of the stars. If the stars' cores are rotating much more rapidly than their surfaces, this could account for this age discrepancy.

The goal of gyrochronology in general is to provide a means of predicting the age of a star given observations of its colour (or mass, or temperature), and rotation period. The discrepancies in period-colour relations between clusters, *Kepler* stars and nearby field stars in the above analysis does not bode well for the current, simple gyrochronology model. Although the individual gyrochronology ages predicted for most of the *Kepler* stars lie within the errorbars of the asteroseismic ages, there is an unmistakable systematic offset which cannot be described by the model. Until now it has been hoped that one single relation between period, mass and age could be

---

<sup>3</sup>Posterior samples for this project are available at <https://github.com/RuthAngus/Gyro>.

used to describe all F, G and K MS stars. Even though we are unable to identify the cause of the discrepancy between the *Kepler* stars and the field stars, a different gyrochronology relation still seems to be required for each cluster. This may simply be due to a metallicity discrepancy, but this still demonstrates that the simple gyrochronology relation used here is inadequate.

The ‘narrowness’ of the gyrochronology relation has hitherto been an unknown; do the three properties, age, mass and rotation period, truly lie on an infinitely narrow plane? Unfortunately we cannot fully answer this question here as the asteroseismic ages are noisy and observational and intrinsic scatter are ambiguously interwoven. However, if this sample of stars is subject to detection bias then, by definition, there must be a broad range of rotation periods for stars of a given mass and age. A future study might include an extra parameter that describes the ‘width’ of the gyrochronological plane and attempt to detect an element of scatter above the noise level. Does age depend solely on rotation period and mass or do other variables influence stellar spin down, perhaps only becoming important after many Myrs? The gyrochronology model calibrated here neglects the effects of metallicity. This property is bound to have an effect on the angular momentum evolution of a star since it impacts internal stellar structure. A future study, investigating the impact of metallicity on rotational evolution will be essential for improving our ability to date stars using gyrochronology.

The picture of gyrochronology will become clearer as the sample of asteroseismic stars with individual mode analysis grows and their age uncertainties shrink. The best targets for asteroseismic studies are relatively inactive since they allow the easier detection of Solar-like oscillations. Inactive stars are also the best targets for gyrochronology as they are usually old and slowly rotating. However these targets are not well suited for rotation period measurements which are most easily and precisely determined for active, rapidly rotating stars. K2, the repurposed *Kepler* mission, will provide new targets for rotation studies; in particular, some relatively old clusters have been and will continue to be monitored by the spacecraft. The observing seasons of K2 are relatively short ( $\sim 90$  days) and fields will only be observed once, so the maximum rotation periods measurable from K2 light curves will be considerably

shorter than with *Kepler*. However these clusters may still be extremely useful for gyrochronology.

## 2.4 Conclusions

We have calibrated the relation between rotation period,  $B - V$  colour and age for MS stars with  $T_{\text{eff}} < 6250$  K, using 310 *Kepler* asteroseismic targets, supplemented with 6 field stars and 50 cluster stars. Unlike previous gyrochronology calibrations, our sample covers a large range of ages, observational uncertainties on all parameters were accounted for and the posterior PDFs of model parameters were explored using MCMC. Incorporating observational uncertainties into the model fitting process was an essential component of our analysis since these uncertainties, particularly on the asteroseismic ages, were large. Three populations: hot dwarfs, cool dwarfs and subgiants, were modelled simultaneously in order to account for potential misclassifications that might have arisen from large observational uncertainties. Posterior probability distributions of the gyrochronology parameters were explored using MCMC and the impact of leaving out subsets of the sample was assessed, leading us to find that a single relation between rotation period, colour and age does not adequately describe the cluster data. For this reason, only the Hyades and Coma Berenices clusters were used to calibrate the model. Fitting equation 2.1 to the Hyades and Coma Berenices clusters, the *Kepler* asteroseismic stars, local field stars and the Sun resulted in bimodal posterior PDFs. The *Kepler* asteroseismic stars were not well described by the same gyrochronology relation as the Sun, field stars and clusters. There are several potential explanations for the cause of this phenomenon. Some of the most likely are: the asteroseismic ages or the photometric rotation periods could be systematically biased; the rotation periods might be subject to detection bias; and the different populations could have different physical properties (e.g. metallicity), not accounted for in the gyrochronology model. If detection bias is responsible and there truly *is* a broad range of rotation periods for stars of a given age and colour, this still has negative implications for Gyrochronology which relies on the assumption

that these three properties lie on a neat plane.

The careful and well motivated modelling techniques used in this work allowed us to identify the potential shortcomings of the current gyrochronology model. Only by conducting MCMC parameter exploration and studying the resulting posterior PDFs were we able to see that one global gyrochronology relation cannot describe all subsets of our sample. The results are somewhat unsettling—no single model describes the sample as a whole, even when dropping some of the clusters and allowing for rather complex population subsets. This tells us that the current model is probably not good enough, but it - or something very similar - is the model everyone has been using in gyrochronology studies so far. Previous studies did not model multiple cluster and field populations simultaneously, or do the modelling in such a detailed way, so the problems did not come through so clearly. In the future better physical models may be developed, better ages may be calculated and more sensitive period search methods, not so susceptible to detection bias may become available. When they do, the prescription we provide will enable the the period-colour-age relation to be modelled more sensitively.

Since the work in this chapter was published (Angus et al., 2015b), the result seen here: that *Kepler* asteroseismic stars rotate unexpectedly rapidly given their age and mass, was also found by van Saders et al. (2016). Instead of using the entire ensemble of short cadence asteroseismic targets from Chaplin et al. (2014), they used only ‘boutique’ targets. These are the highest S/N stars with individual oscillation mode frequencies detectable in their power spectra. These targets are also not well described by existing gyrochronology models. Whereas I showed that the Barnes (2007) empirical form did not fit the data, van Saders et al. (2016) compare rotation periods to predictions made using their theoretically motivated model, equation 1.2.3 in chapter 1. They cannot reproduce the observations using their model and present

the following alteration which *does* describe the data:

$$\frac{dJ}{dt} = \begin{cases} f_K K_M \omega \left( \frac{\omega_{crit}}{\omega_{\odot}} \right)^2, & \omega_{crit} \leq \omega \frac{\tau_c}{\tau_{c,\odot}}, Ro \leq Ro_{crit} \\ f_K K_M \omega \left( \frac{\omega \tau_c}{\omega_{\odot} \tau_{c,\odot}} \right)^2, & \omega_{crit} > \omega \frac{\tau_c}{\tau_{c,\odot}}, Ro \leq Ro_{crit} \\ 0, & Ro > Ro_{crit} \end{cases}, \quad (2.16)$$

where  $K_M$  is defined in equation 1.2.3. By introducing a threshold Rossby number,  $Ro_{crit}$ , above which there is no angular braking, they are able to reproduce the observations. By fitting their model to the *Kepler* targets and the Sun they find  $Ro_{crit} = Ro_{\odot} = 2.16$ .

In order to test this theory it is imperative that we obtain new observations—in particular, rotation periods of old stars with precise and reliable ages. *LSST* may provide rotation periods for old field stars (see chapter 6. In the interim, *K2* is observing several open clusters that will add to our arsenal of precisely dated stars with rotation periods, provided we can extract rotation period signals from *K2*'s noisy light curves.

# Chapter 3

## Systematics-insensitive periodic signal search with K2

### Abstract

From pulsating stars to transiting exoplanets, the search for periodic signals in *K2* data, *Kepler*'s 2-wheeled extension, is relevant to a long list of scientific goals. Systematics affecting *K2* light curves due to the decreased spacecraft pointing precision inhibit the easy extraction of periodic signals from the data. We here develop a method for producing periodograms of *K2* light curves that are insensitive to pointing-induced systematics; the Systematics-Insensitive Periodogram (SIP). Traditional sine-fitting periodograms use a generative model to find the frequency of a sinusoid that best describes the data. We extend this principle by including systematic trends, based on a set of ‘Eigen light curves’, following Foreman-Mackey et al. (2015), in our generative model as well as a sum of sine and cosine functions over a grid of frequencies. Using this method we are able to produce periodograms with vastly reduced systematic features. The quality of the resulting periodograms are such that we can recover acoustic oscillations in giant stars and measure stellar rotation periods without the need for any detrending. The algorithm is also applicable to the detection of other periodic phenomena such as variable stars, eclipsing binaries and short-period exoplanet candidates. The SIP code is available at <https://github.com/RuthAngus/SIPK2>.

### 3.1 Introduction

The excellent precision achieved by the original *Kepler* mission relied on extremely precise pointing, for which three reaction wheels were required. After the failure of

one of these wheels, the *Kepler* team devised a new pointing scheme in which the spacecraft is stabilized by the Solar wind for ecliptic plane viewing zones (Howell et al., 2014). In this configuration the spacecraft is able to maintain an unstable equilibrium, with the two functioning reaction wheels controlling pitch and yaw whilst the spacecraft slowly rolls about the boresight. The spacecraft fires its thrusters once every  $\sim 6$  hours (Vanderburg and Johnson, 2014, hereafter VJ14) to correct for this slow drift and, as stars move across pixels with different sensitivities, their flux varies. The extraction of high-precision photometry from *K2* target pixel files, despite the reduced pointing precision, is a requirement for many fields of research and several methods for the extraction and detrending of *K2* light curves have already been developed. For example, VJ14 and Crossfield et al. (2015) use simple aperture photometry and correct the light curve of each star individually and Aigrain et al. (2015a) use a Gaussian process to model the non-linear dependence of stellar flux on the roll angle of the telescope.

Whilst these methods successfully remove most systematic trends and produce light curves suitable for exoplanet search and some stellar variability studies, residual systematics can still affect the light curves on timescales relevant to asteroseismology and stellar rotation. In particular, the  $\sim 6$  hour thruster firing signal may still appear with high power in the periodograms of these detrended light curves (see figure 3-1). A detrending method for *K2* light curves, specifically intended for the asteroseismic analysis of giant stars has been developed by Lund et al. (2015), in which the systematics due to roll are corrected, again on a star-by-star basis and any remaining periodic signals at  $47 \mu\text{Hz}$  (6 hour period) or its harmonics are removed by prewhitening. The method developed here, the Systematics-Insensitive Periodogram (SIP) produces periodograms of *K2* light curves without the need for detrending or prewhitening.

### 3.1.1 Asteroseismology

As well as providing data that have lead to the discovery thousands of exoplanets, the original *Kepler* mission revolutionized many fields of stellar astronomy, particularly

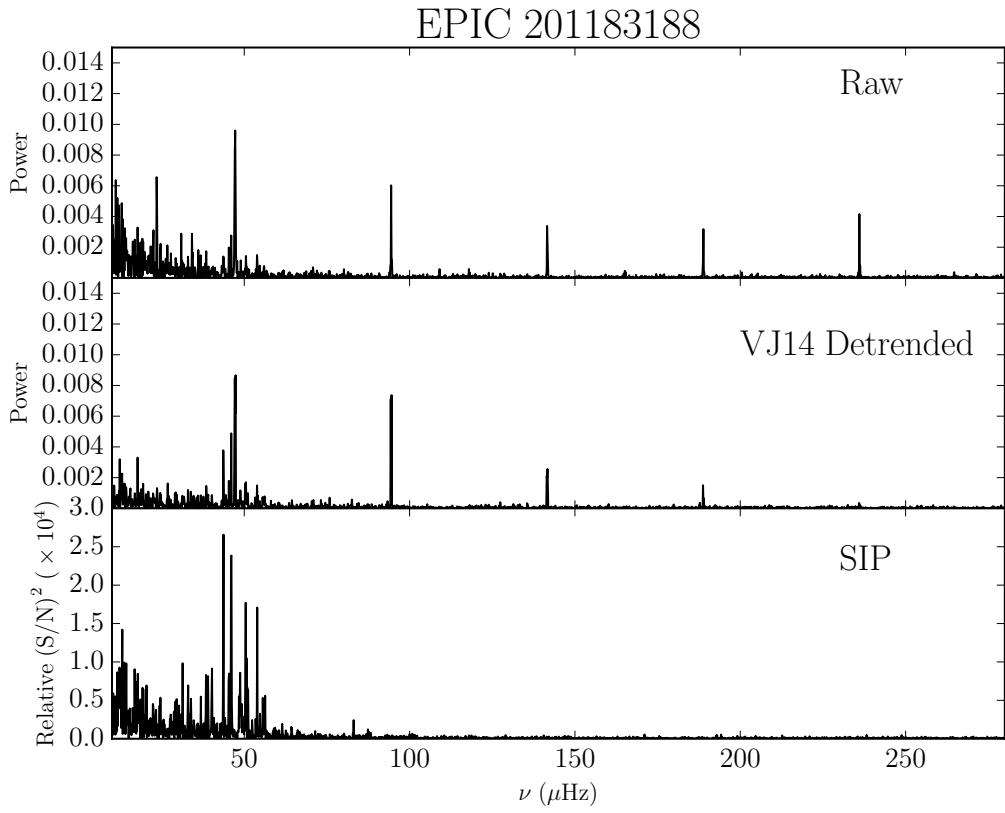


Figure 3-1: LS periodograms of the raw (Top) and VJ14-detrended (middle) *K2* light curves of EPIC 201183188. The bottom panel shows the SIP for this target. Peaks at  $\sim 47 \mu\text{Hz}$  and its harmonics produced by the regular spacecraft thruster fires are still present in the LS periodogram of the detrended data, but do not appear in the SIP.

asteroseismology. Fundamental stellar parameters—in some cases, extremely precise ones—can be calculated for *Kepler* asteroseismic stars from the power spectra of their light curves. Although Sun-like stars oscillate at high frequencies and require short-cadence observations, pulsations of giant stars lie below the Nyquist frequency set by the 28.5 minute sampling rate of long cadence *Kepler* data:  $283 \mu\text{Hz}$ . Asteroseismic analysis of data from the original *Kepler* mission is traditionally conducted upon detrended light curves. For short cadence *Kepler* data, this detrending method is described in García et al. (2011). Due to the precise pointing of the original *Kepler* mission, systematics present in these light curves, caused by temperature fluctuations and minor pointing shifts, are relatively low amplitude.

However, this is not the case for *K2* light curves: the precision over a 6 hour timescale is estimated to be 4 times worse in *K2* data (Howell et al., 2014), therefore new approaches to the treatment of systematics are necessary. Figure 3-1 demonstrates the need for careful systematics treatment of *K2* photometry for asteroseismology. The top panel shows a Lomb-Scargle<sup>1</sup> (LS) periodogram of the raw, simple aperture photometry<sup>2</sup> of EPIC 201183188, a pulsating giant star. The large peaks at  $\sim 47 \mu\text{Hz}$  and its harmonics are caused by the regular thruster fires of the spacecraft. The bottom panel shows the LS periodogram of this light curve, after it has been detrended using the method of VJ14. The large peaks are still present in the detrended light curve. The dominant source of these peaks is the removal of the outlying data points that appear in the *K2* light curves every 6 hours, caused by the spacecraft thruster fires. The rapid motion of the spacecraft results in stellar flux being smeared out over the detector and the circular apertures do not adequately encompass the resulting point spread function. It is the removal of these data points that produces the  $47 \mu\text{Hz}$  peak in the VJ14 periodogram. While the noise source at  $47 \mu\text{Hz}$  does not interfere with the detection of high-signal-to-noise transit events for periods greater than  $\sim 1$  day (Vanderburg et al., 2015), it does hamper the detection of smaller signals, particularly on time scales comparable to that of thruster fires. These peaks

---

<sup>1</sup>All LS periodograms produced in this project were produced using the gatspy Python module: <https://github.com/astroML/gatspy/tree/master/gatspy/periodic>

<sup>2</sup>The method used to extract this photometry is described in §3.2

lie in an important region of parameter space for giant star asteroseismology and could affect the stellar parameters measured for thousands of giants if not dealt with appropriately.

### 3.1.2 Stellar rotation

Stellar rotation studies have hugely benefitted from the era of high-precision space photometry. Active regions on the surface of rotating stars produce periodic variations in flux and stellar rotation periods can therefore be measured from *Kepler* light curves. Stellar rotation is a field of active interest as the rotation period of star can be used to infer its age via gyrochronology (Angus et al., 2015b; Barnes, 2007; Epstein and Pinsonneault, 2014; Skumanich, 1972), is thought to be tied to the stellar magnetic dynamo, and could even reveal dynamical interactions with companion stars or planets (e.g. Béky et al., 2014; Poppenhaeger and Wolk, 2014). Current methods for measuring rotation periods from *Kepler* light curves include periodogram (e.g. Reinhold et al., 2013), AutoCorrelation Function (ACF) (McQuillan et al., 2013a) and wavelet (e.g. García et al., 2014) analysis, or some combination thereof. Stellar variability is not typically sinusoidal, therefore sine-fitting periodograms are not perfectly suited to measuring rotation periods (McQuillan et al., 2013a). For this reason, the ACF method is often favoured over the periodogram method. However, because autocorrelation is performed directly on detrended light curves, and cannot be written down as a generative model, it is not possible use autocorrelation techniques on untreated *K2* data. A quasi-periodic Gaussian process is a much better effective model for stellar variability than a sinusoid, however we choose to focus on the more generally applicable (and computationally tractable) sine-wave periodogram, leaving the Gaussian process periodogram for future consideration.

In this article we focus on the examples of asteroseismology and stellar rotation, however many other fields of astronomy utilize periodic information in *K2* light curves. These include studies of eclipsing binaries, variable stars, exoplanets, white dwarfs and even AGN. The development of tools for extracting periodic information from *K2* data is essential if it is to be as revolutionary in time-domain astronomy as the

original *Kepler* mission was.

In §3.2 we outline the method behind the SIP. In §3.3 we apply the SIP to real *K2* light curves, using some giant asteroseismic pulsators and rotating stars as test cases and then provide the results of some simple tests which show exactly *how* ‘insensitive’ the SIP is to systematic features. Finally, we demonstrate the SIP’s usefulness regarding other periodically varying objects in this section, before presenting our conclusions in §3.4.

## 3.2 Method

The method implemented in this article is an extention of the planet-search algorithm developed by Foreman-Mackey et al. (2015) (hereafter FM15). All targets observed by *Kepler* move on the CCD in the same way, therefore the systematics affecting each individual star’s light curve have shared properties. The FM15 method uses this fact by decomposing the light curves into a set of ‘Eigen Light Curves’ (ELCs) using Principle Component Analysis (PCA), which can be used to model any individual star’s light curve with very little loss of information. This process is similar to the method used to produce PDC-MAP data for the original *Kepler* mission (Smith et al., 2012; Stumpe et al., 2012). The resulting ELCs from campaign 1 can be used to model any campaign 1 *K2* light curve, (campaign 0 ELCs for campaign 0, etc) and specifically, can model the data in combination with an arbitrary physical model.

In order to construct sets of ELCs for campaigns 0 and 1, FM15 downloaded the target pixel files for all stars in these two fields. The position of each star was predicted using the World Coordinate System (WCS) and 10 circular apertures placed around the star with radii varying from 1 to 5 pixels in steps of 0.5 pixels. Following the procedure of VJ14, the aperture producing the light curve with the lowest Combined Differential Photometric Precision (CDPP) within a 6 hour window (Christiansen et al., 2012) was selected<sup>3</sup>. PCA was then performed on the full set of targets in

---

<sup>3</sup>The simple aperture photometry light curves for campaigns 0 and 1 are available at <http://bbq.dfm.io/ketu/lightcurves/>

order to produce ELCs.

FM15 used 150 of these ELCs, plus a transit model, in order to search for exoplanet candidates without the need for a ‘detrending’ step. The number 150 was selected as a compromise between capturing as much of the variability as possible whilst maintaining computational efficiency. The likelihood of the data, conditioned on the ELC-plus-transit model was calculated over a fine grid of periods and transit depths, resulting in the detection of 36 new exoplanet candidates. We use a very similar technique to find periodic signals in K2 data. The primary difference is that we use a sinusoid rather than a transit model. This model is linear, therefore the likelihood function conditioned on a specific frequency can be calculated and the systematics model marginalized over analytically.

Following the notation in FM15, our model for the  $k$ th star can be written

$$\mathbf{f}_k = \mathbf{A}\mathbf{w}_k + \text{noise}, \quad (3.1)$$

where  $\mathbf{f}_k$  is the vector of  $N$  flux values,

$$\mathbf{f}_k = (f_{k,1}, f_{k,2}, f_{k,3}, \dots, f_{k,N})^T \quad (3.2)$$

at times

$$\mathbf{t}_k = (t_1, t_2, t_3, \dots, t_N)^T. \quad (3.3)$$

$\mathbf{A}$  is the design matrix:

$$\mathbf{A} = \begin{pmatrix} x_{1,1} & x_{2,1} & \cdots & x_{150,1} & 1 & \sin(2\pi\nu t_1) & \cos(2\pi\nu t_1) \\ x_{1,2} & x_{2,2} & \cdots & x_{150,2} & 1 & \sin(2\pi\nu t_2) & \cos(2\pi\nu t_2) \\ \vdots & & & & & & \\ x_{1,N} & x_{2,N} & \cdots & x_{150,N} & 1 & \sin(2\pi\nu t_N) & \cos(2\pi\nu t_N) \end{pmatrix} \quad (3.4)$$

where the  $x_{ij}$ s are the ELCs<sup>4</sup>, with  $i$  denoting the ELC number and  $j$  the time index. The design matrix contains the basis functions of the linear model. The basis

---

<sup>4</sup>Campaign 0 and 1 ELCs are available at <http://bbq.dfm.io/ketu/elcs/>

functions for the systematic features in the light curves are the ELC values at each time index, the sine and cosine terms are the basis functions of the sinusoidal signal of interest, and the ‘1’s describe a linear offset. Any  $K2$  light curve can be reproduced as a linear combination of these basis functions. We are interested in the last two elements of the weight vector: the coefficients of the sinusoidal signal. The maximum likelihood solution for the weight vector,  $\mathbf{w}$  is

$$\mathbf{w}_k^* \leftarrow (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{f}_k. \quad (3.5)$$

Under this linear model with Gaussian uncertainties, the marginalized likelihood for the periodic amplitude is a two-dimensional Gaussian with mean given by the last two elements ( $\mathbf{a}$ ) of  $\mathbf{w}^*$  and covariance given by the bottom right two-by-two block ( $\mathbf{S}_a$ ) of  $(\mathbf{A}^T \mathbf{C}^{-1} \mathbf{A})^{-1}$ , where the  $\mathbf{C}$  matrix contains observational uncertainties on the diagonal. These uncertainties are estimated as  $1.48 \times$  the Median Absolute Deviation (MAD), following Aigrain et al. (2015a). Therefore, the signal-to-noise ratio,  $S/N$  of the amplitude measurement is  $\sqrt{\mathbf{a}^T \mathbf{S}_a^{-1} \mathbf{a}}$ . The  $(S/N)^2$  can then be calculated over a user-defined grid of frequencies to produce a SIP. The  $(S/N)$  operation takes into account the goodness of fit, i.e. if the amplitude of the sinusoid at a given frequency is not well constrained, it is penalized. The SIP algorithm scales linearly with the number of frequencies evaluated and with the cube of the number of basis functions used. As an example of the typical computation time, calculating a single SIP for a  $K2$  campaign 1 light curve, over a grid of 1000 frequencies with 150 basis vectors, takes 2-3 seconds on a 2.7 GHz CPU.

### 3.3 Application to real light curves

An example LS periodogram of the raw  $K2$  photometry for giant star, EPIC 201183188 is shown in figure 3-1. Peaks appearing at  $47 \mu\text{Hz}$  and its harmonics are produced by the regular  $\sim 6$  hour thruster fires that repoint the spacecraft. These peaks are also present in periodograms of the VJ14 detrended light curves. The presence of

systematic signals at these timescales are problematic for asteroseismic analysis since they lie in a region of frequency space that is often populated by giant asteroseismic modes. It is possible to remove these signals by ‘prewhitening’ the data, i.e. subtracting a sinusoid of that frequency from the data, however this process will artificially suppress all signals, both systematic and astrophysical, at that frequency. The SIP method eliminates the necessity for any such procedure. The bottom panel of figure 3-1 shows the SIP for the same star, demonstrating the ability of the SIP method to produce periodograms that are free from thruster firing signals.

In order to search for high signal-to-noise asteroseismic modes in the giant star candidates of GO1059, we searched for a power excess in the SIPs using the method of Huber et al. (2009): autocorrelation functions were calculated for sections of the SIP in order to search for regions of increased correlation and locate the frequency of maximum power. The increased correlation arises from the even frequency spacing of acoustic modes, and the frequency of maximum correlation at the location of the power excess corresponds to the large frequency separation,  $\Delta\nu$ . Figures 3-2(a) to 3-2(e) show example power spectra of 6 targets for which we detect pulsations using this method.

The top panel of figure 3-3 shows the VJ14-detrended light curve of an active, rotating star, EPIC 201133037, with a linear trend subtracted off. The brightness fluctuations clearly visible in the light curve of this target are produced by cool active regions on the stellar surface, which reduce the stellar flux periodically. The rotation period of this star is therefore around 20 days. The middle and bottom panels show an ACF and LS periodogram of the detrended light curve. The top panel of figure 3-4 shows the raw light curve of the same target in grey, with the conditioned light curve in black. This conditioned light curve was produced by removing the best fitting systematic trends, described by a certain combination of the ELCs, at the best fitting period of the sinusoid. The bottom panel shows the SIP.

Each of these three methods measures a rotation period of around 20 days for this target. This example demonstrates the ability of the SIP to recover rotation periods that agree with those measured from detrended light curves by autocorrelation.

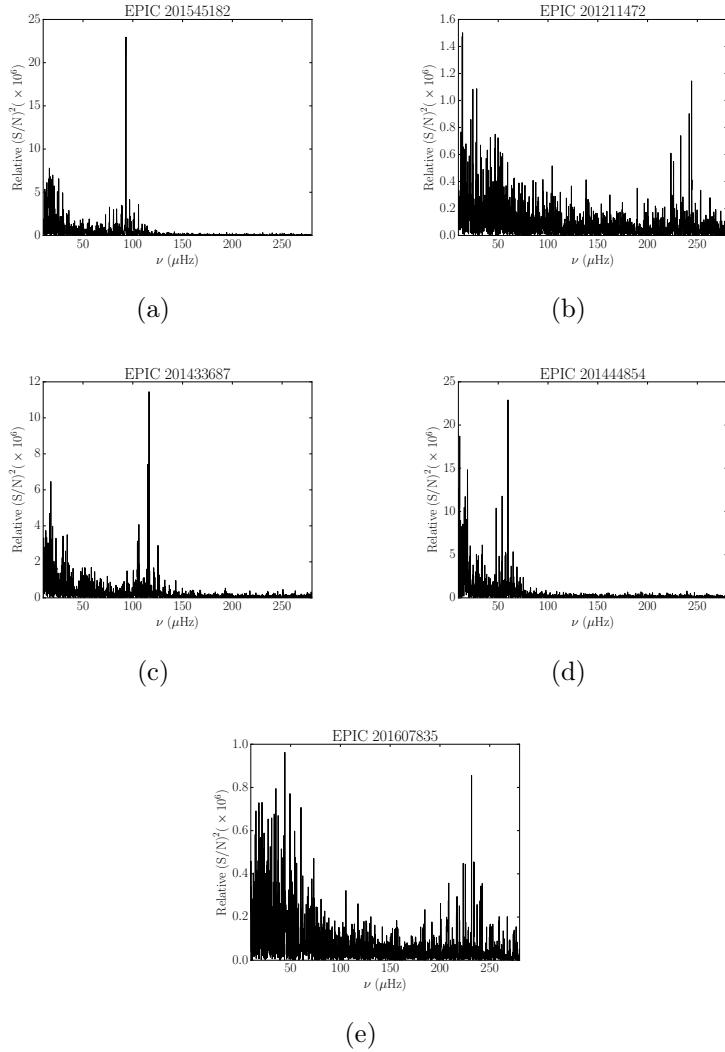


Figure 3-2: SIPs of 6 long cadence *K2* giants with asteroseismic oscillations. These were selected from the guest observing program, GO1059 and identified using the method of Huber et al. (2009).

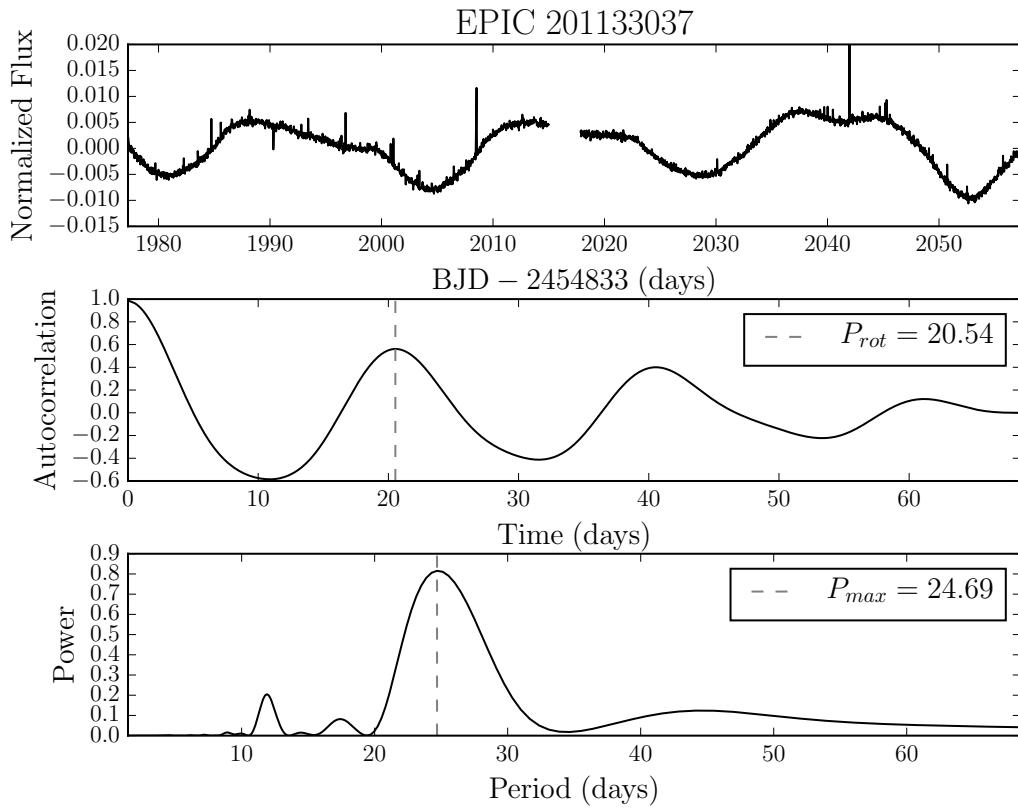


Figure 3-3: *Top:* Light curve of EPIC 201133037, detrended using the method of VJ14. *Middle:* Autocorrelation function of the detrended light curve. The autocorrelation function method measures a rotation period of 21 days for this star. *Bottom:* The LS periodogram of the detrended light curve. The highest peak in the periodogram is located at 25 days.

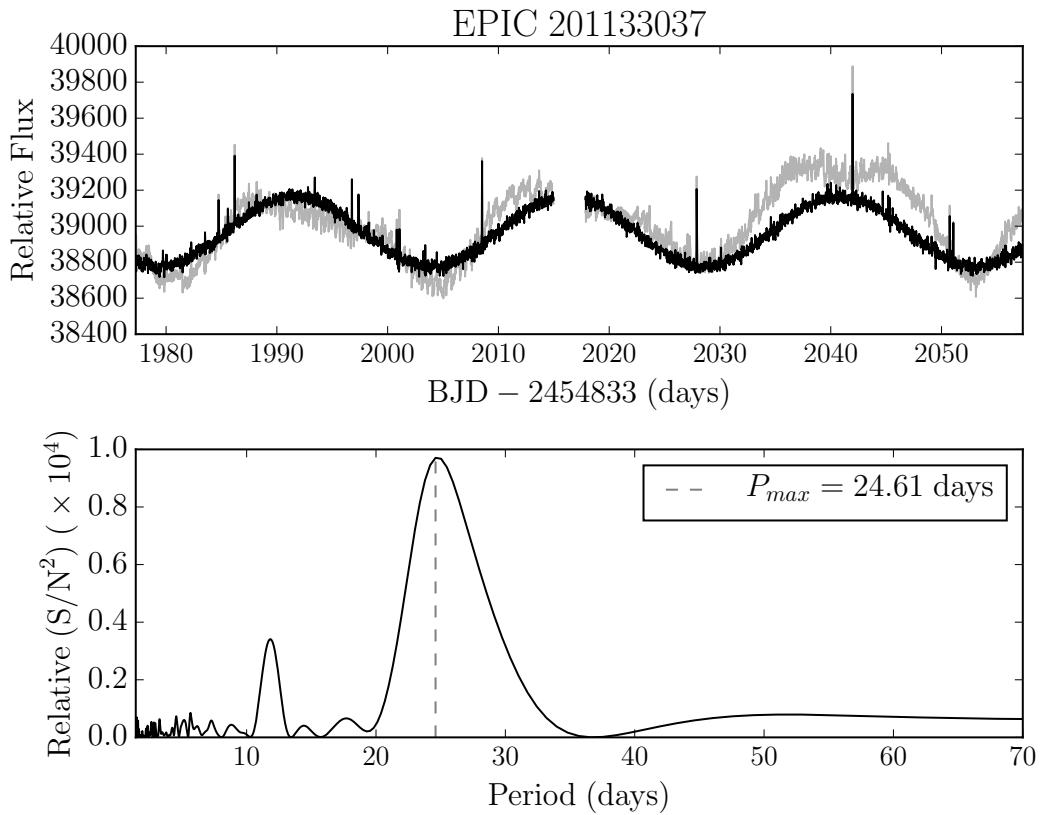


Figure 3-4: *Top:* The raw light curve of EPIC 201133037 is shown in grey and the conditioned light curve is shown in black. The conditioned light curve is produced by removing the trends that best describe the data, at the best fitting frequency. *Bottom:* An SIP of the raw light curve, produced by modelling the data using the top 150 ELCs plus a sine and cosine function at a range of frequencies. The highest peak in the SIP is located at 25 days.

We also include an example that demonstrates the ability of the SIP to outperform a periodogram of detrended data. Figure 3-5 shows the light curve, ACF and LS periodogram of another rotating star, EPIC 201142043 and figure 3-6 shows its SIP. This star shows lower amplitude variability than the previous example and the careful treatment of systematics is much more important. Whereas the ACF method is able to measure a rotation of  $\sim 3$  days for this star, the LS periodogram of the detrended light curve incorrectly measures a period of 59 days. Although there is a small peak at the rotation period of the star, it is not the dominant periodic signal. The SIP method is, by definition, insensitive to these long-term systematics and is able to measure a period of  $\sim 2$  days. This example further demonstrates the fact that long-term systematic trends caused by slow pointing variations are often not removed by conventional detrending methods. The 59 day signal is almost certainly a systematic trend and not an astrophysical signal because it does not appear in the SIP. It is well described by the ELCs and must therefore be common to many stars. Although these examples demonstrate that the SIP can provide rotation periods, in practice it is likely to suffer from overfitting. Several ELCs contain smoothly varying signals on time scales of a few days to tens of days which look similar to signals produced by stellar rotation. In some cases it may be possible to reproduce stellar rotation signals in a *K2* light curve using only the ELCs, with no additional sinusoid. We are still investigating the nature of this potential over-fitting problem, but a solution may to regularise the SIP — apply priors to the weights of the ELCs.

We have shown that the SIP method is able to measure stellar rotation periods and does better than producing periodograms from detrended data. However, it has been shown that the ACF method often performs better than periodogram methods in general for measuring stellar rotation periods (Mazeh et al., 2015; McQuillan et al., 2013a, 2014). For stars with relatively high-amplitude variability, for which perfect removal of systematic trends is less important, performing the ACF method on detrended data is likely to produce similar results to the SIP method. The SIP method is ideally suited to low-amplitude cases, where systematic trends could drastically influence rotation period measurements. Whilst the SIP method may outperform ACF

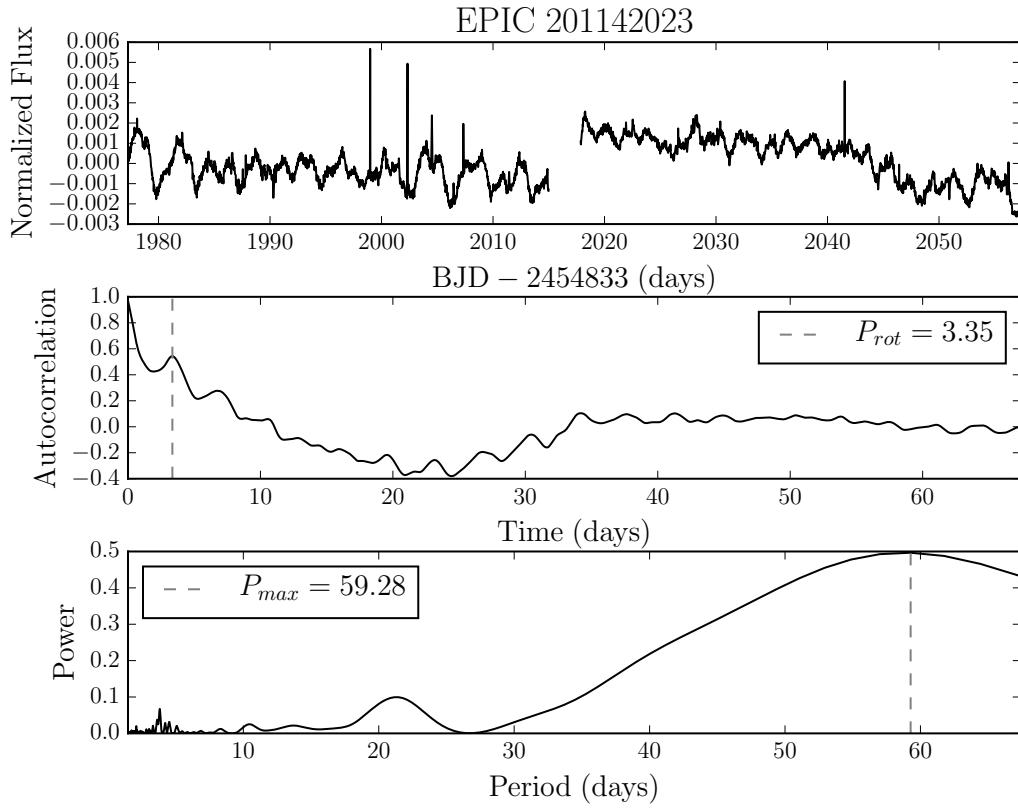


Figure 3-5: *Top:* Light curve of EPIC 201142043, detrended using the method of VJ14. *Middle:* Autocorrelation function of the detrended light curve. The autocorrelation function method measures a rotation period of 3 days for this star. *Bottom:* The LS periodogram of the detrended light curve. The highest peak in the periodogram is located at 59 days and is likely to be a systematic trend produced by spacecraft pointing variations.

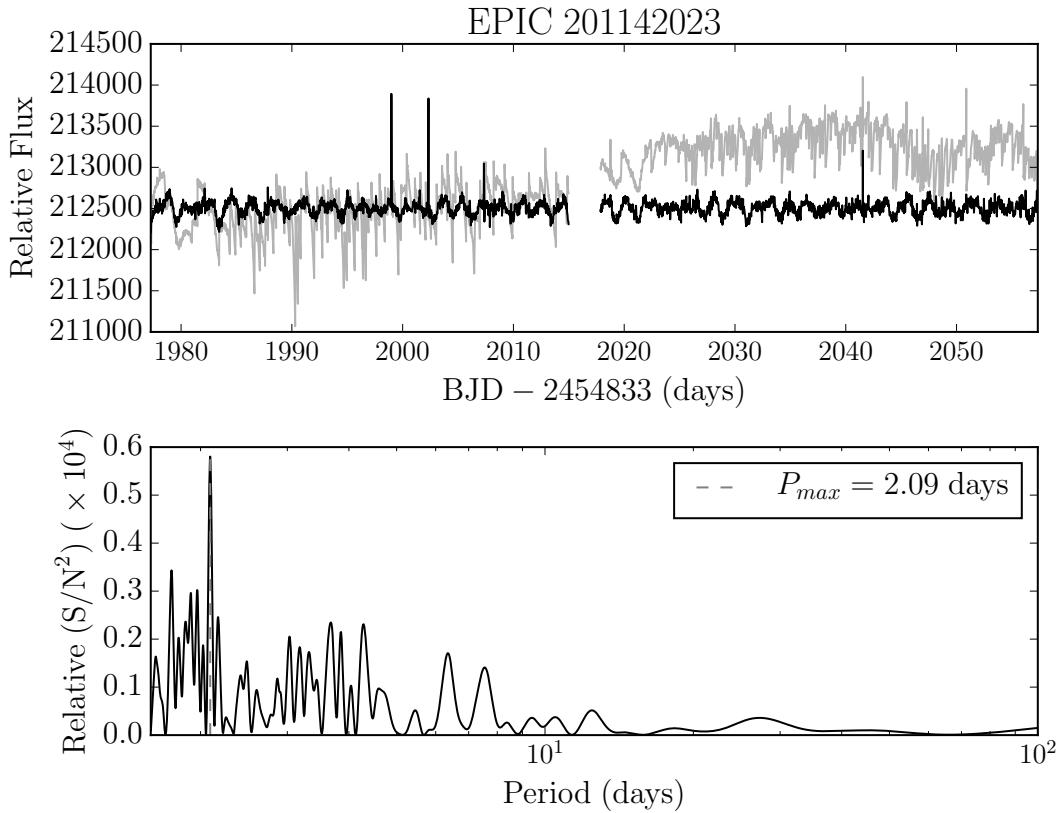


Figure 3-6: *Top:* The raw light curve of EPIC 201142043 is shown in grey and the conditioned light curve is shown in black. The conditioned light curve is produced by removing the trends that best describe the data, at the best fitting frequency. *Bottom:* An SIP, produced by modelling the data as a linear combination of the top 150 ELCs plus a sine and cosine function at a range of frequencies, measuring a rotation period of 2 days. The SIP is, by definition, insensitive to the long-timescale systematics that dominate the LS periodogram of the detrended data, shown in figure 3-5.

in the low-amplitude cases, any ‘marginal’ rotation period measurements calculated using either method should be treated with caution unless a representative uncertainty is provided. In general neither ACF nor periodogram methods are equipped to provide such uncertainties. In practice, we recommend using both the SIP and ACF methods, in combination with a by-eye check, to measure rotation periods for *K2* stars.

### 3.3.1 Tests and discussion

In order to demonstrate the consistent ability of the SIP method to remove the signal at  $47\mu\text{Hz}$ , corresponding to the periodic  $\sim 6$  hour thruster firings, we computed SIPs for 4923 targets from the GO1049 proposal: “Galactic Archaeology on a grand scale” (PI: Stello, D.). For each target, an SIP of its raw photometry and a LS periodogram of its VJ14 light curve was calculated for frequencies between 40 and  $54\mu\text{Hz}$ . Both the height and frequency of the highest peak in the SIP and the highest peak in the LS periodogram were recorded. A histogram of the frequencies of the highest peaks in the SIPs of all 4923 targets is shown in the top panel of figure 3-8. The bottom panel shows the histograms of peak heights within the correspondingly colored ranges indicated in the top panel. This figure shows that while there are a greater number of maximum peaks around  $47\mu\text{Hz}$ , the S/Ns of these peaks are comparable to those found just above and just below this frequency. Figure 3-9 shows the equivalent results for the VJ14 light curves. There is a significant number of large peaks at  $\sim 47\mu\text{Hz}$  in the LS periodograms of the detrended light curves; the highest peak in the LS periodograms was almost always located at  $\sim 47\mu\text{Hz}$ . Furthermore, the distribution of peak power within the range  $46.5\text{--}48\mu\text{Hz}$  is skewed towards higher powers, i.e. a substantial fraction of the peaks at  $\sim 47\mu\text{Hz}$  have a large power. The SIP method is able to consistently remove the  $47\mu\text{Hz}$  signal which is present in almost every VJ14 light curve.

We performed an injection and recovery experiment in order to test the detection efficiency of the SIP algorithm. 4000 sinusoids with frequencies ranging from 10-270  $\mu\text{Hz}$  and amplitudes ranging from 1 to 100 parts-per-million were injected into the

raw *K2* light curve of target star EPIC 201121245, a relatively non-variable giant with low-amplitude acoustic oscillations. In order to recover the injected signals we calculated a SIP of the original light curve, subtracted this from the SIP of injected light curve, and searched for excess power in the residuals. This allowed us to perform injection and recovery tests on this target star without being affected by the star's own intrinsic variability. We then measured the position of the highest peak in the resulting residual SIP and recorded the successful detections, defined as those that lay within  $1 \mu\text{Hz}$  of the injected value. SIPs were computed over a grid of frequencies ranging from  $10$ - $270 \mu\text{Hz}$ , with a spacing of  $0.1 \mu\text{Hz}$ . The resulting detection efficiency map is shown in figure 3-7. This figure demonstrates that the SIP can recover the frequencies of signals with amplitudes less than 10 ppm.

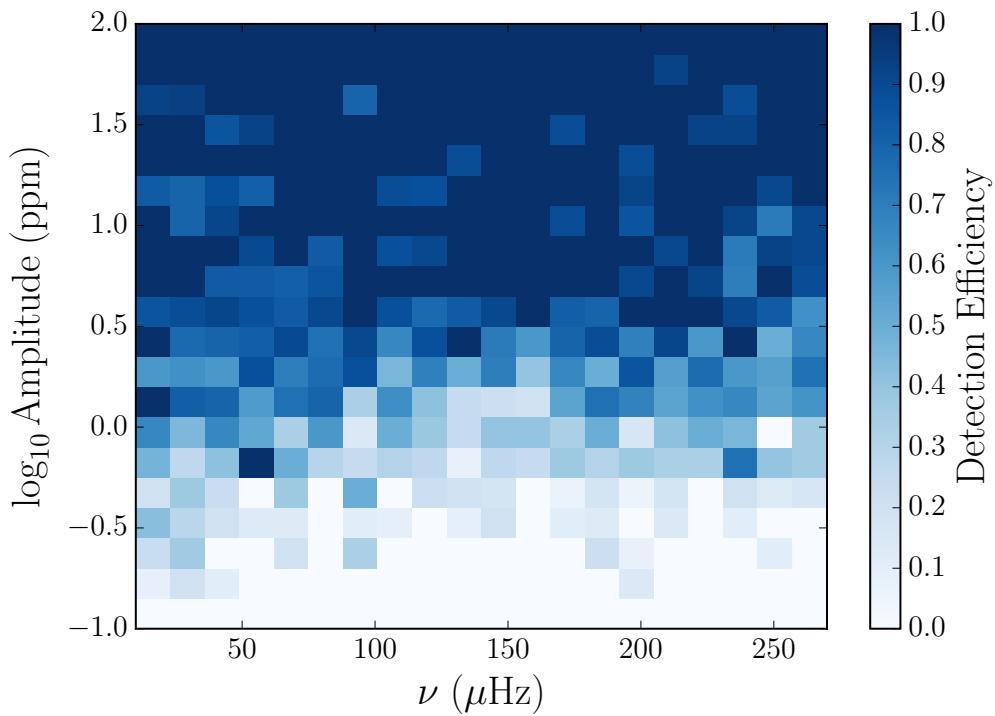


Figure 3-7: A map of the detection efficiency of the SIP algorithm as a function of frequency and injection amplitude. The SIP is capable of recovering signals with amplitudes less than 10 ppm.

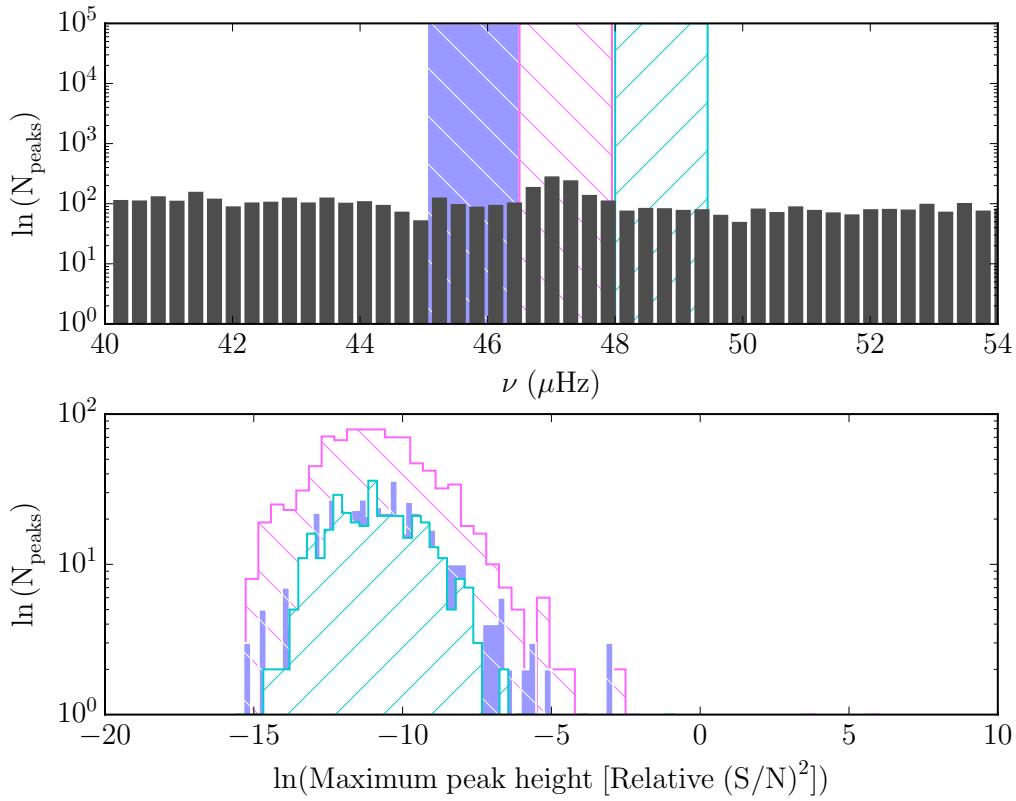


Figure 3-8: *Top:* Histogram of the frequencies of the highest peaks in the SIPs of 4923 *K2* targets within the range 40–54  $\mu\text{Hz}$ . *Bottom:* Histograms of peak heights within the correspondingly colored ranges indicated in the top panel. Whilst there is a larger number maximum peaks around 47  $\mu\text{Hz}$  (the frequency corresponding to the 6 hour thruster fire) the amplitudes of these maximum peaks are comparable to the maximum peak heights just above and just below this frequency.

Figure 3-10 shows the conditioned light curve and SIP of the short-period, disintegrating planet candidate discovered by Sanchis-Ojeda et al. (2015). This planet has a period of around 9 hours, short enough to be detectable with a periodogram, as was demonstrated for a number of ultra-short period *Kepler* exoplanets by Sanchis-Ojeda et al. (2014). The top panel of this figure shows the *K2* light curves of these objects, conditioned on the highest S/N sinusoidal signal in the periodograms. The light curve was produced by subtracting the trends that best describe the data at the highest S/N period found in the SIP.

Photometric variability in dwarf stars on timescales less than 8 hours, often known

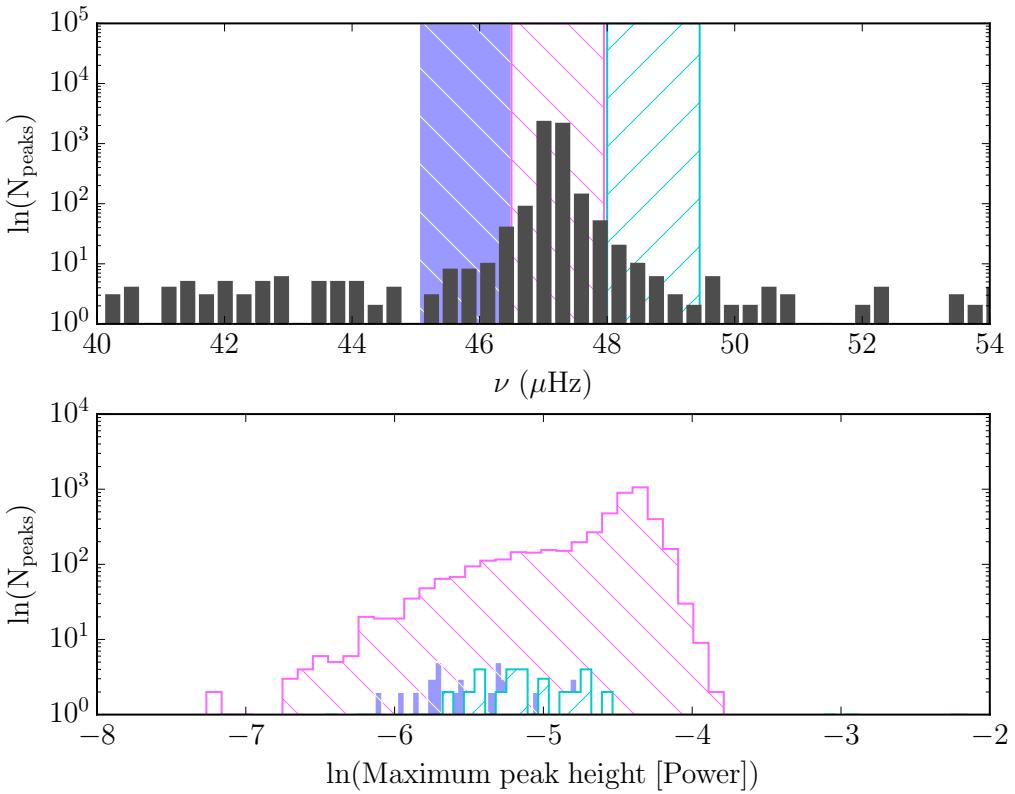


Figure 3-9: *Top:* Histogram of the frequencies of the highest peaks in the LS periodograms of the Vanderburg and Johnson (2014) light curves of 4923  $K2$  targets within the range 40–54  $\mu\text{Hz}$ . *Bottom:* Histograms of peak heights within the correspondingly colored ranges indicated in the top panel. The frequency of maximum peak height was  $\sim 47 \mu\text{Hz}$  in almost every periodogram. Furthermore, the distribution of maximum peak height within the range 46.5–48  $\mu\text{Hz}$  is skewed towards higher powers, i.e. a large fraction of the peaks at  $\sim 47 \mu\text{Hz}$  have a large power.

as flicker, has been linked to surface gravity (Bastien et al., 2013; Kipping et al., 2014). The metrics used to quantify photometric variability include finding the range in intensity, counting the number of zero crossings and calculating the root-mean-square (RMS) of the light curve. Although these features are related to signal processing, they are operations performed on detrended light curves, not inferred from periodograms. However, it may be possible to derive a property of the periodogram that scales with the density or surface gravity of a star, for example, the mean excess power at frequencies near those relevant to granulation timescales. The SIP method

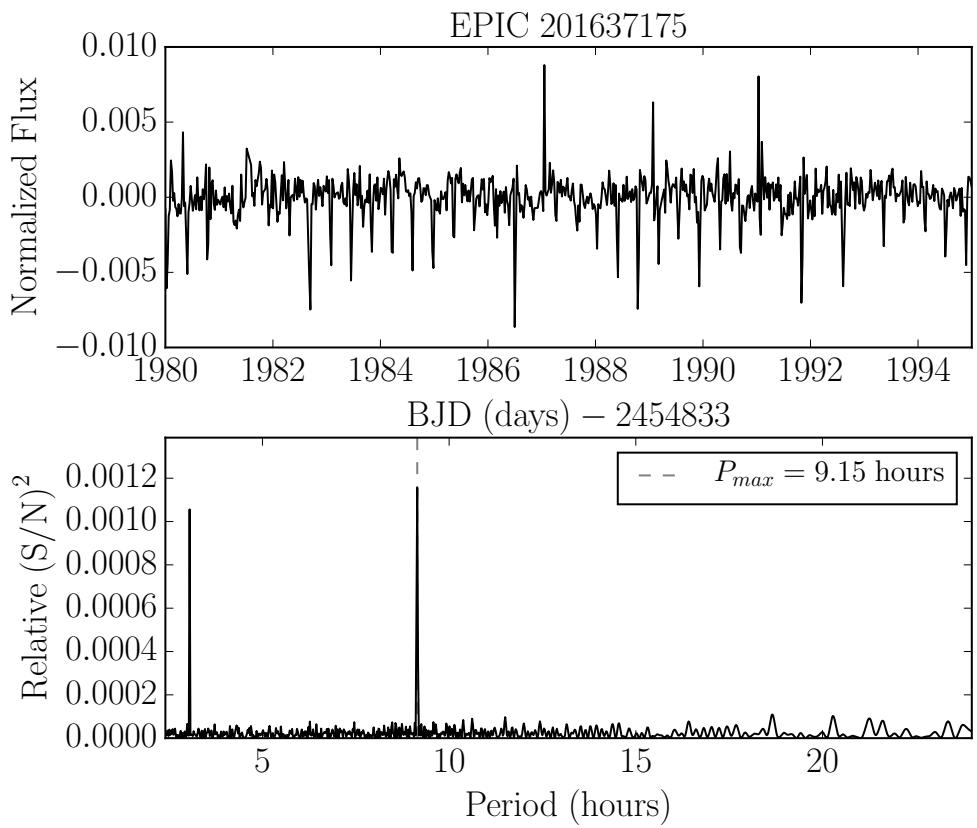


Figure 3-10: *Top:* The light curve of exoplanet candidate, EPIC 201637175, conditioned on the highest amplitude sinusoidal signal found in the periodogram. *Bottom:* The systematic-insensitive periodogram of this target.

presented here would be useful for such a technique.

### 3.4 Conclusions

We demonstrate that modelling campaign 1 *K2* photometry as a linear combination of 150 PCA components plus a sinusoid can produce periodograms that are almost completely free from instrumental systematic signals, without the need for detrending. We find that the  $47 \mu\text{Hz}$  signal, generated by the spacecraft thruster fires is not present in the vast majority of Systematics-Insensitive Periodograms (SIPs) for more than 4000 targets selected from the *K2* guest observer program, GO1059, "Galactic Archaeology on a grand scale" (PI: Stello, D.). The SIP method is highly successful for campaign 1 targets where the large number of stars, observed for a baseline of 80 days ensures that most of the systematics are captured in the ELCs and we anticipate that it will be equally effective for the up-and-coming campaigns.

The SIP method is capable of detecting periodicities in *K2* data in the region of frequency space relevant to the study of asteroseismic oscillations in giant stars and for any signals with a timescale close to 6 hours. It may be useful for measuring stellar rotation periods and is an improvement upon a simple LS periodogram of detrended data, but rotation period inference with the SIP may suffer from over-fitting. We intend to address this problem in future by developing a regularised version of the SIP. In practice, the best approach for measuring rotation periods in *K2* data is likely to be a combination of the SIP method and the ACF method, where autocorrelation is performed on detrended light curves.

Of course, as mentioned in chapter 1, a sinusoid is actually *not* a good model for stellar rotation signals. For this reason, the SIP's usefulness for measuring rotation periods is limited. In the following chapter I introduce an alternative method for inferring rotation periods, using a semi-parametric function (a Gaussian process). It will be possible to combine this Gaussian process model with the same noise model used by the SIP to produce a GPSIP (Gaussian Process Systematics-Insensitive Periodogram). Indeed, any model could be used in place of a sinusoid, although linear models have an advantage in computational efficiency and a GPSIP would be extremely expensive. Still, it would be worth exploring the performance of the GPSIP

in the future.

The SIP code is available for public use and can be found at <https://github.com/RuthAngus/SIPK2>.



# Chapter 4

## Stellar rotation period inference using Gaussian processes

### Abstract

The light curves of spotted, rotating stars are often non-sinusoidal and quasi-periodic—spots are not static on the stellar surface and have finite lifetimes causing stellar flux variations to slowly shift in phase. A strictly periodic sinusoid is therefore not a representative generative model for the light curve of star with a rotational signal. Ideally, a physical model of the stellar surface would be conditioned on the data, however the parameters of such models can be highly degenerate (*e.g.* Jeffers and Keller, 2009; Russell, 1906). Instead, we use an appropriate *effective* model: a Gaussian Process (GP) with a quasi-periodic covariance kernel function. By modelling the covariance matrix of the light curve with a GP, a highly flexible semi-parametric function, we avoid the necessity to choose a ‘best fitting’ functional form, whilst sampling directly from the posterior probability distribution function of the periodic parameter and marginalising over the other kernel hyperparameters. We used 300 simulated light curves with a range of rotation periods to test the GP model. We attempted to recover the rotation periods using three methods: our GP method, a sine-fitting periodogram method and an autocorrelation function method. The posterior probability distribution of the rotation period parameter was sampled using the affine invariant ensemble MCMC sampler `emcee` (Foreman-Mackey et al., 2013), and the GP operations were performed using the `george` python package (Foreman-Mackey et al., 2014b). Rotation periods inferred via this method are more precise and accurate than the periodogram and ACF methods. Furthermore, the improvement is expected to be even more dramatic when applied to real, noisy *Kepler* light curves, since the GP method is well suited to modelling rotation signals and correlated noise simultaneously.

## 4.1 Introduction

The light curves of spotted, rotating stars are often non-sinusoidal and Quasi-Periodic (QP). These stars vary in brightness due to active regions on their surfaces which rotate in and out of view. The non-sinusoidal quality is caused by the complicated surface spot patterns and the quasi-periodicity is produced both by the finite lifetimes of these active regions and the presence of differential rotation on the stellar surface. A strictly periodic sinusoid is therefore not a good model for stellar light curves. In an ideal world, a physical model of the stellar surface would be conditioned on the data. A physically realistic, generative model would perfectly capture the complexity of shapes within stellar light curves as well as the quasi-periodic nature, allowing for extremely precise probabilistic period recovery. However, such physical models require many free parameters in order to accurately represent a stellar surface and some of these parameters are extremely degenerate. In addition to global stellar parameters such as inclination and rotation period, each spot or active region should have (at minimum) a longitude, latitude, size, temperature and lifetime. Considering that many stars have on the order of hundreds of spots, the number of free parameters quickly becomes unwieldy, especially if the posterior PDFs of these parameters are explored with MCMC. Simplified spot models, such as the one described in Lanza et al. (2014) where only two spots are modelled, have produced successful results, however these simplified models sacrifice some precision due to lack of model flexibility. Instead of using a physical model for stellar light curves, we choose to use an *effective* model: one which captures the behaviour but is not physically motivated, although the parameters of this model may be *interpreted* as physical ones. An ideal effective model for the light curve of a spotted, rotating star is one with a small number of non-degenerate parameters that is flexible enough to perfectly capture non-sinusoidal, QP behaviour. These requirements are perfectly fulfilled by a Gaussian process (GP) model.

The standard methods used for measuring rotation periods include detecting peaks in a Lomb-Scargle (Lomb, 1976; Scargle, 1982) (LS) periodogram (e.g. Reinhold et al.,

2013), Auto-Correlation Functions (ACFs) (McQuillan et al., 2013a) and wavelet transforms (García et al., 2014). The precision of the LS periodogram and wavelet methods are limited by the suitability of the model choice. A sinusoid is used in the case of the LS periodogram and the wavelet method relies on a choice of mother wavelet that is assumed to describe the data over a range of transpositions (see, *e.g.* Carter and Winn, 2010). In contrast, the ACF method is much better suited to signals that are non-sinusoidal. In fact, it doesn't matter what shape the signal is: as long as it is approximately periodic the ACF will display peaks located at the rotation period. A drawback of the ACF method however, is that it requires data to be evenly-spaced<sup>1</sup> which is not the case with *Kepler* light curves (although in many cases it can be approximated as uniformly sampled). An ACF is also an operation performed on the data, not a generative model of the data and is therefore not inherently probabilistic. For this reason it is very difficult to estimate the uncertainty on an ACF rotation period measurement. Many rotation periods in the literature have been inferred by measuring the position of the first peak in an ACF, however this approach can be dangerous. The exponential decay in correlation can shift the peak position shortwards of its true value, leading to an underestimate of the period. We return to this point in section §5.3.

The motivation for developing a GP-based method for rotation period inference is, firstly to measure more accurate and precise rotation periods using a better-suited generative model than a sinusoid for the reasons explained above. Secondly, in order to infer *probabilistic* periods, i.e. to estimate the posterior PDF of the rotation period and thereby measure a realistic and representative uncertainty. And thirdly, to allow for an additional correlated noise model to be included during regression, the parameters of which could be marginalised over.

GPs are commonly used in the machine learning community and increasingly in other scientific fields, for example biology and geophysics (where GP regression is called ‘kriging’). More recently, GPs have been used in the astronomy literature (see *e.g.* Aigrain et al., 2016; Evans et al., 2015; Gibson et al., 2012; Haywood, 2015; Hay-

---

<sup>1</sup>Edelson and Krolik (1988) describe a method for computing ACFs for unevenly-spaced data.

wood et al., 2014; Rajpaul et al., 2015, 2016). They are useful in regression problems involving any stochastic process, specifically when the probability distribution for the process is a multi-variate Gaussian. If the probability of obtaining a data-set is a Gaussian in  $N$  dimensions, where  $N$  is the number of data points, that data-set can be described as, and with, a Gaussian process. An in-depth introduction to Gaussian processes is provided by Rasmussen and Williams (2005).

GP models parameterise the covariance between data points and a kernel function provides the form of covariance matrix parameterisation. For example, take the time-series in figure 4-1. This is the *Kepler* light curve of KID 1430163, a relatively active star that rotates once every  $\sim 4$  days. The stochastic variability in this time-series is typical of *Kepler* FGK stars. Clearly, data points in this light curve are correlated. Points that are close together in time are tightly correlated and those that are far apart are loosely correlated. It is the way in which the covariance varies with the separation between data points that is modelled when using GP regression. It is not the data but the *covariance structure* that is modelled. This fact is what gives GPs their flexibility—they can model any time series with a similar covariance structure. In addition, a very simple function can usually capture the covariance structure of a light curve, whereas a much more complicated one may be required to model the time-series itself. The light curve of this star has been modelled with two different GP kernel functions in figure 4-1, shown in blue and pink. Both provide adequate fits to the data, however only the periodic kernel function, ‘QP’ is a useful model because it has a periodic parameter. I return to this point shortly.

A range of kernel functions could be used to describe stellar variability. For example, the simplest and most commonly used kernel function, the ‘Squared Exponential’ (SE) produces an adequate fit to the KID 1430163 light curve. The SE kernel function is defined as,

$$k_{i,j} = A \exp\left(-\frac{(x_i - x_j)^2}{2l^2}\right). \quad (4.1)$$

Here  $A$  is the amplitude of covariance,  $l$  is the length scale of exponential decay and  $x_i - x_j$  is the separation between data points. The SE kernel function has the

advantage of being very simple—it has just two parameters, a covariance amplitude and length scale:  $A$  and  $l$ . If  $l$  is large two data points far apart in  $x$  will be tightly correlated, and if small they will be loosely correlated. Another property of the SE kernel function is that it produces functions that are infinitely differentiable. It is therefore possible to model a data set and its derivatives simultaneously. The SE kernel function may be an adequate model of the covariance in stellar light curves, but it is not a *useful* one for the problem of rotation period inference because it does not have a parameter that describes a period. In order to infer rotation periods it is necessary to use a periodic kernel function. For this reason, we use the ‘Quasi-Periodic’ kernel function. The QP kernel function is defined as

$$k_{i,j} = A \exp \left( -\frac{(x_i - x_j)^2}{2l^2} - \frac{\sin^2(\frac{2\pi}{P})}{\Gamma^2} \right). \quad (4.2)$$

It is the product of the SE kernel function, which describes the overall covariance decay, and an exponentiated, squared, sinusoidal kernel function that describes the periodic covariance structure.  $P$  can be interpreted as the rotation period of the star and  $\Gamma$  controls the number of zero crossings per period. This kernel function allows two data points that are separated in time by one rotation period to be tightly correlated, while points separated by half a period can be weakly correlated. Both the QP and SE kernel functions were used to produce the blue and pink models shown in figure 4-1.

In order to infer a stellar rotation period from a light curve, we fit a GP model with a QP kernel function to the data. The likelihood of the model, conditioned on the data could then be maximised in order to find the maximum likelihood value for  $P$ . In this study however, the full posterior PDFs of the parameters are explored using MCMC. This latter approach comes at a cost: a GP model is expensive to compute once as it requires a matrix inversion and determinant evaluation, let alone the many thousands of times that is necessary to fully explore the posteriors of the parameters. However, fully exploring the posterior PDF of  $P$ , allows for an accurate estimate of the uncertainty on the rotation period.

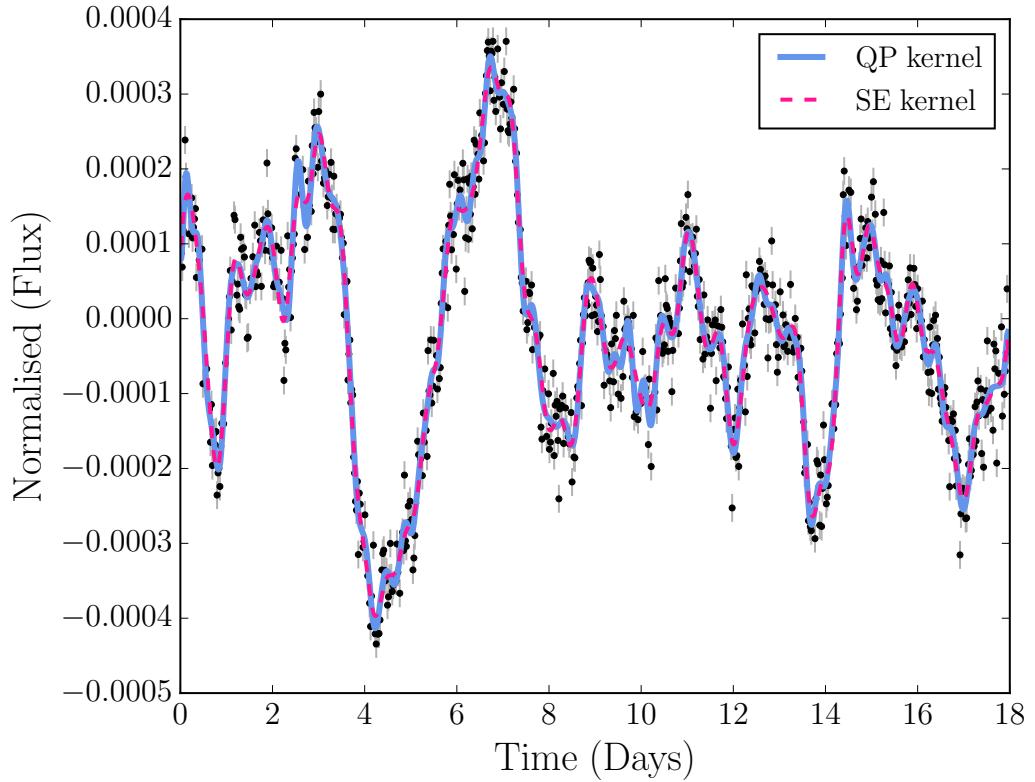


Figure 4-1: Light curve of KID 1430163, an active star with a rotation period of  $\sim 4$  days. The blue solid line shows a fit to the data using a Gaussian process model with a QP covariance kernel function and the pink broken line shows a fit with a SE kernel function. Both kernel functions provide an adequate fit to the data, however we use the QP function because it has a useful parameter that corresponds to the star's rotation period.

## 4.2 Rotation period injection and recovery

In order to test the GP method we attempted to measure rotation periods of a set of simulated light curves. We simulated 300 light curves using a spot model similar to that used for the Aigrain et al. (2015a) ‘hare and hounds’ rotation period recovery experiment. These light curves were simulated by placing dark, circular spots with slowly evolving size, on the surface of bright, rotating spheres, ignoring limb-darkening effects. Aigrain et al. (2015a) simulated one thousand light curves in order to test the ability of participating teams to recover both the stellar rotation periods and the rotational shear: the amplitude of surface differential rotation. Unlike in their study we are not interested in recovering differential rotation in this work, so did not include it in our simulations. We opted to use only solid-body rotators because differential rotation may produce some additional scatter in the rotation period measurements recovered. This code can also be adjusted to produce more realistic light curves by altering spot lifetimes. Stars with spot lifetimes that are long relative to their rotation periods will have highly periodic light curves. Those with short spot lifetimes will be more quasi-periodic. We fixed the mean spot lifetime at an arbitrary value of 30.5 days for all light curves in these initial tests. In future we plan to include both differential rotation and variable spot lifetimes in our light curve model. Light curves were simulated with a real *Kepler* long-cadence time array, with one data point every thirty minutes over a four year duration. The rotation periods of the simulations were randomly drawn from a log-uniform distribution between 0.5-60 days. Figure 4-2 shows an example of a simulated, noise-free light curve with a period of 17.4 days.

Initial tests were conducted on noise-free light curves, in order to provide proof-of-concept. We attempted to recover the rotation periods of these 300 light curves using three rotation period recovery methods: the ACF method, the LS periodogram method and the GP method.

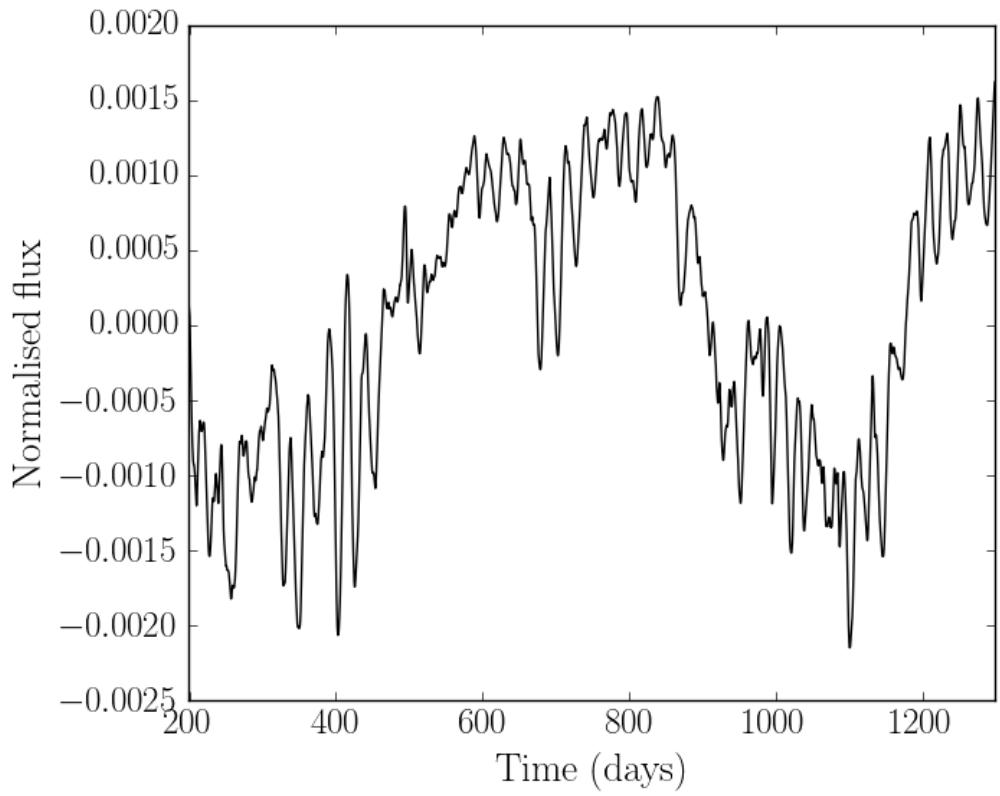


Figure 4-2: An example simulated, noise-free light curve. This ‘star’ has a rotation period of 17.4 days.

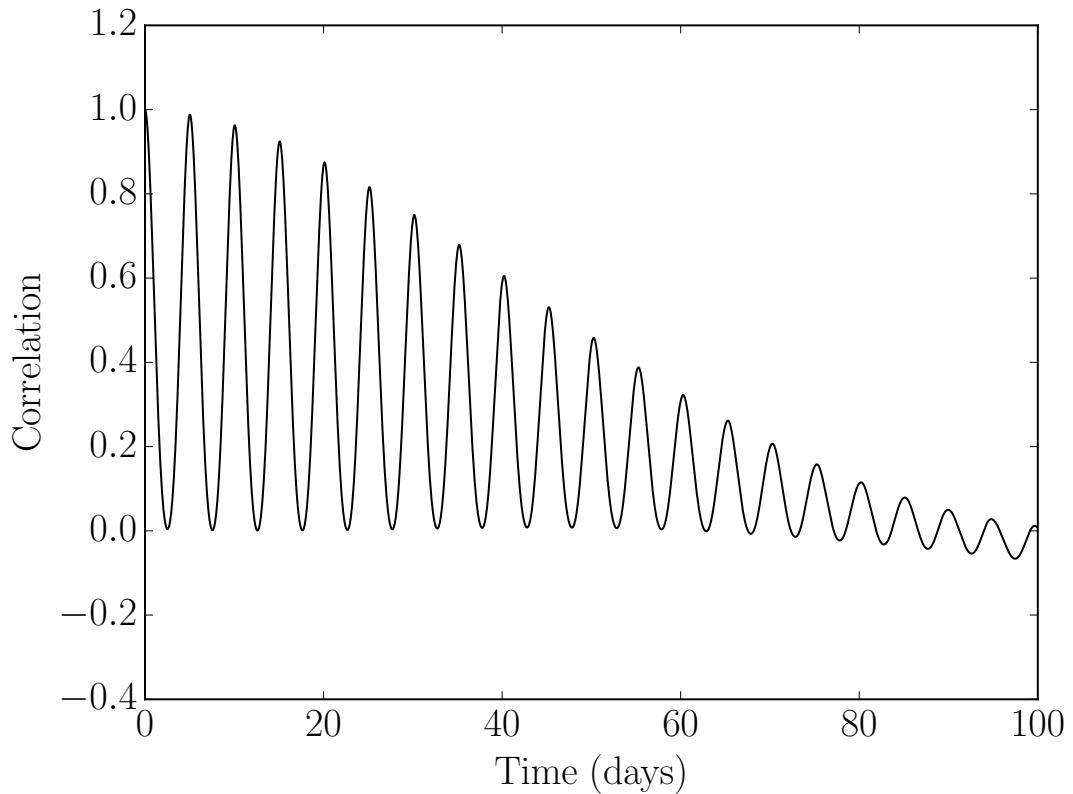


Figure 4-3: The autocorrelation function of the simulated, noise-free light curve depicted in figure 4-2.

#### 4.2.1 ACF

We calculated an ACF for each light curve using the method of McQuillan et al. (2013a). In this method, an ACF is calculated for each light curve and smoothed by convolving with a Gaussian with  $\sigma = 9$  days. A rotation period is estimated as the lag-time of the first peak in the ACF, unless the second peak is larger in which case *that* lag-time is interpreted as the true period. The second peak in an ACF can be larger than the first if there are two active regions at or near opposite longitudes on the surface of the star—this would produce a light curve with two dips per rotation period. An example ACF of the light curve in figure 4-2 is shown in figure 4-3.

The ACF method has proven to be extremely useful for measuring rotation periods. The catalogue of rotation periods of *Kepler* stars provided in McQuillan et al.

(2013a) has been widely used by the community and has provided ground-breaking results for stellar and exoplanetary science. The results of the ACF method as tested in Aigrain et al. (2015a) were positive (see, for example their figure 8) as it produced a large number of accurate rotation period measurements. Another advantage of the ACF method is that it is conveniently fast to implement.

We applied the ACF method to our sample of 300 noise-free, simulated light curves. Periods measured using the ACF method are plotted against the original rotation period values used to generate the light curves in figure 4-7. A noteworthy feature of this figure is that many of the points fall below the  $x = y$  line, *i.e.* the recovered rotation periods are a little shorter than the true periods. We believe this is a feature of the peak position measurement method that is performed on the ACF. ACFs of stellar light curves can be roughly described as a cosine function superimposed on top of a decaying exponential. In such a function the peak positions can be shifted towards the left—the short period end—because the decaying exponential raises the left side of each peak more than the right. It is possible to model this effect of course, however the standard practice is to simply measure the peak position without taking it into account. We are still investigating the origin and implications of this effect.

#### 4.2.2 The sine-fitting periodogram method

For each simulated light curve, a Lomb-Scargle (LS) periodogram (Lomb, 1976; Scargle, 1982) was computed over a grid of 10,000 periods, evenly spaced in frequency, between 1 and 100 days. The period of the highest peak in the periodogram was adopted as the rotation period. The resulting recovered rotation periods are plotted as a function of the true periods in figure 4-5. These recovered rotation periods are in general more accurate than the ACF results: they do not systematically over or under predict rotation period. They are however less precise than the ACF results—the 8.03 day RMS of residuals is five times larger.

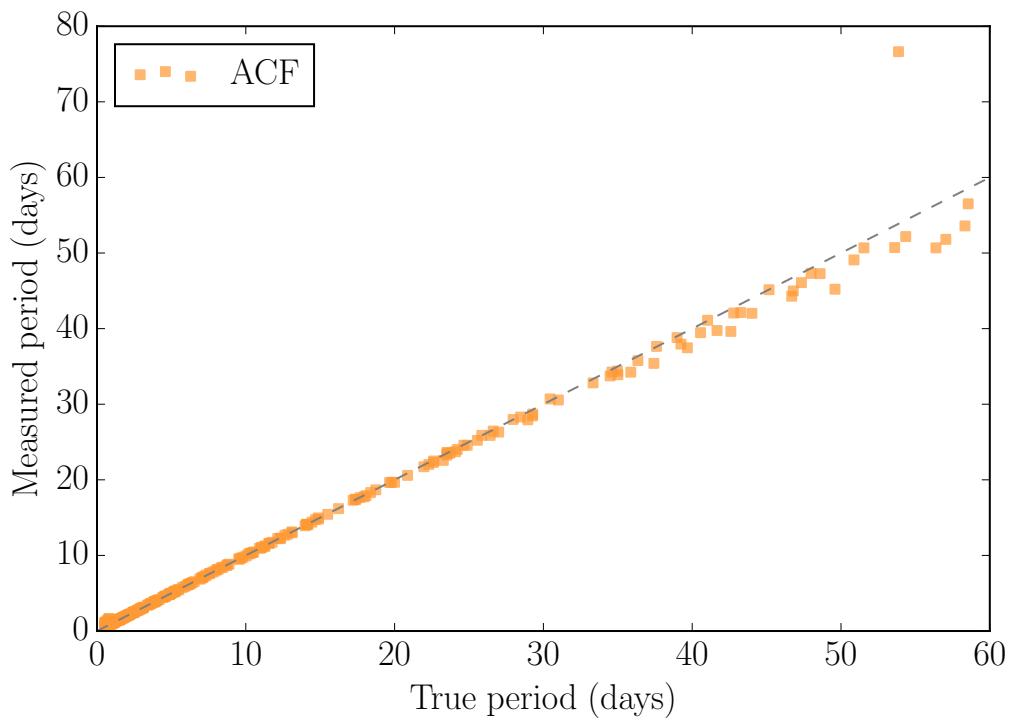


Figure 4-4: Rotation periods measured using the ACF method as a function of the ‘true’ period used to generate 300 simulated, noise-free light curves.

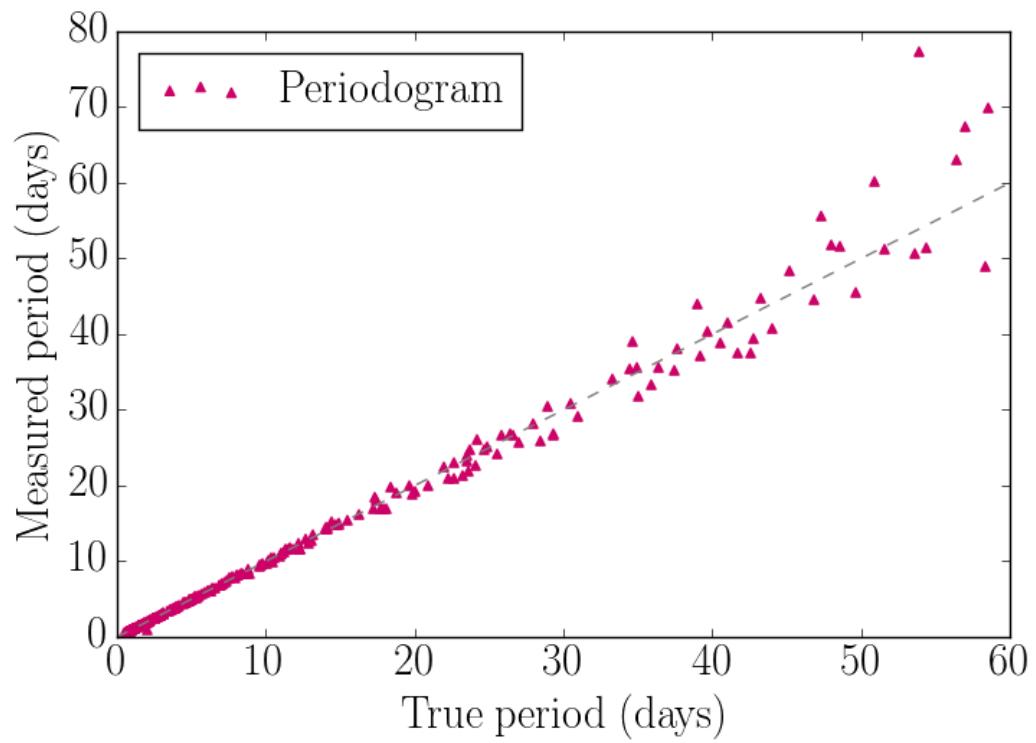


Figure 4-5: Rotation periods measured using the LS periodogram method as a function of the ‘true’ rotation period used to generate 300 simulated light curves.

### 4.2.3 The GP method

In order to recover the rotation periods of the simulated light curves using Gaussian processes, we sampled the posterior PDFs of the kernel hyperparameters described in equation 4.1. The likelihood function for a GP is similar to the simple Gaussian likelihood function that is used for optimisation problems where the uncertainties are Gaussian and uncorrelated. The latter can be written

$$\ln \mathcal{L} = -\frac{1}{2} \sum_{n=1}^N \frac{(y_n - \mu)^2}{\sigma_n^2} - \frac{N}{2} \ln(2\pi\sigma_n^2), \quad (4.3)$$

where  $y_n$  are the data,  $\mu$  is the mean model and  $\sigma_n$  are the Gaussian uncertainites on the data. The equivalent equation in matrix notation is

$$\ln \mathcal{L} = -\frac{1}{2} \mathbf{r}^T \mathbf{C}^{-1} \mathbf{r} - \ln |\mathbf{C}| + \text{constant}, \quad (4.4)$$

where  $\mathbf{r}$  is the vector of residuals and  $\mathbf{C}$  is the covariance matrix,

$$\mathbf{C} = \begin{pmatrix} \sigma_1^2 & \sigma_{2,1} & \cdots & \sigma_{N,1} \\ \sigma_{1,2} & \sigma_2^2 & \cdots & \sigma_{N,2} \\ \vdots & & & \\ \sigma_{1,N} & \sigma_{2,N} & \cdots & \sigma_N^2 \end{pmatrix} \quad (4.5)$$

In the case where the uncertainties are uncorrelated, the noise is ‘white’, (which is a frequent assumption made by astronomers and is sometimes justified) and the off-diagonal elements of the covariance matrix are zero. However, in the case where there is evidence for correlated ‘noise’<sup>2</sup>, as in the case of *Kepler* light curves, those off-diagonal elements are non-zero. With GP regression, a covariance matrix generated by the kernel function,  $\mathbf{K}$  replaces  $\mathbf{C}$  in the above equation. Incidentally, this approach is the reverse of the regression techniques usually employed by astronomers. In most problems in astronomy one tries to infer the parameters that describe the mean model and, if correlated noise is present, to marginalise over that noise. Here, the parameters

---

<sup>2</sup>In our case the ‘noise’ is actually the model!

describing the correlated noise are what we are interested in and our mean model is simply a straight line at  $y = 0$ .

One could either maximise this likelihood function in order to find the best-fit hyper-parameters of the covariance kernel function or, as in our approach, sample from the posterior PDFs of the hyper-parameters. The advantage of the maximum-likelihood method is that the best-fit parameters will be found much faster, but the uncertainties on the rotation period will not be constrained. In order to measure accurate uncertainties, the posterior PDFs of the parameters must be explored. Since obtaining accurate uncertainties on rotation periods is one of our main motivations for this method, we use MCMC despite the added computational expense.

We use the ACF period as an initial guess for the rotation period (this decision is discussed further in section §5.3). We use a uniform prior over rotation period with bounds described below and assert that the covariance decay timescale parameter,  $l$  must be greater than the rotation period. This represents our assumption that the evolution timescale of active regions is greater than stellar rotation periods. This assumption may be more valid for late spectral types—hot stars have smaller active regions which are likely to be shorter-lived. For the remaining hyperparameters we use the following initial values and log-uniform prior distributions:

$$\begin{aligned} A_{initial} &= e^{-5}, \sim \exp(U[-20 : 20]) \\ l_{initial} &= e^7, \sim \exp(U[-20 : 20]), l < P \\ \Gamma_{initial} &= e^{0.6}, \sim \exp(U[-20 : 20]) \\ \sigma_{initial} &= e^{-16}, \sim \exp(U[-20 : 20]) \\ P_{initial} &= P_{ACF}, \sim U([1 - 0.4]P_{ACF} < P_{ACF} < [1 + 0.4]P_{ACF}). \end{aligned} \tag{4.6}$$

$\sigma$  is an additional white noise term added to the diagonal elements of the covariance matrix. It is the fraction by which the observational uncertainties have been underestimated. If the errorbars reported on the data are too small, this parameter will be non-zero. In practice this parameter should always be non-zero when performing GP regression. This is because the covariance matrix must be positive definite, however

matrix inversion performed using most solvers is approximate, not exact, therefore slight deviations from positive definitism can arise. Including a small amount of extra variance in the model allows enough flexibility that the matrix inversion algorithms do not run into numerical issues.

We subsampled the light curves in order to reduce computation time. The subsampling is altered for each light curve and depends, again, on the ACF period estimate. We found that retaining only 20 data points per rotation period, based on the ACF period estimate, significantly reduced computation time whilst maintaining performance. The `george` (Foreman-Mackey et al., 2014b) python package was used to implement the GP model which uses the fast matrix solver, HODLR (Ambikasaran et al., 2014). Matrix operations performed with HODLR scale as  $N \log^2(N)$ , where  $N$  is the number of data points. We used `emcee` (Foreman-Mackey et al., 2013), an affine invariant ensemble sampler to explore the posterior PDFs of the model parameters. The rotation period was taken as the median value of the posterior PDF. The resulting measured rotation periods are compared to the true rotation periods in figure 4-6. The root-mean-square (RMS) of residuals for the GP results is 0.26 days. The results of all three methods are plotted on the same axes in figure 4-7.

The marginal posterior distributions of the QP kernel hyperparameters for an example noise-free light curve with a 14.4 day rotation period, are shown in figure 4-9. The light curve itself with the best fit GP model is shown in figure 4-8.

#### 4.2.4 Real *Kepler* data

The noise-free light curves were injected into real *Kepler* light curves in order to test the performance of the GP method in a realistic case. Ideally, simulated light curves would have been injected into the *Kepler* pixel-level data, then the same detrending method that is applied to the *Kepler* light curves would be applied to these light curves before attempting to measure rotation periods for these stars. However, since this work is designed to be more of a demonstration of the overall efficacy of the GP periodogram, rather than the development of a *Kepler* specific method, this is beyond the scope of our demonstrations.

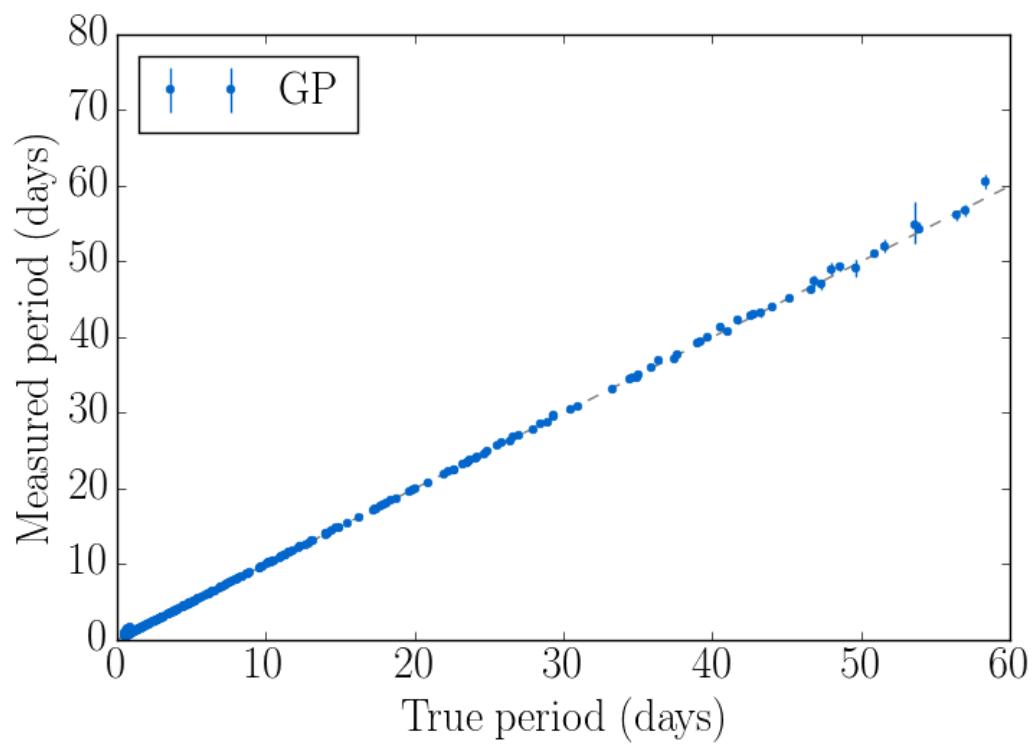


Figure 4-6: Measured vs true rotation period for 300 simulated light curves using the GP method.

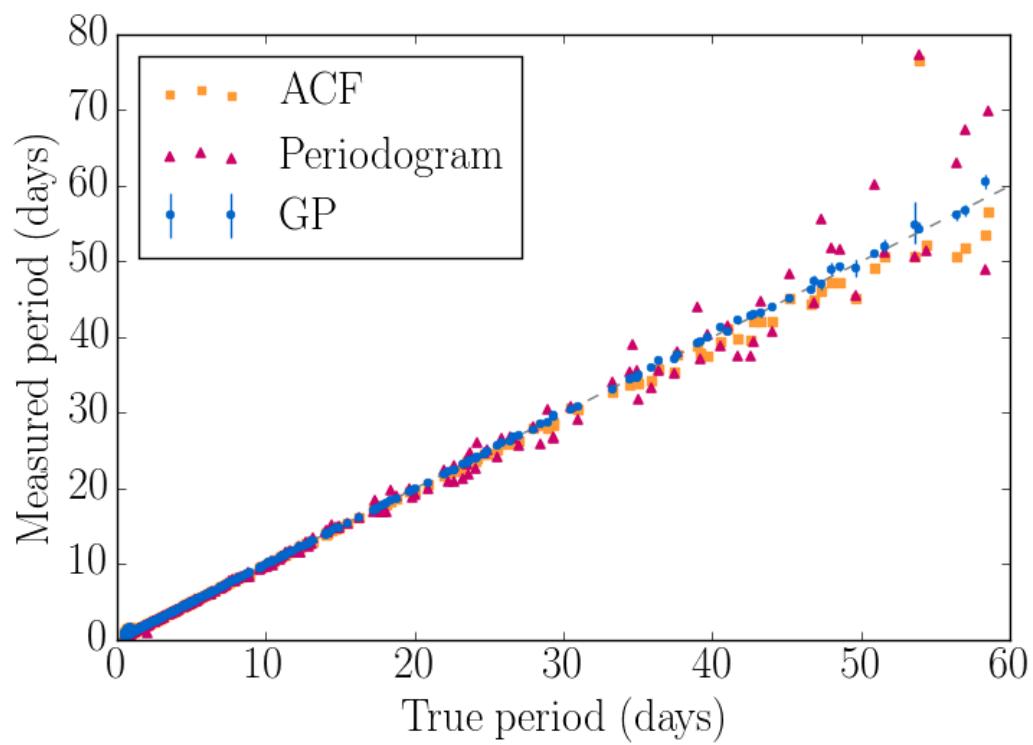


Figure 4-7: Measured vs true rotation period for 300 simulated light curves from the three different methods described in the text.

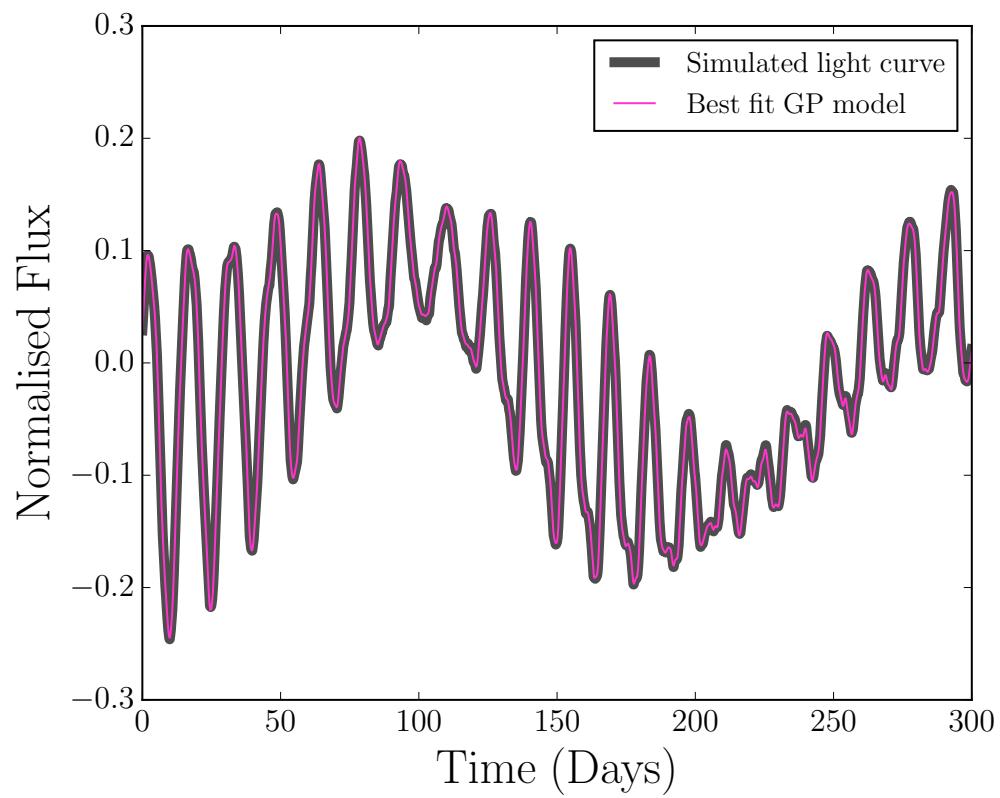


Figure 4-8: A simulated light curve with a rotation period of 14.4 days. The pink curve shows the GP model with the best-fit parameters. The marginal posteriors of the GP hyper-parameters used to produce this fit are shown in figure 4-9.

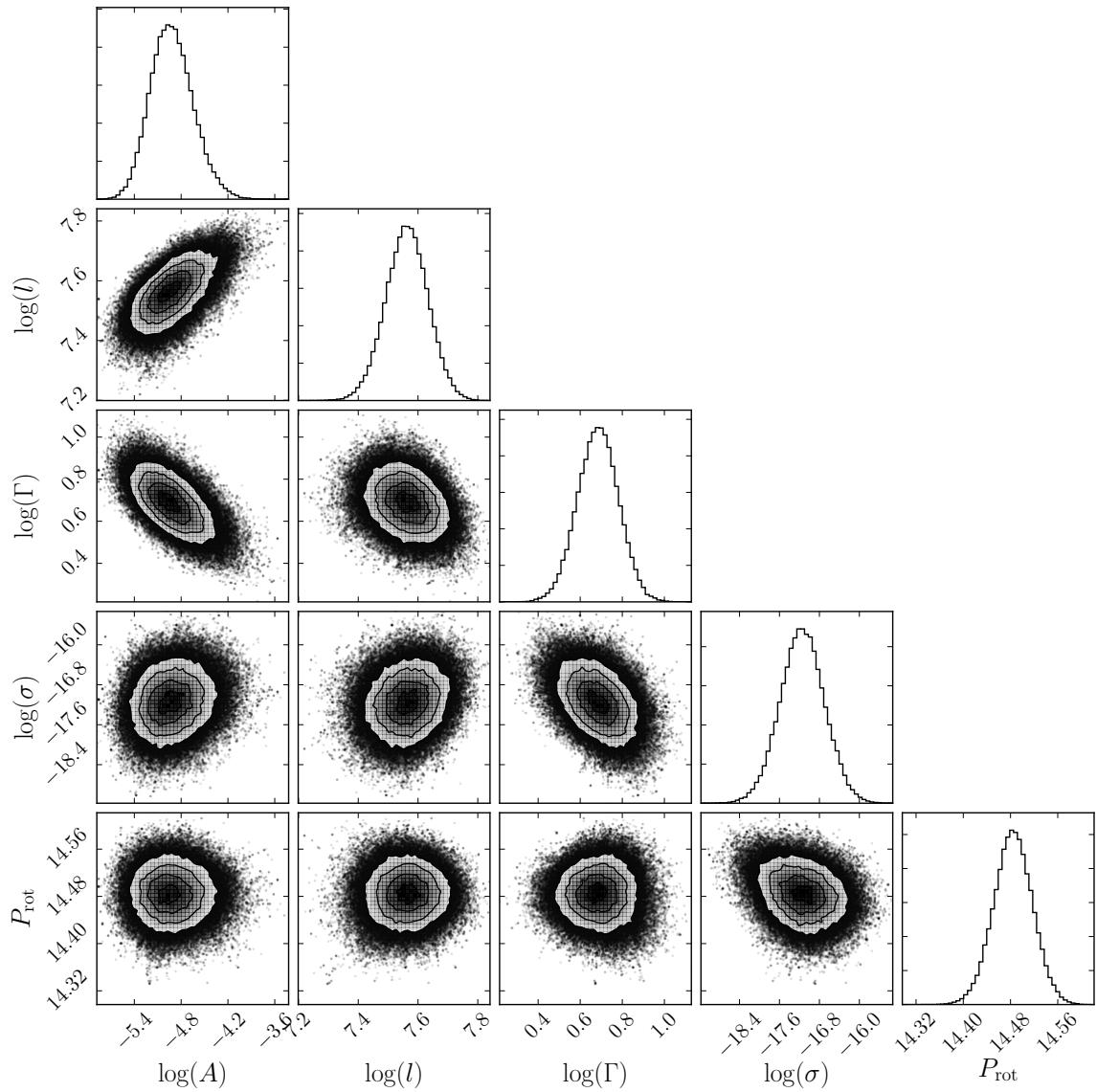


Figure 4-9: Marginal posterior PDFs of the QP GP model parameters.  $\sigma$  is an additional white noise term added to the diagonal elements of the covariance matrix. It is the fraction by which the observational uncertainties have been underestimated. If the errorbars reported on the data are too small, this parameter will be non-zero. Since this light curve was simulated without noise, this parameter is very close to zero. The true rotation period of this star is 14.4 days.

Unfortunately, automating the GP method so that it can run on large numbers of real *Kepler* light curves is difficult. The added noise reduces the acceptance fraction within the MCMC chains, significantly increasing the required time for convergence. Despite this, the method works well on individual targets where chains can be run for a long time and subsampling can be reduced. We used a simulated, noisy light curve from Aigrain et al. (2015a) that was generated by injecting a noise-free simulation into a real *Kepler* light curve in order to preserve realistic *Kepler* noise properties.

*Kepler* light curves are big data and extracting the maximum amount of information from them requires either large numbers of CPU hours, or using work-arounds. One feature of *Kepler* light curves that works to our advantage is that they are naturally broken up into smaller time units by quarterly (three month) breaks. Splitting the data set into quarters, rather than modelling the entire time series contiguously speeds up the computation time as inverting several small matrices is faster than inverting one large one. The *Kepler* quarter divisions are natural places to split the time series because the spacecraft rotates by ninety degrees every quarter (three months), placing each star on a new CCD module. Pixel response functions and background flux differs from pixel-to-pixel and module-to-module, so noise properties of *Kepler* light curves change every quarter. Additionally, changes in the spacecraft’s orientation and position during quarterly re-pointings temporarily affect the temperature of the CCD, producing systematic features in the light curves at the start of some quarters. We model each quarter separately but the parameters of the GP kernel function are global, *i.e.* we do not use a separate period parameter for each quarter—there is just one period parameter for an entire light curve. It would be possible to model the time series with a mixture of some global parameters and some quarter-specific parameters, for example one might expect that the amplitude of covariance,  $A$  or white noise level,  $\sigma$  to vary on a quarterly basis. However, since there are seventeen quarters this would lead to thirty-seven parameters, and in the interest of minimising computation time (adding parameters leads to longer MCMC burn in and convergence time), we choose to use global parameters only. Unfortunately, the application of this method to noisy test cases is still under development and cannot provide re-

sults here. We hope to develop this method further and apply it to real light curves in the near future.

## 4.3 Discussion

The main drawback of the GP method is computation time. Because GPs are so expensive to compute, it is necessary to come up with shortcuts in order to perform inference on *Kepler* light curves, each of which compromises accuracy. The shortcuts that we employ here include:

- Initialisation. The choice of initial parameter values makes a difference to computation time: the closer they are to the maximum likelihood parameters the shorter the MCMC burn-in time. Using the ACF period to initialise is not ideal as results will not be fully independent of the ACF period unless MCMC chains are run for an infinitely long time.
- Subsampling. Not only is some information always lost by subsampling, but the choice of subsampling frequency will also effect the results. We chose to retain only 20 data points per period — based on the ACF period — which, again means that our results are not independent of the ACF results.
- Priors. We chose to use uniform, bounded priors to reduce parameter space and therefore convergence time. Unfortunately, if the true period lies outside the bounded region, it will never be found.

These shortcuts are only necessary when working with a large number of stars (of order hundreds or thousands). If only interested in single or small numbers of targets it may be possible to reduce the extremity of these shortcuts or to avoid them altogether.

### 4.3.1 The ACF method

The main advantage of the ACF method over the GP method is speed: it is extremely fast and therefore useful to get a quick estimate of a period. We have demonstrated

that the ACF method may produce rotation periods that are slightly systematically biased towards lower rotation periods. This effect comes from the process of extracting the period from the ACF itself: typically superimposed onto a decaying exponential, that first peak in the ACF gets shifted towards faster periods. One may be able to get around this by modelling the ACF as a sum of a cosine and exponential function (a similar function to GP kernel function) but in practice ACFs tend to have a more complicated structure and it is not possible to model them with a simple function. This observation leads to the idea of modelling the ACF with a Gaussian processes, which then logically progresses to a using a Gaussian process with a Gaussian process kernel function. Although this may be an attractive idea, in practice a GP would not necessarily produce a positive semi-definite covariance kernel<sup>3</sup>. An alternative approach may be to correct for rotation period bias after the peak position has been measured by inferring a correction factor. Yet another approach could be to measure the mean separation between adjacent peaks in the ACF: the first peak should be most strongly affected by the exponential decay, the second peak less so, and so on.

### 4.3.2 Initialisation

Using the ACF to initialise the MCMC chains is not ideal because, of course, one becomes reliant on the assumption that the ACF period is close to the true period. Of course, if you had infinite CPU time, this would not be a problem as you would eventually sample the entire posterior PDF of the period parameter, however in practice this is likely to be an issue. The only way to get around this problem is to run the MCMC chains for as long as possible, or, alternatively to use a sampler that is designed to move around the parameter space much more quickly than `emcee`; nested sampling for example. We chose to use the ACF periods rather than the periodogram periods to initialise as, although there was some systematic bias present in these results, that bias was small and there were fewer large outliers.

---

<sup>3</sup>One that can be inverted; a necessary regression operation.

### 4.3.3 Kernel function choice and interpretation of hyper-parameters

The QP kernel function represents a simplistic effective model of a stellar light curve. It adequately describes the data, captures that all-important periodic quality and is relatively simple, with only a few hyper-parameters. It also satisfies the requirement to produce positive semi-definite covariance matrices. Whilst the QP kernel function evidently captures the periodic qualities of light curves adequately, it is still a somewhat arbitrary choice. Another valid choice would be a squared cosine function multiplied by a squared exponential,

$$k_{i,j} = A \exp\left(-\frac{(x_i - x_j)^2}{2l^2}\right) \cos^2\left(\frac{2\pi}{P}\right) \quad (4.7)$$

This function produces a positive semi-definite matrix and has the  $P$  parameter of interest. It may in fact be even *more* suited to modelling stellar light curves as it describes a Gaussian in frequency space. It is easy to imagine a differentially rotating star with a period that is a Gaussian in frequency space: the mean frequency would be the frequency at the most active latitude, at or near which spots spend the majority of their time and the tails would be occupied by spots that drift near the equator or poles. The main difference between this cosine and the QP function is that the cosine function allows negative covariances and the QP function does not. Is it realistic to allow negative covariances? In practice, the ACFs of *Kepler* light curves often go negative. However, many stars have two active regions on opposite hemispheres that produce two brightness dips per rotation. If the covariance is forced to be negative for two data points that are separated by half a rotation period, those light curves with two peaks per rotation period may not be well modelled. It would be very worthwhile to test this assumption and this alternative kernel function in future. If CPU time were not limited there may be some benefit to performing formal model comparison with different kernel functions. However, the evidence integral is an ambitious calculation for models with likelihood functions that take milliseconds to compute, let alone those involving GPs with light curves containing thousands of data points which can take minutes.

Clearly, stellar rotation periods are well represented by the  $P$  parameter in our QP kernel function, as evidenced by its impressive ability to recover the true rotation periods from the simulated light curves in this work. However, it is not clear whether the QP kernel function is the *best* function to use. There may be an alternative function which is better suited to capturing stellar variability and is able to recover periods even more precisely than the QP kernel. There may also be an alternative function which is more physically motivated, that captures not just the rotation period but also (for example) the spot lifetime or differential surface rotation. This is beyond the scope of this work but we hope that these questions will be answered in the near future.

#### 4.3.4 Future work

The next stages of improving this method are listed as follows:

- To optimise the tradeoff between computational efficiency and accuracy. We have performed limited tests to explore the optimum subsampling strategy and prior bounds. These should be explored more thoroughly in future. For example, it would be useful to know how the precision and accuracy of the recovered rotation period varies as a function of subsampling frequency
- To perform model selection with different kernel functions. Again, these tests have not been performed due to the cost of calculating the fully marginalised likelihood with GPs. This calculation may be prohibitively expensive, however simpler model selection tests can be performed. For example, the relative precision of rotation periods recovered using alternative kernel functions could be tested.
- To design and implement a physically motivated kernel function. We have only explored the physical interpretation of one parameter in our kernel function,  $P$ . However, the other parameters may also be related to some physical processes. For example, the overall timescale for covariance fall-off,  $l$  may be related to spot

lifetimes. A star with long spot lifetimes will show little variation in the overall shape and amplitude of its light curve between rotations and  $l$  will be large. In contrast, the light curve of a star with short spot lifetimes may display non-repeating patterns and amplitudes that vary rapidly between rotations. In this case  $l$  will be small. The  $\Gamma$  parameter is related to the number of zero crossings within one rotation period: when  $\Gamma$  is small there are many zero crossings and vice versa. Since the number of zero crossings per rotation period is related to the number of active regions on the surface of the star, this parameter may also be of physical interest. In addition, instead of interpreting the parameters of the QP kernel function used here, it may be possible to design an entirely new kernel function, based on the physical processes that drive the light curve variability. This idea is being explored by another member of my research group.

- To develop a detection criterion to assess whether a rotation period was measured. Another general problem in rotation period inference is deciding whether a real rotation period was measured at all. Detection thresholds are usually set using the amplitudes of peaks in LS periodograms or ACFs in order to eliminate contaminants. Using a GP model it may be possible to do this via model selection: *i.e.* are the data better described by a periodic model, or a non-periodic model? There is more than one way to implement such a model comparison, one way could be to use a sum of two kernel functions: one periodic and one non-periodic, each with its own amplitude parameter. If there is periodicity in the data, the amplitude of the periodic kernel will be larger than the amplitude of the non-periodic kernel and vice versa.
- To build in a noise model for *Kepler* data. Another huge advantage of the GP method is, because it is a *generative* model of the data, the rotation period signal can be modelled at the same time as systematic noise features. One can then marginalise over the parameters of the noise model. This approach would be extremely advantageous for *Kepler* data since long-term trends are often removed by the *Kepler* detrending pipeline. Marginalising over the noise

model at the same time as inferring the parameters of interest will insure that the periodic signal is preserved.

## 4.4 Conclusions

We simulated 300 noise-free, *Kepler*-like light curves using a spot model and attempted to recover the rotation periods used to generate them. Three methods were compared: a LS periodogram method, an ACF method and our GP method. The GP method produced the most precise and accurate rotation periods of the three techniques. This is because a GP is a semi-parametric model that is well suited to signals that are non-sinusoidal and quasi-periodic, because it does not require the assumption of regular time sampling and because a full posterior PDF of the rotation period parameter is produced. Not only this, unlike the other two methods, the GP method provides accurate rotation period uncertainties.

The main drawback of this method is the computation time. We demonstrated that the GP method is capable of recovering signals from noise-free light curves, but were unable to automate the GP method for light curves with realistic noise properties. It may be that this method must remain a ‘boutique’ method: ideal for small numbers of targets but not well suited to the entire *Kepler* catalogue. We continue to develop this method and remain hopeful that we can, eventually infer a probabilistic rotation period for thousands of *Kepler* stars.

# Chapter 5

## Probabilistic Inference of basic stellar parameters: application to flickering stars

### Abstract

The relations between observable stellar parameters are usually assumed to be deterministic. That is, given an infinitely precise measurement of independent variable, ‘ $x$ ’, and some model, the value of dependent variable, ‘ $y$ ’ can be known exactly. In practice this assumption is rarely valid and intrinsic stochasticity means that two stars with exactly the same ‘ $x$ ’, will have slightly different ‘ $y$ ’s. The relation between short-timescale brightness fluctuations (flicker) of stars and both surface gravity (Bastien et al., 2013) and stellar density (Kipping et al., 2014) are two such stochastic relations that have, until now, been treated as deterministic ones. We recalibrate these relations in a probabilistic framework, using Hierarchical Bayesian Modelling (HBM) to constrain the intrinsic scatter in the relations. We find evidence for additional scatter in the relationships, that cannot be accounted for by the observational uncertainties alone. The scatter in surface gravity and stellar density does not depend on flicker, suggesting that using flicker as a proxy for  $\log g$  and  $\rho_*$  is equally valid for dwarf and giant stars, despite the fact that the observational uncertainties tend to be larger for dwarfs.

## 5.1 Introduction

Accurate stellar characterization plays a vital role for many active research fields within astronomy. For example, stellar populations, galactic archaeology, the study of binary stars, asteroseismology and exoplanet studies all rely on inferences of basic stellar parameters to varying degrees. Empirically-derived and reliable estimates are of particular value, increasing our confidence in the end-product results built upon these inputs.

Basic stellar parameters, such as effective temperature and surface gravity, can be inferred using one (or more) of several types of observations, such as spectroscopy, photometry, interferometry, etc. This inference can be performed by invoking theoretical models or by building an empirical calibration library. For example, an observed stellar spectrum could be matched against a library of theoretical spectra generated using stellar atmosphere models, or, against a library of observed spectra of “standard stars”, serving as calibrators. Regardless of the approach, be it theoretical or empirical, the methods used for the inference of stellar parameters are traditionally “deterministic”. In this context, a deterministic model can be loosely described as one where a particular observational input always returns a single-valued output for a parameter of interest, i.e. nature itself has no variance and the underlying model is considered to be a perfect description of reality.

An alternative approach for inferring model parameters is to allow relationships between observables to be stochastic. In recent years, there has been a shift towards such methods in several areas of astronomy, particularly within the exoplanet community. For example, Wolfgang et al. (2015) considered that the mass-radius relationship of exoplanets is stochastic, since a particular sized planet could have a range of planet masses due to unmodeled variances in compositions, environment and other complications. These recent demonstrations in exoplanetary science have prompted us to consider the need for treating the parent stars in the same probabilistic framework, with potential applications spanning many fields of astronomy.

The demand for probabilistic stellar parameters is not only motivated by the

fact that probability distributions are far more representative of our ‘beliefs’ about astrophysical parameters, it also has a practical purpose. When using data published in the astronomical literature to, for example, infer relationships between parameters that are themselves the product of an inference process (for example, exoplanet transit depth and period), inference can be performed as the final stage in a hierarchical treatment (see, e.g. Foreman-Mackey et al., 2014a). Studies such as these benefit from posterior PDF samples, rather than point estimates of inferred properties.

One of the more recent tools developed to characterize stars is known as “flicker” (Bastien et al., 2013). Flicker is a measure of the scatter on an 8-hour timescale (denoted as  $F_8$ ) in a broad visible bandpass time series photometric light curve, such as that from *Kepler* or the upcoming TESS mission.

Short-term brightness fluctuations of *Kepler* stars that are traced by flicker are produced by a combination of acoustic, p-mode oscillations and granulation. Convective material wells up at the stellar photosphere; rising material is hot and bright and sinking material is cooler and darker. Hot plasma rises at the centres of convective cells, ‘granules’, and cool material sinks at the edges, as shown in figure 5-1. The granules evolve over time, producing fluctuations in the disk-integrated stellar luminosity which is evident in *Kepler* light curves. The characteristic timescales and amplitudes of these fluctuations depend on fundamental stellar parameters (e.g. Samadi et al., 2013a,b). In fact, Kallinger and Matthews (2010) and Mathur et al. (2011) demonstrate that the characteristic granulation timescale scales with the inverse of the peak p-mode frequency,  $\nu_{max}$ . Similarly, the root-mean-square (RMS) of brightness fluctuations scale as  $\nu_{max}^{-1/2}$  (Chaplin et al., 2011; Mathur et al., 2011). Since  $\nu_{max}$  is proportional to surface gravity<sup>1</sup>, it follows that granulation-induced short-term brightness variations are related to  $\log g$ . Bastien et al. (2013) set out to show, not only that this was the case, but that granulation was detectable in *Kepler* long cadence data, thus allowing  $\log g$  to be inferred for thousands of targets lacking short cadence data. Bastien et al. (2013) developed a statistic of *Kepler* long cadence light curves that traces the amplitude of brightness variations produced by granulation:

---

<sup>1</sup> $\nu_{max}$  is related to the acoustic cut-off frequency which scales as  $\frac{g}{T_{\text{eff}}^{1/2}}$ .

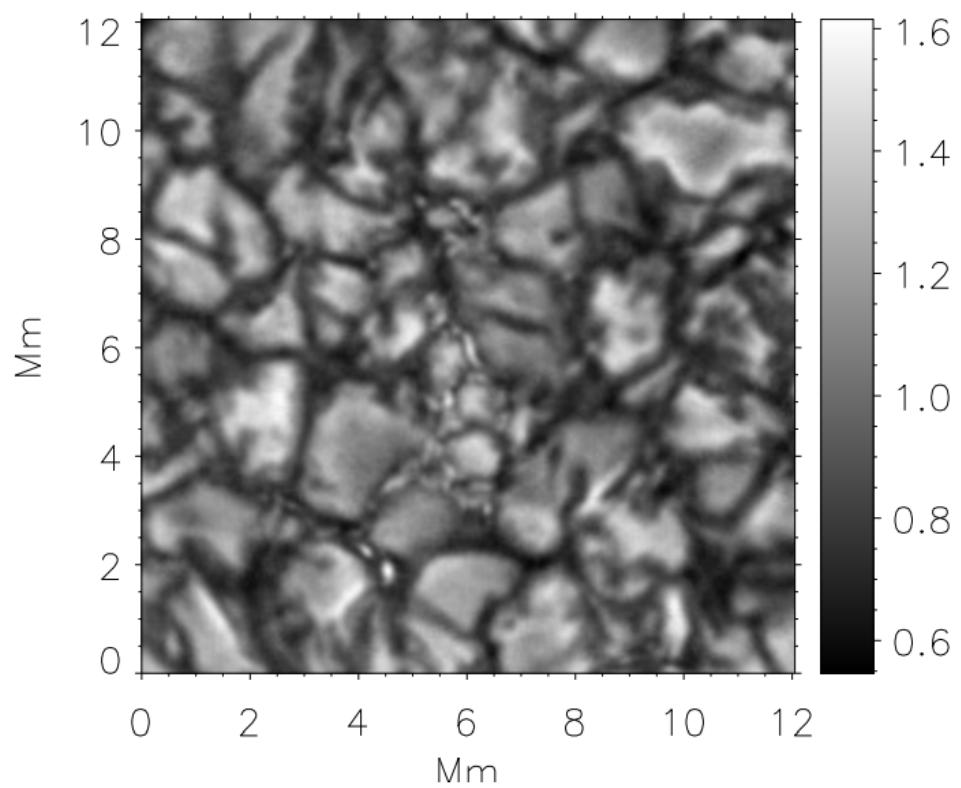


Figure 5-1: An image of the Solar photosphere showing granulation cells. Figure from Cegla et al. (2013).

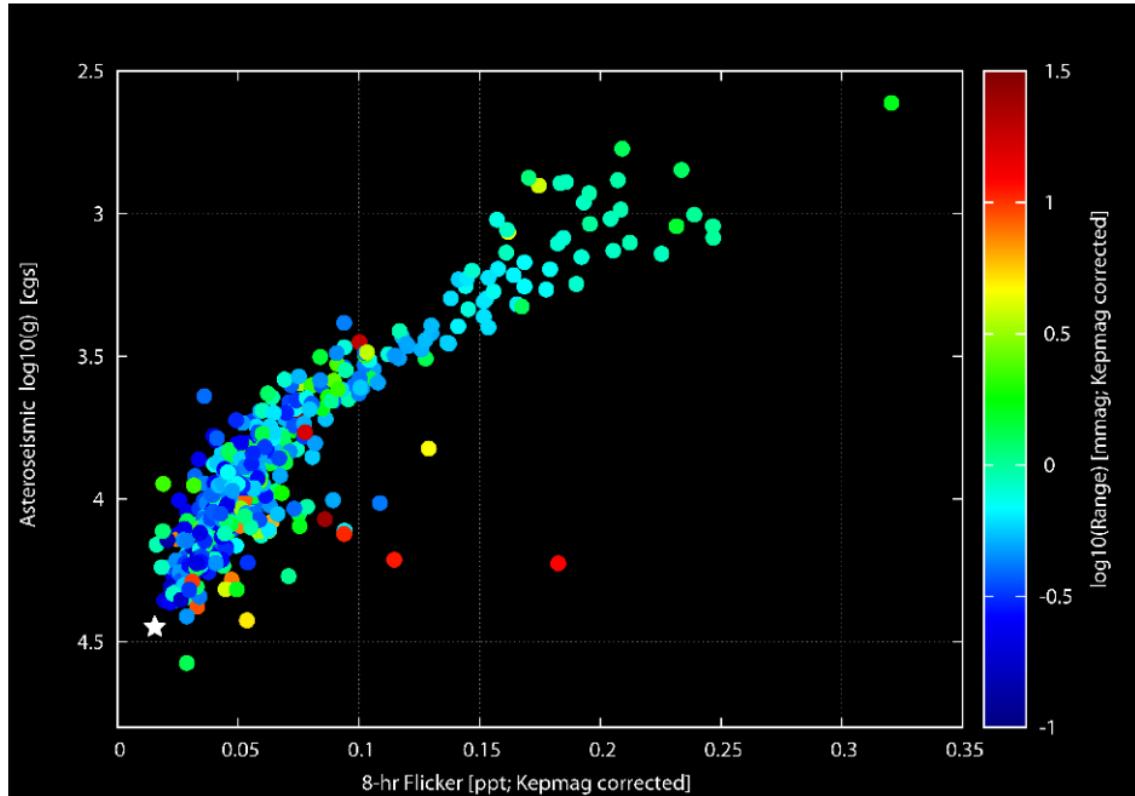


Figure 5-2: Figure 2 from Bastien et al. (2013). This figure demonstrates the impressive correlation between  $\log g$  (from asteroseismic analysis of *Kepler* light curves) and flicker.

the RMS on time scales shorter than 8 hours, known as ‘8-hour flicker’, or  $F_8$ . As shown in Bastien et al. (2013), flicker displays a remarkable correlation to the asteroseismically determined surface gravities ( $\log g$ ) — see figure 5-2. Conversely, the observation implies that flicker can be used to empirically infer surface gravities at the level of  $\sim 0.1$  dex, an attractive proposition given the wealth of photometric light curves available through the array of exoplanet transit missions flying and scheduled to launch.

Cranmer et al. (2014) demonstrated that models of stellar surface granulation indeed reproduce a flicker effect in close agreement with that observed by Bastien et al. (2013), providing a physically-plausible explanation. Since surface gravity is highly correlated with mean stellar density ( $\rho_*$ ) on evolutionary tracks, Kipping et al. (2014) showed that flicker can be also be used to infer  $\rho_*$ , which is more useful for

exoplanet transit analysis (Seager and Mallén-Ornelas, 2003).

Whether one calibrates flicker to  $\log g$  or  $\rho_*$ , there are several aspects of the problem which are attractive for our purposes of a simple demonstration of probabilistic inference of stellar parameters. Firstly, in log-log space the relationship is very simple, appearing to be linear (Kipping et al., 2014). Secondly, there is a sufficiently large number of points in the sample (439 stars) to constrain a population-based model. Thirdly, there is significant excess scatter around the best-fitting relation implying that a deterministic model is inadequate. This is not surprising given that granulation is a complex and messy process for which one should not expect any parametric model to provide a perfect description. Finally, the physical processes that produce surface granulation, of which flicker is an observational tracer, may be more or less noisy for different types of stars. We will test whether flicker has greater predictive power in certain regions of parameter space; i.e. is flicker significantly more informative for subgiants than for dwarfs? For these reasons, we identify the calibration of flicker to  $\log g$  and  $\rho_*$  as a well-posed problem to first demonstrate probabilistic inference in the arena of stellar characterization.

## 5.2 Probabilistic calibration

### 5.2.1 Calibration Data

For our calibration data, we used a sample of *Kepler* stars with both asteroseismic and flicker measurements available. Chaplin et al. (2014) report asteroseismic  $\rho_*$  estimates (and the associated uncertainties) for 518 *Kepler* stars. The authors report three different sets of results, depending on the choice of  $T_{\text{eff}}$  and [Fe/H], and in this work we elected to use values reported in their Table 6 over Table 5, and Table 5 over Table 4. Table 6 contains a subset those stars that appear in Tables 4 and 5 with more precise, spectroscopic parameters. Tables 4 and 5 report asteroseismic values based on two different sets of effective temperature estimates, from two different methods. We additionally used the 71 additional planet hosting stars with asteroseismology

reported in Huber et al. (2013) but not reported in Chaplin et al. (2014). Values for flicker and “range” were taken from Kipping et al. (2014), based upon the methods described in Bastien et al. (2013). In order to use the same data set as Kipping et al. (2014) who removed outliers from the sample, we only include targets in our calibration for which:

- Range of variability,  $R_{\text{var}}$  defined in Bastien et al. (2013) to be the difference between the 5th and 95th percentiles of stellar flux  $< 1000 \text{ ppm}$
- $4500 < T_{\text{eff}} < 6500 \text{ K}$
- $K_P < 14$
- $1.2 < \log_{10}(F_8 \text{ [ppm]}) < 2.2$

We use the same sample for our calibration of  $\log g$ , except that we exclude the Huber et al. (2013) data, since these authors do not provide estimates of  $\log g$ <sup>2</sup>.

### 5.2.2 Hierarchical Bayesian Model

We model the stochastic relationship between  $F_8$ ,  $\log g$  and  $\rho_*$ , accounting for the fact that there exists some intrinsic scatter in the dependent variable. There are two excellent reasons for modelling the relation stochastically; firstly, if the intrinsic scatter is ignored and the relation between variables is assumed to be deterministic, those data points with smaller measurement uncertainties may have an unrepresentative greater weighting during the fitting process (Hogg et al., 2010a). Secondly, we are interested in producing probability distributions over stellar densities and surface gravities, as opposed to point estimates, and propagating these probability distributions through to subsequent analyses. Several recent studies have required posterior Probability Distribution Function (PDF) samples, in order to conduct their hierarchical analyses (e.g. Angus et al., 2015b; Foreman-Mackey et al., 2014a; Rogers, 2015).

---

<sup>2</sup>Whilst we could compute  $\log g$  ourselves from the reported masses and radii, this could only be done under the incorrect assumption of zero covariance between  $M_*$  and  $R_*$ .

The two models we use to describe the relationships between  $F_8$ ,  $\log g$  and  $\rho_\star$  are

$$\log_{10}(F_8) \sim \mathcal{N}(\alpha_\rho + \beta_\rho \log_{10}(\rho_\star), \sigma_\rho^2), \quad (5.1)$$

and

$$\log_{10}(F_8) \sim \mathcal{N}(\alpha_g + \beta_g \log_{10}(g), \sigma_g^2). \quad (5.2)$$

The free parameters of the two models are  $\alpha_\rho$ ,  $\beta_\rho$ ,  $\sigma_\rho$ ,  $\alpha_g$ ,  $\beta_g$  and  $\sigma_g$ . These relations are Gaussian distributions with means given by the equation of a straight line and standard deviations which describe the intrinsic scatter about the mean. We used the MCMC package, `emcee` (Foreman-Mackey et al., 2013) to explore the posterior PDFs of our model parameters.

We also tested a model in which the additional scatter depends on flicker itself, defined as

$$\log_{10}(F_8) \sim \mathcal{N}(\mu = \alpha_\rho + \beta_\rho \log_{10}(\rho_\star), \sigma_\rho^2 + \gamma_\rho \log_{10}(F_8)), \quad (5.3)$$

for flicker vs  $\rho_\star$  and similarly for  $\log g$ . This model allowed us to determine whether the magnitude of additional scatter varied as a function of flicker. In other words, whether flicker was a better proxy for  $\log g$  or  $\rho_\star$  for either dwarf or giant stars. We found that the maximum *a-posteriori* values for the  $\gamma$  parameters were consistent with zero:  $\gamma_{rho} = 0.006 \pm 0.02$ ,  $\gamma_g = -0.01 \pm 0.01$ , and interpret this as evidence for a constant intrinsic scatter level across evolutionary stages.

We used a likelihood function which accounts for 2-D uncertainties but does not allow the intrinsic scatter to be a function of the dependent or independent variables. For the relation between flicker and  $\rho_\star$ , this likelihood function can be written as

$$\begin{aligned} \ln [p(F_8 | \rho_\star, \alpha_\rho, \beta_\rho, \sigma_\rho)] &\propto \\ -\frac{1}{2} \sum_{n=1}^N &\left[ \frac{[F_{8n} - (\alpha_\rho + \beta_\rho \rho_{\star n})]^2}{[\beta_\rho \sigma_{F8,n}^2 + \sigma_{\rho*,n}^2 + \sigma_\rho^2]} + \ln(\sigma_{F8,n}^2) + \ln(\sigma_{\rho*,n}^2) + \ln(\sigma_\rho^2) \right] \end{aligned} \quad (5.4)$$

and similarly for  $\log g$ . We found that the posterior PDFs for the model parameters obtained using this likelihood function were consistent with those obtained using a model that only accounts for the uncertainties on the flicker measurements. The median values of the model parameters differed by around  $0.05\sigma$  for the  $\alpha$  and  $\beta$  parameters, by  $0.3\sigma$  for  $\sigma_\rho$  and by  $0.8\sigma$  for  $\sigma_g$ . Since they are so dependent on the observational uncertainties, the parameters that describe the intrinsic scatter in the relations are more sensitive to whether the uncertainties in the  $x$ -direction are included. Accounting for uncertainties on  $y$  and  $x$  is not essential in this case but is still good practice and will, in general, produce more accurate model parameters and uncertainties.

We used the uninformative prior for the parameters of a straight line for data with unknown uncertainties, outlined in VanderPlas (2014),

$$p(\alpha, \beta, \sigma) \propto \frac{1}{\sigma} (1 + \beta^2)^{-3/2}. \quad (5.5)$$

We also tested uniform, flat priors as defined below:

$$\begin{aligned} \alpha, \beta &\sim U(-10 : 10) \\ \log(\sigma_{rho}), \log(\sigma_g) &\sim U(-10 : 10). \end{aligned} \quad (5.6)$$

We found that the results were relatively insensitive to the choice of prior, with median parameter values differing by only around  $0.05\sigma$ . MCMC chains were run until the Gelman & Rubin convergence criterion,  $\hat{R}$  reached a value of less than 1.002 and the number of autocorrelation times<sup>3</sup> was greater than 35. The Gelman & Rubin convergence diagnostic,  $\hat{R}$  (Gelman and Rubin, 1992) is defined as

$$\hat{R} = \frac{\hat{V}}{W} = \frac{m+1}{m} \frac{\hat{\sigma}_+^2}{W} - \frac{n-1}{mn}, \quad (5.7)$$

---

<sup>3</sup>One autocorrelation time is defined as the number of steps taken between independent MCMC samples.

Table 5.1: Median parameter values with  $1\sigma$  uncertainties.

Parameter	Median value
$\alpha_\rho$	$1.31 \pm 0.01$
$\beta_\rho$	$-0.53 \pm 0.01$
$\sigma_\rho$	$0.060 \pm 0.003$
$\alpha_g$	$4.91 \pm 0.05$
$\beta_g$	$-0.83 \pm 0.01$
$\sigma_g$	$0.060 \pm 0.003$

where

$$\hat{\sigma}_+^2 = \frac{n-1}{n}W + \frac{B}{n}. \quad (5.8)$$

$B/n$  is the variance between MCMC chains,

$$\frac{B}{n} = \frac{1}{m-1} \sum_{j=1}^m (\bar{\psi}_{j.} - \bar{\psi}_{..})^2, \quad (5.9)$$

and  $W$  is the variance within each MCMC chain.

$$W = \frac{1}{m(n-1)} \sum_{j=1}^m \sum_{t=1}^n (\psi_{jt} - \bar{\psi}_{j.})^2. \quad (5.10)$$

In the above relations  $n$  is the number of steps,  $m$  is the number of chains,  $\psi$  is a model parameter ( $\alpha$  for example) and  $\bar{\psi}_{..}$  is the mean.

Figures 5-3 and 5-4 show the data with the best-fit models. The shaded regions show the  $1$  and  $2\sigma$  confidence interval which are representative of the intrinsic scatter in the relations. The marginal posterior PDFs of the model parameters for  $\rho_*$  are shown in figure 5-5. The marginal posterior PDFs for  $\log g$  are similarly Gaussian and, as with  $\sigma_\rho$ ,  $\sigma_g$  is clearly greater than zero. We checked the consistency between the two relations by calculating flicker values for the Sun, finding  $F_8 = 1.24 \pm 0.07$  and  $F_8 = 1.21 \pm 0.1$  from Solar density and surface gravity measurements, respectively. All the code used for this project and several ipython notebooks explaining our analysis are available at <https://github.com/RuthAngus/flicker>.

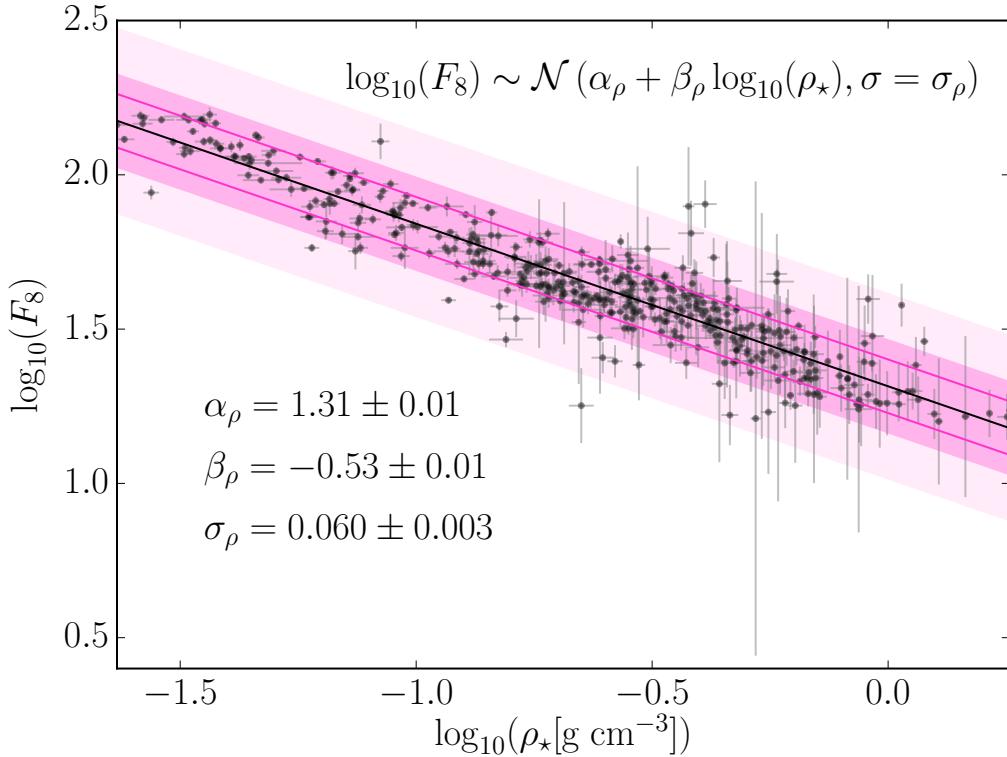


Figure 5-3: Stellar density vs. flicker. This figure shows the model, conditioned on the data. The solid black line shows the model with the best-fitting parameter values quoted in the text. The solid pink lines show the  $1\sigma$  region where the extra scatter is not included and the pink shaded regions show the  $1$  and  $2\sigma$  regions with the additional scatter.

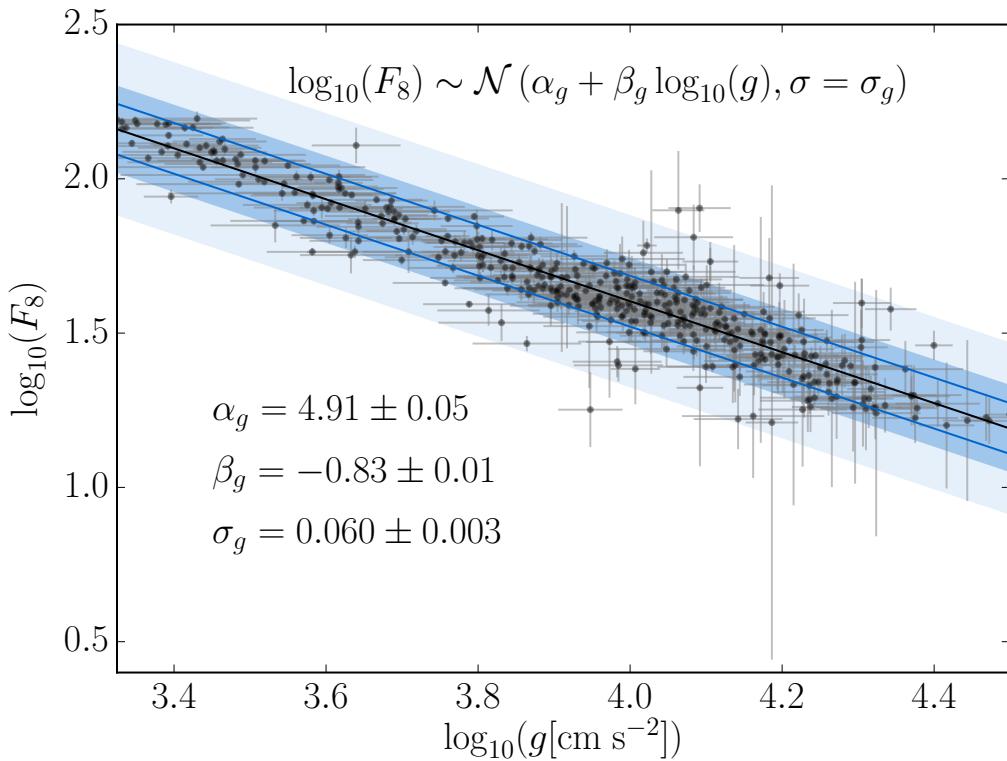


Figure 5-4:  $\log g$  versus flicker. As in 5-3 this figure shows the model, conditioned on the data. The solid black line shows the model with the best-fitting parameter values quoted in the text. The solid blue lines show the  $1\sigma$  region where the extra scatter is not included and the blue shaded regions show the  $1$  and  $2\sigma$  regions with the additional scatter.

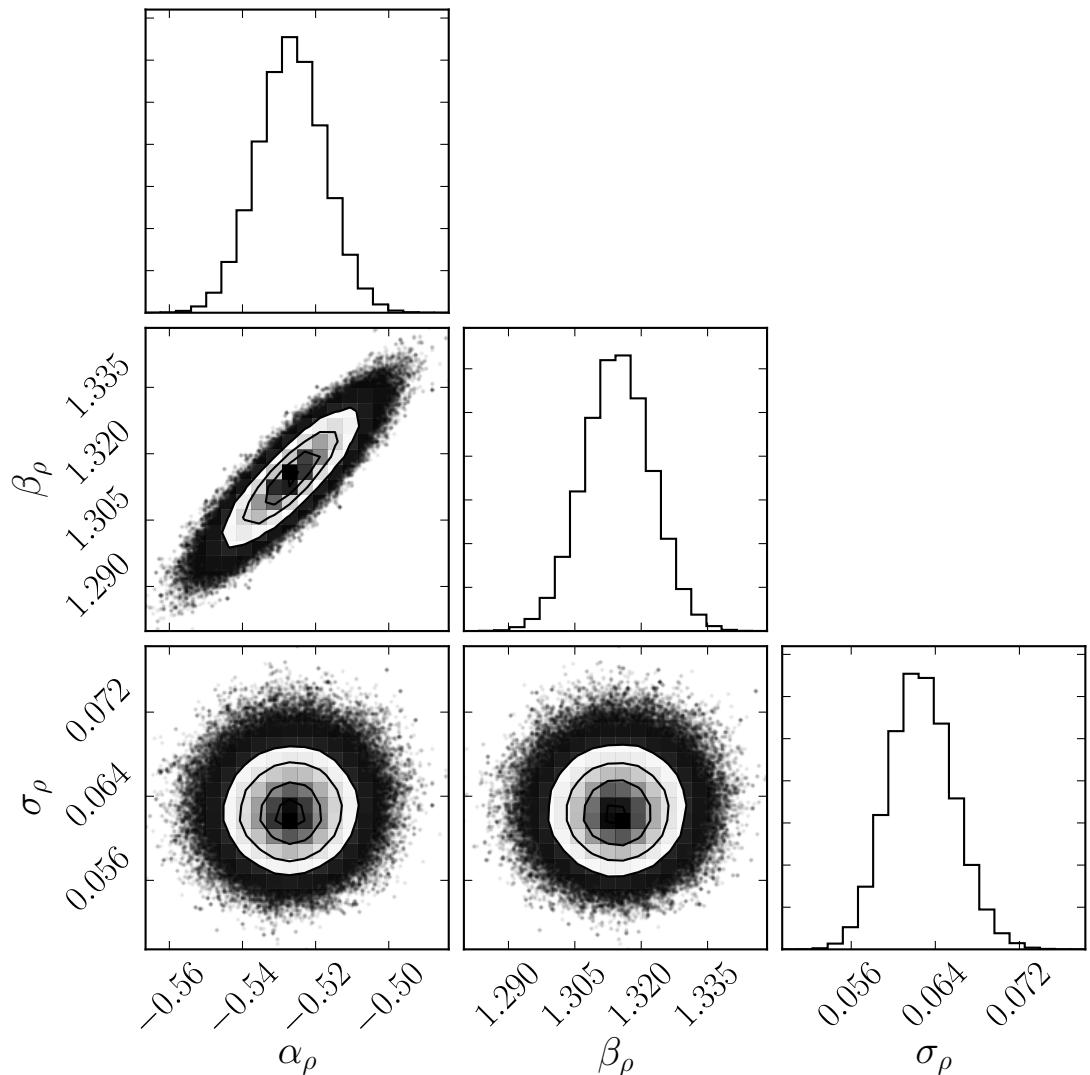


Figure 5-5: Marginal posterior PDFs of the model parameters for  $\rho_*$ . This figure was generated using the `corner` python package (Foreman-Mackey et al., 2014).

## 5.3 Discussion

We have recalibrated the relation between short timescale brightness fluctuations in the *Kepler* light curves of stars (flicker) with both stellar density and surface gravity, whilst including parameters to describe the intrinsic scatter in these relationships, presented in table 5.2.2. The terms  $\sigma_\rho$  and  $\sigma_g$  are both non-zero, suggesting that there *is* an additional source of scatter in the relations, not accounted for by the observational uncertainties alone. This is either caused by intrinsic scatter in the physical relationship between flicker and density and  $\log g$ , produced by some physical process that is not accounted for in the model, or by an underestimation of the observational uncertainties. We also tested a model with both an additional variance term *and* a term that included flicker-dependent variance. We found that the need for additional flicker-dependent variance was not supported by the data, indicating that the intrinsic scatter in the relations between flicker,  $\log g$  and  $\rho_*$  does not depend on evolutionary state.

This is a simple ‘fitting a line to data’ exercise but it continues the discussion of probabilistic modelling that is an active topic within the fields of exoplanet and stellar astronomy. We used Hierarchical Bayesian Modelling (HBM) to constrain the intrinsic scatter in the relationship between flicker, surface gravity and density and included the effects of the non-negligible two-dimensional observational uncertainties. Relationships between astronomical parameters are almost always non-deterministic; an element of stochasticity effects the physical parameters of stars so one can never perfectly predict  $y$  given an observation of  $x$ . We advocate a probabilistic approach in both the ‘fitting the model to data’ step, *and* when using an empirically calibrated model to predict parameter values. The fitting stage benefits because if the relationships between parameters are falsely assumed to be deterministic, they will be skewed by data points with uncertainties that only represent measurement error and no additional scatter. The prediction stage benefits from the stochastic treatment both because a probability distribution is in many ways more representative of an observation than a point estimate, and because posterior PDF samples can be used

in subsequent studies (provided the prior used during the fitting process is described).

We provide posterior PDF samples at <https://zenodo.org/deposit/105051/>. Whenever a prediction for the surface gravity or density of a star is required, for a given estimate of flicker, we recommend using these posterior samples within the calculation of  $\rho_*$  or  $\log g$  and its (Monte Carlo) uncertainty. These posterior samples will naturally fold in the covariances between parameters. Simple analytical uncertainty propagation is only valid when uncertainties are Gaussian and uncorrelated which is rarely true and certainly not the case when the model is a straight line (the slope and intercept are always correlated<sup>4</sup>). A flicker value with uncertainties (or even better: posterior PDF samples), input into our model will result in a probability distribution over stellar densities or surface gravities which reflects both the uncertainties on the flicker measurement, the uncertainties on the model parameters *and* the intrinsic scatter in the flicker- $\rho_*$ - $\log g$  relations.

---

<sup>4</sup>Although they can be orthogonalised via a change of variables.



# Chapter 6

## LSST

The following project was motivated by a call for community input on the optical and infrared follow-up requirements for stellar rotation and magnetism studies with *LSST*. I present some of the initial findings of this project here and intend to continue to develop this project in future. This work was undertaken in collaboration with others but most of the analysis was performed by myself. Exceptions are clearly indicated.

### Abstract

The Large Synoptic Survey Telescope (*LSST*) will provide sparse but precise ground-based photometry for billions of cool dwarfs in clusters and the field. We explore the potential of *LSST* for large-scale rotation period measurement with an emphasis on applications to gyrochronology, the method of inferring stellar ages from rotation periods. With its ten year baseline, *LSST* light curves will be sensitive to long rotation periods which are characteristic of old and low-mass stars. New asteroseismic data from the *Kepler* spacecraft have revealed that magnetic braking may cease at around Solar Rossby number, implying that gyrochronology is not applicable to old stars. By measuring rotation periods of old, slowly rotating, low-mass stars we can decisively test the age-rotation relations at all ages. Of particular interest are the open clusters with precisely measured isochronal ages. These clusters will allow us to recalibrate the age-rotation relations in the old, low-mass regime, provided we can measure the photometric rotation periods of their members. Using representative distributions of stellar ages and spectral types from *TRILEGAL* outputs, we simulated thousands of light curves using a gyrochronology relation, a simple star-spot model and approximate *LSST* cadence. By running a period recovery pipeline, we characterise the sensitivity of *LSST* to rotation periods as a function of spectral type, apparent magnitude and rotation period. Using the full ten-year data set we find that *LSST* is most sensitive to rotation periods between 10 and 20 days. Its sensitivity

falls at short periods due to the sparsity of its sampling and at longer periods due to the lower variability amplitudes of slow rotators as well as their smaller apparent magnitudes.

### 6.0.1 Introduction

The Large Synoptic Survey Satellite (*LSST*) is a 8.4 metre telescope with a 9.6 square degree field of view in Cerro Pachón, Chile, currently under construction. It is designed to observe 18,000 square degrees in the southern sky (south of +10 degrees, declination) in six Sloan Digital Sky Survey (*SDSS*) filters: *ugrizy*. During its main mode of operation (90% of the time), *LSST* will perform two fifteen second exposures per visit, with one thousand visits per night and will have a faint limit of around 24.5 in r-band. For the remaining 10% of the time, *LSST* will focus on a small number of ‘deep drilling fields’. These fields are yet to be determined but could be, for example, the Large and Small Magellanic Clouds, the galactic plane, and so on. These fields will receive targeted, repeat observations of, for instance 200 observations over a 40-hour period after which the faint limit could be extended to around 28 apparent magnitudes. First light is currently scheduled for 2021 and data release one of eleven is expected to contain eighteen billion objects (Ivezic et al., 2008).

*LSST* will provide photometric rotation periods for a new region of period-mass-age parameter space. The *Kepler* spacecraft focused on Earth-like planets with Sun-like hosts, thus the majority of its targets were G type, with fewer K and M dwarfs. Unlike *Kepler* however, any target falling within *LSST*’s field of view will be observed — not just those on a predetermined target list. In addition, due to the large collecting area of *LSST*, it will be sensitive to a large number of faint stars, including many K and M dwarfs. Being ground-based, its lifetime does not depend on the reliability of moving parts or fuel and *LSST* will run for 10 years, more than double the length of the *Kepler* prime mission. This long baseline will enable rotation signatures of faint, slowly rotating stars to be detected, populating both low-mass and old regions of the age-rotation parameter space. *LSST* will provide an extremely different but complementary data set to *Kepler*: whereas *Kepler* data are dense and evenly spaced,

*LSST* light curves will have sparse, irregular cadence. Despite the sparsity of *LSST* data, its irregular cadence will move the minimum recoverable period limit towards shorter rotation periods<sup>1</sup>. With the exception of the deep drilling fields, the minimum interval between exposures will be around three days for the majority of targets. This corresponds to a nominal minimum recoverable period of around six days, however, the irregularity of *LSST* cadence, combined with the ten year observing window will reduce this lower bound.

In order to investigate the potential of *LSST* data regarding gyrochronology, it is essential that we estimate the range of rotation periods that will be detectable. In what follows we describe our *LSST*-like, simulated data-set and our rotation period recovery pipeline.

### 6.0.2 Simulations

We developed a simple *LSST* cadence model for our simulations<sup>2</sup>. This model generates a cadence pattern for each object by requiring that it only be observed during the night and whilst that field is visible, *i.e.* for half of the year. Each object is visited every three days on average during the observable season and visits are clustered around a season with a Gaussian shape. A histogram of the number of visits per week as a function of time for a given object or field is shown in figure 6-1. The code used to simulate *LSST* cadence was written by J. Davenport and is available at <https://github.com/jradavenport/MW-Flare>.

We used the *TRILEGAL* (Girardi et al., 2012) galaxy simulation code to generate field stellar populations for five hypothetical *LSST* fields. These fields were centred on the same galactic longitude,  $l = 45$ , but different galactic longitude:  $b = -10, -20, -40, -60, -80$ . In this analysis we present results from the  $b = -80$  field only, leaving a more comprehensive study of different stellar populations to future analysis. *TRILEGAL* takes the coordinates and size of a field within the galaxy

---

<sup>1</sup>By comparison, the interval between *Kepler* long cadence observations is 30 minutes. It is therefore sensitive to variability on timescales as short as 1 hour.

<sup>2</sup>In future we intend to use the *LSST* operations simulator (OpSim Ridgway et al., 2012) to generate a more realistic cadence pattern for each target, depending on its position on the sky.

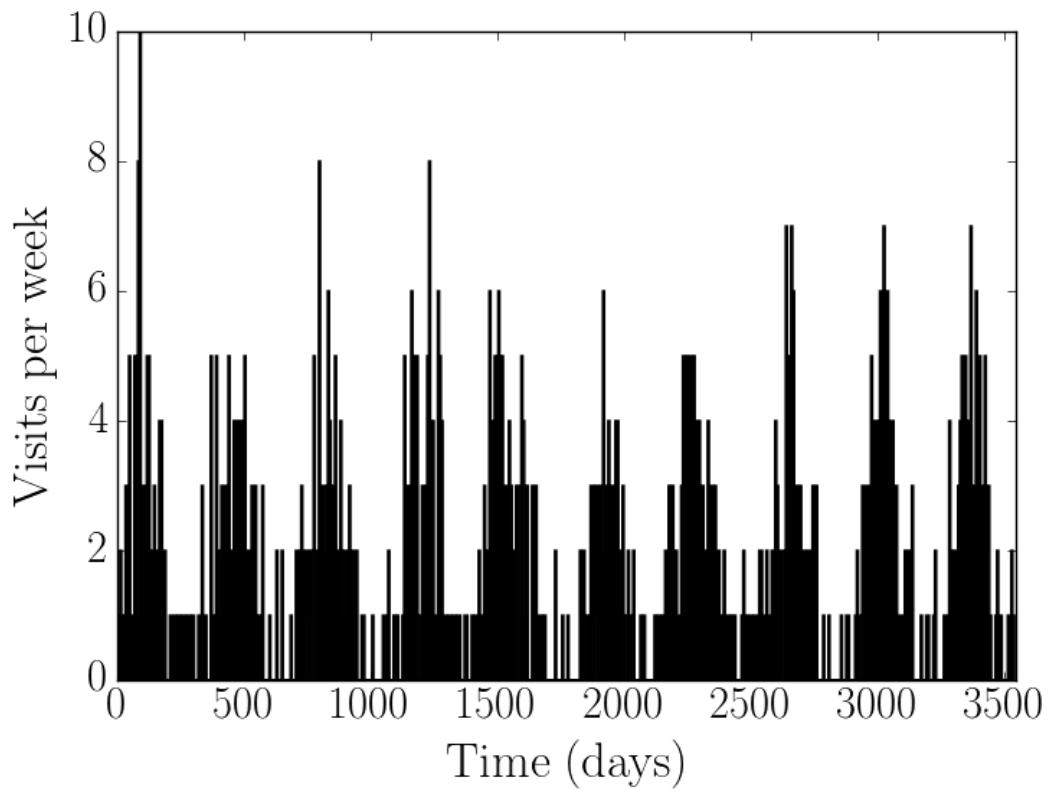


Figure 6-1: A histogram of the number of visits per week as a function of time for a given object or field observed by *LSST* as used in our simulations. Each object is observed only during the night and for half a year at a time. The code used to generate this plot was written by J. Davenport, <https://github.com/jradavenport/MW-Flare>.

as input parameters and simulates the population of stars in that field. It returns a catalogue of simulated stars with their properties: age, effective temperature, *ugriz* magnitudes, and others. We randomly selected 20,000 stars from a field centred on galactic coordinates  $l = 45$ ,  $b = -80$  in order to produce a representative but manageable target sample. Selected stars had  $r$ -band magnitudes between 16 and 28 and  $\log g > 4$ . These target stars were then separated into ‘cool’ ( $T_{\text{eff}} < 6250$ ) and ‘hot’ ( $6250 < T_{\text{eff}}$ ) temperature bins. Rotation periods for the cool stars were calculated using the Angus et al. (2015b) gyrochronology relation which converts *TRILEGAL* ages and  $B - V$  colour (calculated from *TRILEGAL*  $g - r$ ) into rotation periods. Hot stars ( $T_{\text{eff}} \gtrsim 6250$ ) lack a significant convective envelope and since the combination of convective plasma motion and stellar rotation is responsible for magnetic field generation, these stars do not undergo magnetic braking. As such, their rotation periods cannot be estimated using gyrochronology. To assign rotation periods to these hot stars we fit a sum of two Gaussian functions to the rotation periods of hot stars in the McQuillan et al. (2014) catalogue and randomly sampled from the resulting distribution. The distribution of hot star rotation periods is shown in figure 6-2.

The overall distributions of effective temperatures and theoretical rotation periods for stars used in our simulations are shown in figures 6-3 and 6-4.

### Synthesising light curves

Once theoretical rotation periods had been assigned to both hot and cool field stars, we used code similar to that used in Aigrain et al. (2015b) to simulate light curves. These light curves are calculated by placing dark spots on a rotating sphere and integrating the total resulting flux over the surface. Stellar flux variations produced by dark active regions on the surface are typically non-sinusoidal and this simple spot model provides a more accurate representation of stellar light curves than a simple sinusoid. However, it should be noted that this code can be adjusted to produce more realistic light curves by altering spot lifetimes and including differential rotation with spot migration. Stars with spot lifetimes that are short relative to their rotation periods will display quasi-periodic brightness variations in their light curves.

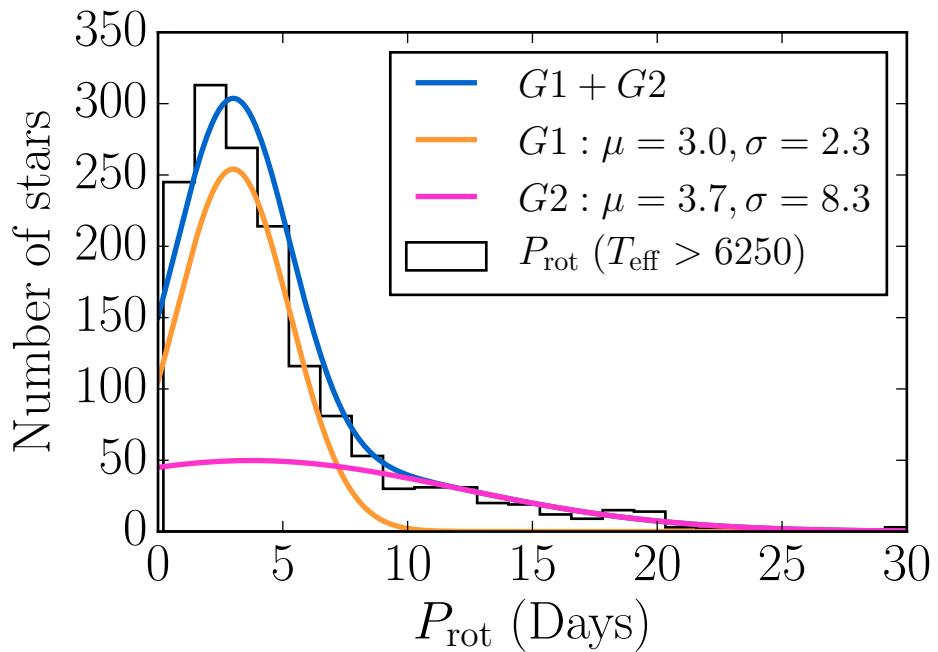


Figure 6-2: A histogram of the rotation periods of stars with  $T_{\text{eff}} > 6250$  K from the McQuillan et al. (2014) catalogue. We fit a sum of two Gaussians to this distribution in order to assign rotation periods to the hot stars in our *TRILEGAL* samples.

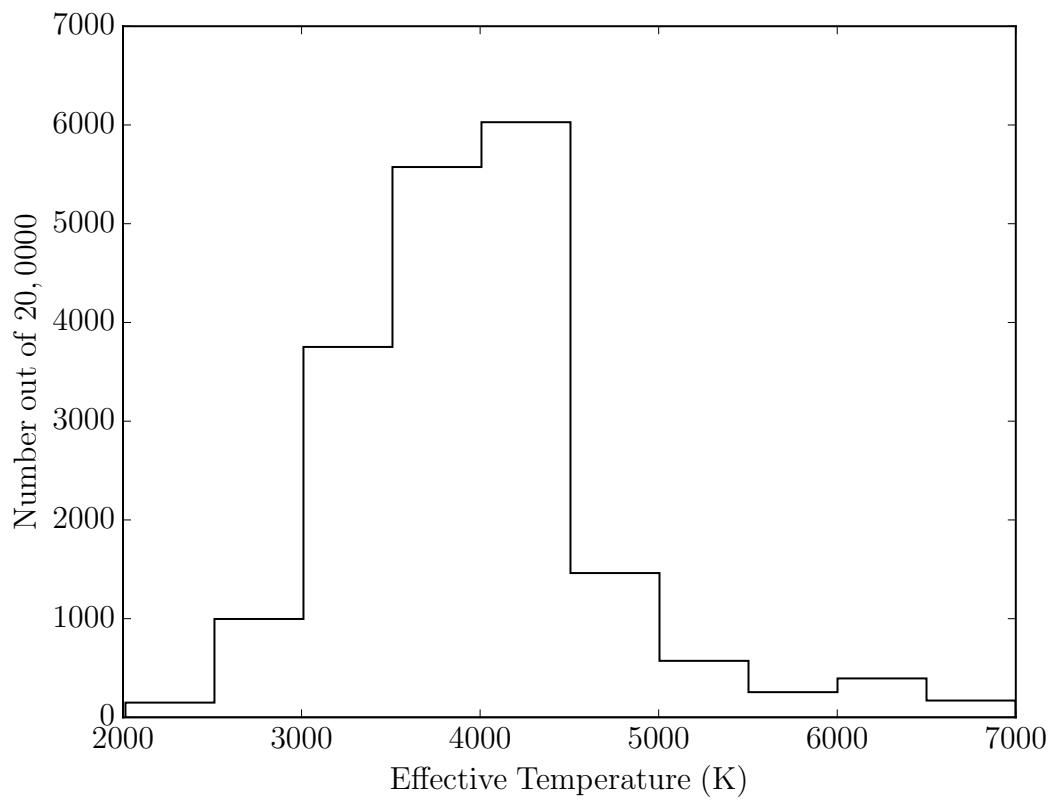


Figure 6-3: The effective temperature distribution for stars used in our simulations. We selected 20,000 stars from a field centred on galactic coordinates  $l = 45$ ,  $b = -80$ .

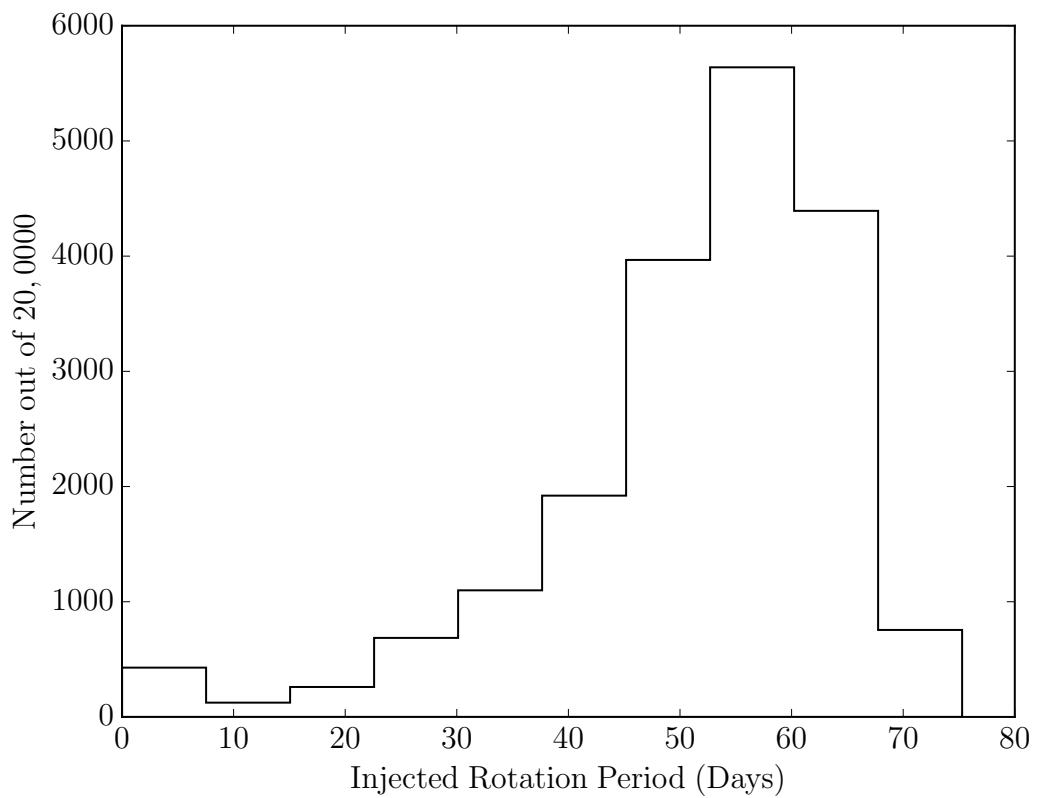


Figure 6-4: The distribution of rotation periods used to generate synthetic light curves. We selected 20,000 stars from a field centred on galactic coordinates  $l = 45$ ,  $b = -80$ .

Including differential rotation will compound this effect and further perturb star spot light curve features from strict periodicity. We fixed the mean spot lifetime at 30.5 days for all simulations and did not include differential rotation. This is a relatively conservative spot lifetime as it is appropriate for G-stars and later types tend have longer spot lifetimes. We intend to include differential rotation in a future analysis.

In order to assign appropriate amplitudes to the simulated light curves, the relation between rotation period, amplitude of variability and  $T_{\text{eff}}$  was approximated for *Kepler* stars with rotation periods measured by McQuillan et al. (2014)<sup>3</sup>. Over a range of temperature and rotation period bins, the mean range of light curve variability was computed. The range of variability, as defined by McQuillan et al. (2014) is the difference between the 5th and 95th percentiles of the light curve flux distribution. Figure 6-5 shows the mean range of variability for stars in different temperature bins, as a function of their rotation periods. This figure shows that cooler stars have larger variability amplitudes than hotter stars, on average. This is to be expected since cooler stars have deeper convective zones and stronger magnetic fields. It also shows that the variability amplitude declines with rotation period (or perhaps age). The numbers indicated in this figure were used to generate synthetic light curves. Each star in the sample was assigned an amplitude that was drawn from a Gaussian centred on the mean value of the appropriate temperature and period bin, with variance corresponding to the variance within each bin. An important caveat is that the amplitudes of variability in the McQuillan et al. (2014) catalogue are *only* reported for the positive detections. In other words, no amplitudes were reported for those *Kepler* stars that did not show clear periodic variability and did not result in a reliable period measurement. The stars without period measurements are likely to either show very low amplitude brightness variations, highly non-periodic variations, to have very long rotation periods, or some combination of these three factors. It is relatively safe to assume however, that the stars without rotation period measurements are likely to have lower amplitudes, on the whole, than the stars with rotation periods. The amplitudes of variability as a function of rotation period and effective temperature,

---

<sup>3</sup>This analysis was performed by collaborator D. Buzasi.

derived from the McQuillan et al. (2014) catalogue are therefore only representative of the most variable targets, not the *Kepler* catalogue as a whole. In the future we intend to infer the amplitude of variability as a function of temperature for all *Kepler* stars, not just those with rotation periods. For now however, we have used a simple approximation to correct for this sample bias. We assume that if a rotation period is not detectable in a *Kepler* light curve, it is unlikely to be detectable in an *LSST* light curve. This assumption may not be valid for slow rotators whose rotational signals could be removed by the *Kepler* detrending pipeline, or difficult to detect across multiple *Kepler* quarters, but should be acceptable for most other stars. With that assumption in place, we simply scaled our results by the fraction of G, K and M type stars with measured rotation periods. We found that the fraction of G dwarfs with measured rotation periods was 0.24, the fraction of K dwarfs was 0.54 and the fraction of Ms was 0.83. These numbers reflect the fact that lower-mass stars are more active. I return to this point in section §6.0.3.

Simulated light curves were sampled using our *LSST* cadence model and finally, white noise was added. The variance of the white noise was different for each star depending on its *r*-band magnitude. We used values provided in Jacklin et al. (2015) to estimate the relation between *r*-band magnitude and white noise variance. Example light curves are shown in figures 6-6 and 6-7.

We attempted to recover the rotation periods of the 20,000 simulated stars using a Lomb-Scargle (LS) periodogram (Lomb, 1976; Scargle, 1982). LS periodograms were computed for each light curve over a uniform grid of 1000 periods ranging from 2 to 100 days. The position of the highest peak in the periodogram was recorded as the measured period. We did not define a detection threshold — a rotation period, or rather, a peak position was recovered from every light curve. We hope to define such a detection threshold in future. As explained previously, a LS periodogram is not ideal for measuring rotation periods from stellar light curves because star spots tend to produce non-sinusoidal, quasi-periodic brightness variations. A commonly used alternative to the LS periodogram is the auto-correlation function (ACF) which is better suited to non-sinusoidal, quasi-periodic signals, as demonstrated by McQuillan

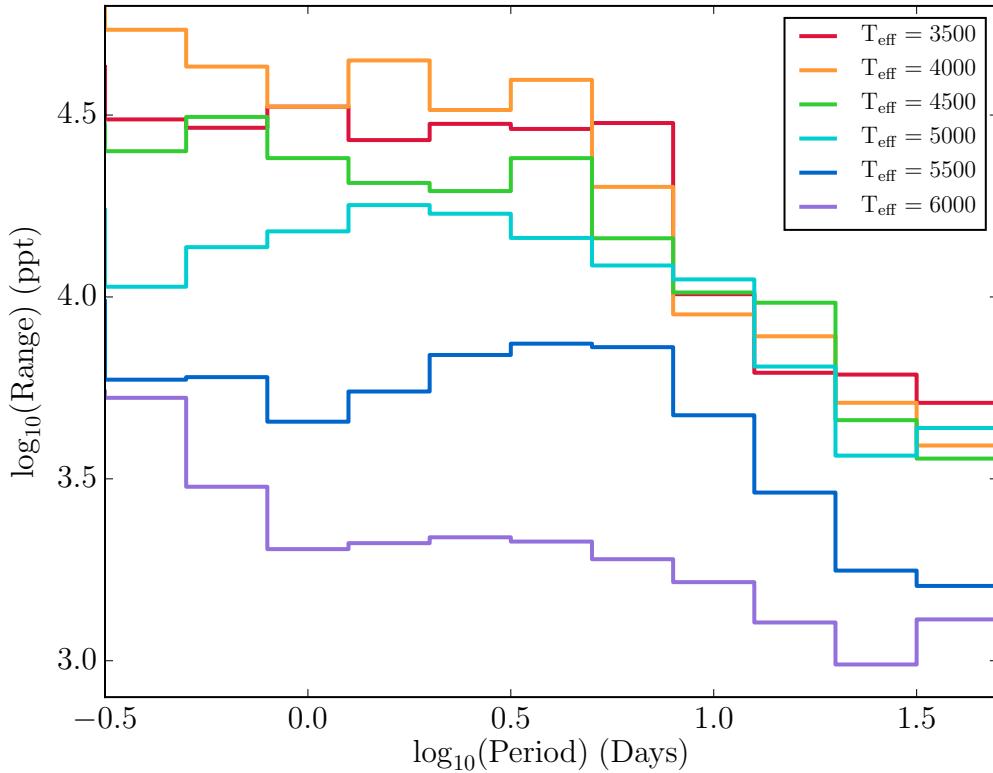


Figure 6-5: Mean amplitudes of variability as a function of rotation period, for stars in the McQuillan et al. (2014) catalogue. For each temperature bin indicated, the mean range of variability (the difference between the 5th and 95th flux percentile) in that bin for all targets *with a measured rotation period* in the McQuillan et al. (2014) catalogue is plotted as a function of their rotation periods. These numbers were used to generate synthetic light curves — stars with temperatures and rotation periods indicated on this plot were assigned amplitudes accordingly.

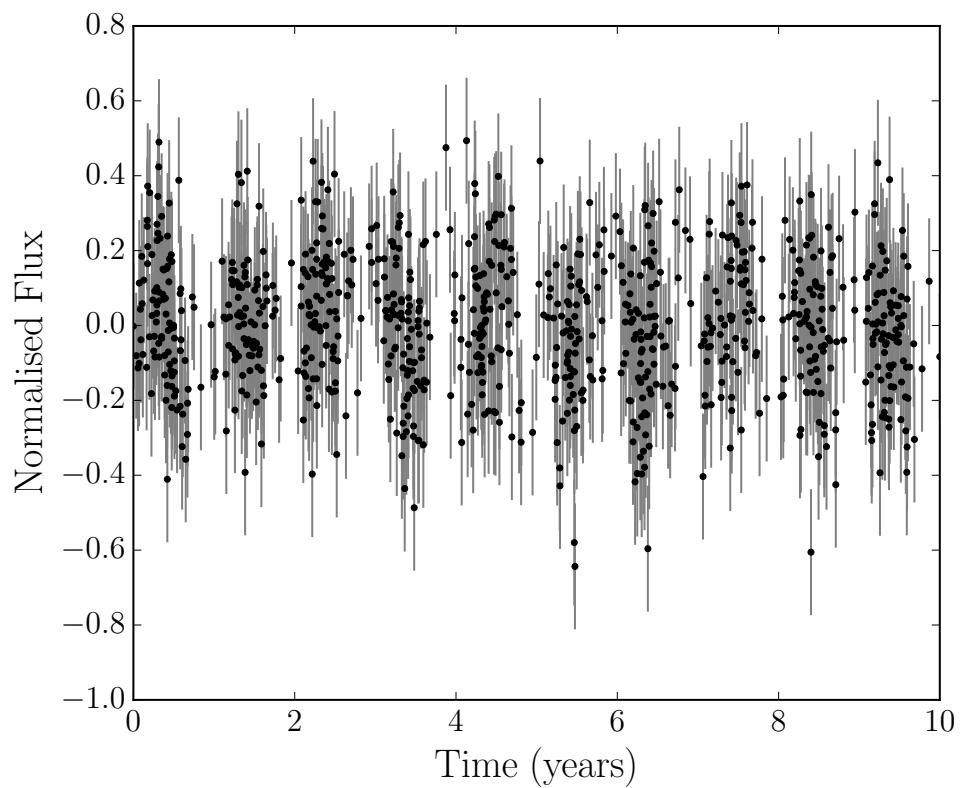


Figure 6-6: An example of a simulated *LSST* light curve. This synthetic star has  $P_{\text{rot}} = 44$  days, an amplitude of 6862 ppm, a temperature of 2904 K and an *r*-band magnitude of 25.6.

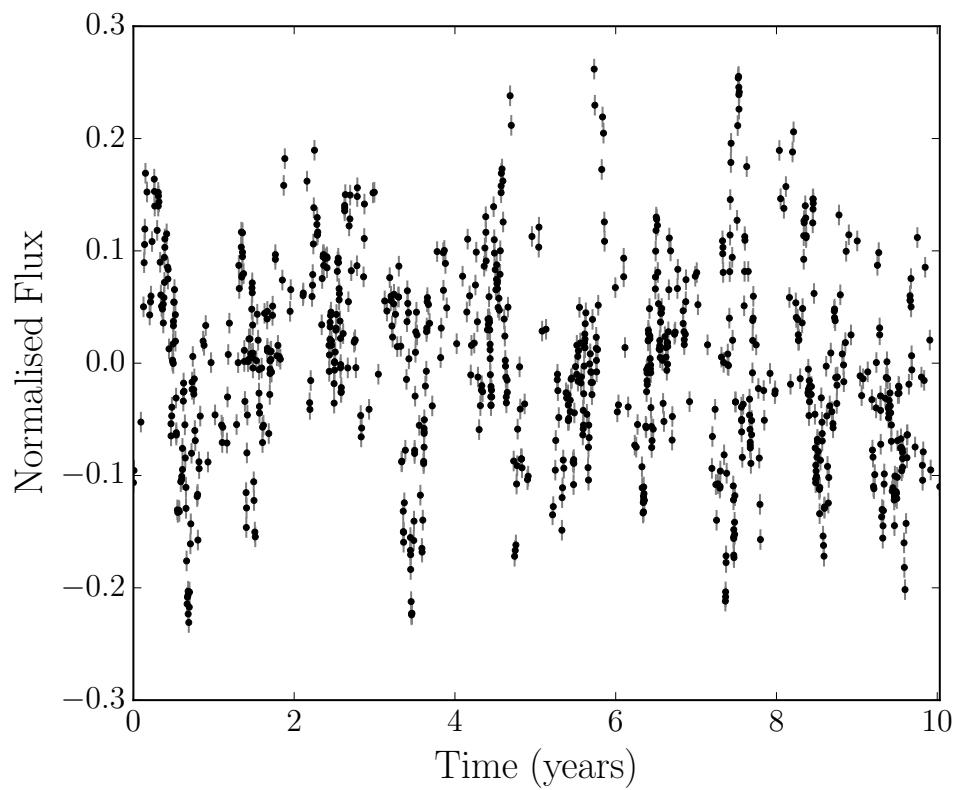


Figure 6-7: This synthetic star has  $P_{\text{rot}} = 31$  days, an amplitude of 7236 ppm, a temperature of 3266 K and an *r*-band magnitude of 20.4.

et al. (2013a). Unfortunately however, evenly spaced data are required to produce an ACF. A modified ACF method developed by ? does not require evenly spaced data and we intend to use this algorithm in a future analysis. A comparison of the LS periodogram and ACF methods by Aigrain et al. (2015b) found good agreement between the two, which is encouraging for this study. Another alternative method that has produced promising results using Gaussian processes (GPs) is currently under development (Angus et al., 2015a), however it is relatively computationally expensive. We have used the LS periodogram in these initial tests because it is fast to compute, but intend to use the GP method in future.

### 6.0.3 Results

The results of our simulations are shown in figures 6-8 and 6-9. These figures show the theoretical rotation periods used to simulate 20,000 light curves, plotted against the rotation periods that were measured from those light curves using a LS periodogram. Data points are coloured according to their temperatures in figure 6-8 and according to their *r*-band magnitudes in figure 6-9. Rotation periods less than around ten days have a low recovery fraction, which is most likely due to the sparcity of *LSST* sampling. The large outliers at long rotation periods are mostly faint, as shown in figure 6-9.

The overall fraction of stars with successfully recovered rotation periods (where ‘successfully recovered’ rotation periods are those measured to within 10% of the injected period) as a function of input period is shown in figure 6-10. The results for *G* ( $5200 < T_{\text{eff}} < 6000$ ), *K* ( $3700 < T_{\text{eff}} < 5200$ ), and *M* ( $2500 < T_{\text{eff}} < 3700$ ) spectral types are shown separately. The percentage of stars recovered in each spectral type bin has been multiplied by the fraction of *Kepler* stars with measured rotation periods in the McQuillan et al. (2014) catalogue. This rescaling should capture the fact that many stars in the *Kepler* data set do not have rotation periods since they have very low variability amplitudes, display non-periodic variability, and/or have very long rotation periods. This rescaling is only appropriate if the assumption that rotation periods are easier to measure in *Kepler* data than *LSST* data which, of course

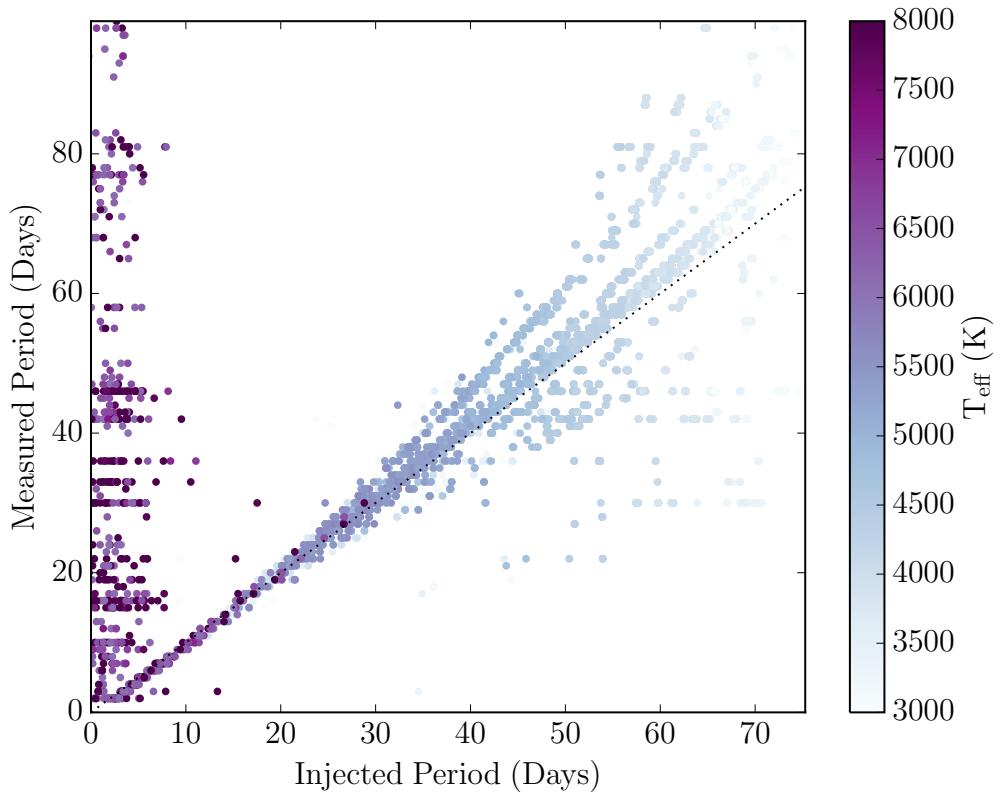


Figure 6-8: Measured versus injected rotation period for 20,000 simulated *LSST* targets in a field centred on  $l = -45$ ,  $b = -80$ . Points are coloured according to their temperatures. Rotation periods less than  $\sim 10$  days have a low recovery fraction, and this is worse for hot stars as they have lower amplitudes of variability. Quantisation appears in the  $y$ -direction due to the relatively coarse frequency grid used to compute LS periodograms.

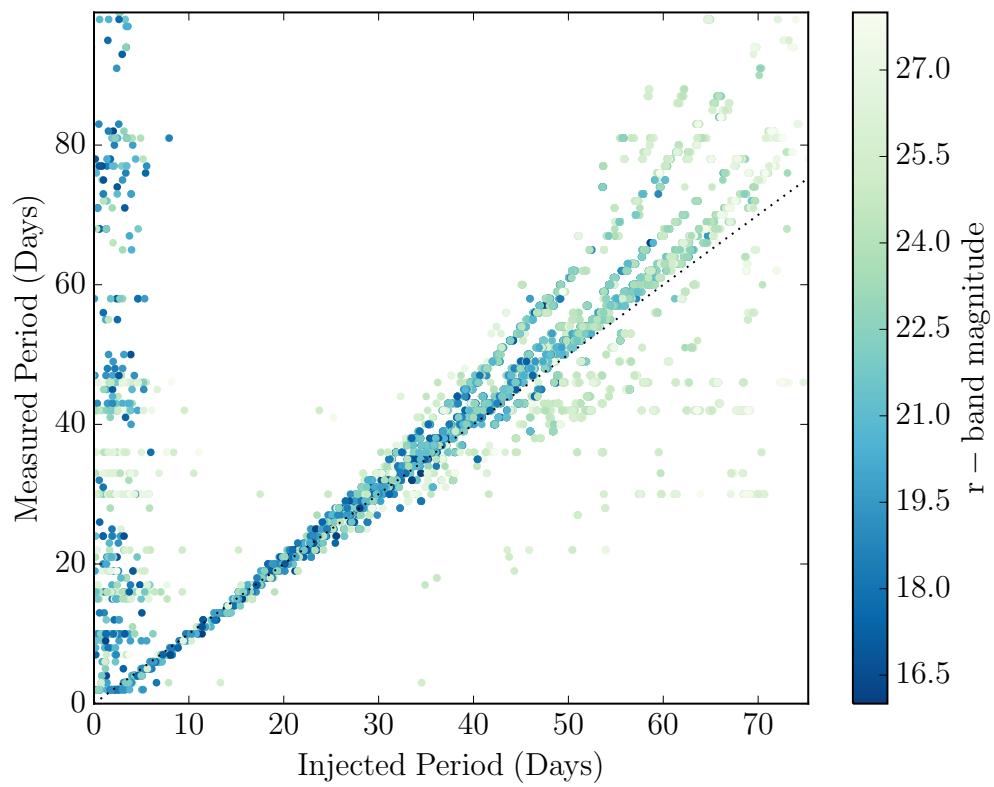


Figure 6-9: Measured versus injected rotation period for 20,000 simulated *LSST* targets in a field centred on  $l = -45$ ,  $b = -80$ . Points are coloured according to the  $r$ -band magnitude. The large outliers at long rotation periods are mostly faint. Quantisation appears in the  $y$ -direction due to the relatively coarse frequency grid used to compute LS periodograms.

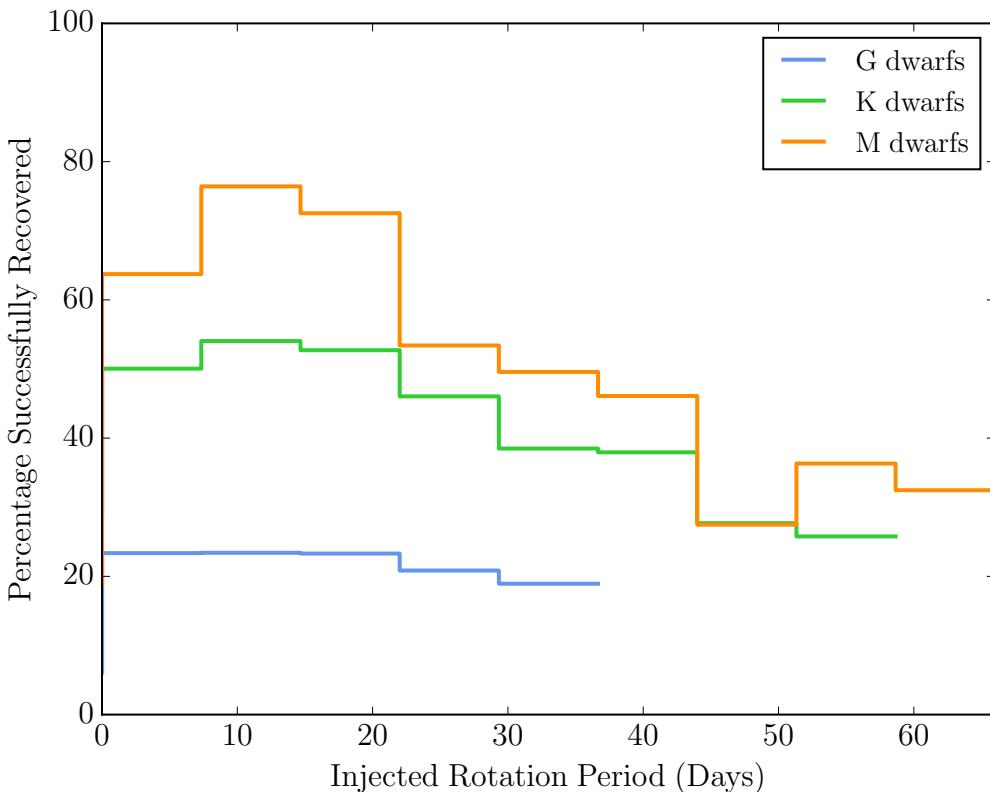


Figure 6-10: The percentage of successfully recovered rotation periods in each indicated bin, as a function of input rotation period. These histograms have been multiplied by the overall fraction of rotation periods measured by McQuillan et al. (2014) for each spectral type. These fractions are 0.24, 0.54 and 0.83 for G, K and M dwarfs respectively.

may not be valid for certain stars — especially the very slow rotators. We hope to use a more sophisticated rescaling in future.

There are no *G* dwarfs with rotation periods longer than 40 days as these stars would be older than the galaxy. This is also the case for *K* dwarfs with rotation periods greater than 60 days. For all spectral types, the rotation period sensitivity of *LSST* peaks at 10-20 days. At shorter rotation periods the sensitivity is limited by the minimum interval between observations: around three days. At longer rotation periods the sensitivity is limited by both the smaller variability amplitudes of slow rotators and the fact that slow rotators tend to be lower mass, therefore fainter.

#### 6.0.4 Conclusions

*LSST* light curves have the potential to provide rotation periods for a new region of stellar age and rotation phase-space. Specifically, it may provide rotation periods for older and lower mass stars than *Kepler* was able to. The *Kepler* detrending algorithms remove long-timescale stellar variability of the kind produced by old and low-mass stars. In addition, the sample will not be skewed towards G-type stars as the *Kepler* target list was, on the contrary: the majority of stars observed by *LSST* will be M-dwarfs as these dominate the mass-function. We have investigated the range of rotation periods that are recoverable in *LSST* data. Effective temperatures, *r*-band magnitudes, ages and metallicities of 20,000 stars in a hypothetical *LSST* field (centred on  $l = 45$ ,  $b = -80$ ), were generated using the *TRILEGAL* stellar population simulation software (Girardi et al., 2012). Theoretical rotation periods were then calculated for these stars using the (Angus et al., 2015b) gyrochronology relation. Based on these theoretical rotation periods, light curves were simulated for the stars using a simplified version of the (Aigrain et al., 2015a) code. A Lomb-Scargle periodogram was computed for each light curve and a period measured from the position of the highest peak. We found that *LSST* is most sensitive to rotation periods between 10 and 20 days. Its sensitivity falls at short periods due to the sparsity of its sampling and at longer periods due to the lower variability amplitudes and smaller apparent magnitudes of slowly rotating stars.

This work was performed as part of a project to understand the ground-based optical and infrared follow-up requirements for studying stellar magnetism and rotation with *LSST*. Going forwards, I intend to expand upon the work presented here by implementing the improvements listed below.

- To use OpSim, the official *LSST* cadence simulator (Ridgway et al., 2012), to generate a more realistic cadence model.
- To use the gyrochronology relation of van Saders and Pinsonneault (2013). This age-rotation relation applies to subgiants as well as dwarfs. Including slightly evolved stars in our models will provide an indication of the level of expect level

of subgiant contamination.

- To use a more realistic spot model. Our current spot model is simplistic—a single spot lifetime is used for all stars, rather than a distribution and differential rotation is not included. Using a model that incorporates differential rotation and variable spot lifetimes will produce more realistic light curves with quasi-periodic features.
- To apply our method to the whole sky. The distribution of stellar spectral types and ages varies with galactic coordinate. For example, there are more young stars at galactic latitudes near zero since star formation is still occurring in the thin disc of the Milky Way. Rotation period recovery success rate depends on both spectral type and age and is therefore likely to vary for different fields in the galaxy.
- To investigate the rotation period recovery potential of the deep drilling fields. These fields will have much shorter intervals between subsequent observations. Very short rotation periods are therefore likely to be detectable within them.
- To incorporate a realistic correlated noise model. Only white noise was included in our simulations, but correlated noise is extremely important in *Kepler* light curves, making it difficult to measure long rotation periods. Correlated noise is also likely to plague *LSST* data and we hope to incorporate a realistic correlated noise model into our simulations if one becomes available.



# Chapter 7

## Conclusions

In chapter 2 I used new results from *Kepler* asteroseismology to recalibrate the relation between rotation period, age and colour. The short cadence *Kepler* light curves of several cool dwarf, Solar-like oscillators were analysed by Chaplin et al. (2014) who provide a catalogue of ages for these targets. We used photometric rotation periods of these stars published in García et al. (2014). They used ACF and wavelet methods to measure the periodicity of star spot-induced brightness modulations in their *Kepler* light curves. We fit a gyrochronology model to these data, of the form  $P = A^n \times a(B - V - c)^b$ , where  $P$  is rotation period (in days),  $A$  is age (in Myr),  $B$  and  $V$  are  $B$  and  $V$  band magnitudes respectively and  $a$ ,  $b$ ,  $c$  and  $n$  are dimensionless free parameters. We found evidence to suggest that this function does not provide a good fit to the data and that some old *Kepler* asteroseismic targets are more rapidly rotating than expected, given their age and colour. van Saders et al. (2016) present an adaptation to their theoretical model which is able to reproduce the trends in this data set. They suggest that there is a critical Rossby number ( $Ro = 2.16$ ) at which the magnetic dynamo that drives angular momentum loss shuts off. New rotation periods from the repurposed *Kepler* mission, *K2* may shed light on this controversial topic. In addition, a simple injection and recovery test performed on *Kepler* data would outline the sensitivity cut-off for rotation period detection in *Kepler* light curves. This test would reveal whether rotation period detection bias may be responsible for producing the results found in chapter 2 and van Saders et al. (2016).

In chapter 3 I present a method for searching for periodic signals (*e.g.* rotation periods) in *K2* light curves without detrending. The pointing precision of *Kepler* was dramatically reduced when its third reaction wheel broke. *K2* light curves are contaminated with high amplitude systematic features as a result. Accurately modelling the noise in these light curves is essential in order to search for exoplanets, perform asteroseismology or measure stellar rotation periods. We constructed a noise model by decomposing all *K2* light curves from campaign 1 into a set of orthogonal basis vectors called ‘Eigen Light Curves’ (ELCs). Instead of detrending, fitting the noise model and subtracting it from the data, we model the noise and the signal simultaneously and marginalise over the noise model. The Systematics-Insensitive Periodogram (SIP) is a method that uses a linear combination of 150 of these ELCs to model a light curve while simultaneously fitting a sinusoid to the data at a given frequency. Finding the amplitudes of the best-fit sinusoids over a grid of frequencies produces a SIP. The SIP is particularly effective for red giant asteroseismology as these signals are typically sinusoidal and plagued by a six-hour thruster firing signal that most detrending algorithms are unable to remove. The capabilities of the SIP are limited for rotation period inference as it is difficult to separate systematics from physical signals on long timescales and because a sinusoid is an imperfect model for stellar rotation.

In chapter 4 I present a new method for inferring precise and accurate rotation periods from *Kepler* light curves using Gaussian processes. A Gaussian process model is a flexible, semi-parametric model that is capable of capturing the non-sinusoidal, quasi-periodic nature of stellar light curves. I compare the GP method to the Lomb-Scargle periodogram and ACF methods, finding it to be more accurate and precise than both. In addition it provides accurate uncertainties because, in our implementation, a full posterior PDF of the rotation period parameter is explored. Unfortunately, while this method works well on noise-free simulations, it is currently extremely computationally expensive when applied to noisy light curves. It is therefore currently most useful for individual targets rather than large ensembles of light curves.

In chapter 5 of this thesis I use *Kepler* flicker measurements from Bastien et al.

(2013) and fundamental stellar parameters from asteroseismology to recalibrate the relation between flicker, surface gravity and stellar density. These relations are stochastic, although they have been treated deterministically in the past (*e.g.* Bastien et al., 2013; Kipping et al., 2014) and there is significant intrinsic scatter that is not accounted for by the observational uncertainties. I use hierarchical bayesian modelling to quantify model these relations and quantify the additional scatter required to explain the observations.

Finally, in chapter 6 I investigate opportunities for inferring rotation periods from *LSST* light curves. *LSST* data will be sparse and unevenly sampled—quite the opposite of *Kepler*. If we can detect long rotation periods of old, low-mass stars in *LSST*'s ten years of observations, currently empty parts of the rotation period-mass-age parameter space will be filled in, with promising implications for gyrochronology. We find that *LSST* is most sensitive to rotation periods between 10 and 20 days, and its sensitivity drops at longer rotation periods since the majority of slow rotators are low-mass and therefore faint.

I began this thesis by introducing the *Kepler* spacecraft which has provided all the data used in this work. I would like to conclude by saying that, in my opinion, *Kepler*'s legacy has to be the greatest of any single purpose astronomical instrument ever built; not just for exoplanets but for stellar astrophysics too. *Kepler* has revived the astronomical community's general interest in stars: after all, one can only understand an exoplanet as well as one understands the star it orbits. Were it not for *Kepler*, this thesis would look very different and I, for one, am very grateful for the unique challenges posed by its rich, unrivalled data set.



# Bibliography

- Abe, F., Bennett, D. P., Bond, I. A., Eguchi, S., Furuta, Y., Hearnshaw, J. B., Kamiya, K., Kilmartin, P. M., Kurata, Y., Masuda, K., Matsubara, Y., Muraki, Y., Noda, S., Okajima, K., Rakich, A., Rattenbury, N. J., Sako, T., Sekiguchi, T., Sullivan, D. J., Sumi, T., Tristram, P. J., Yanagisawa, T., Yock, P. C. M., Gal-Yam, A., Lipkin, Y., Maoz, D., Ofek, E. O., Udalski, A., Szewczyk, O., Źebruń, K., Soszyński, I., Szymański, M. K., Kubiak, M., Pietrzynski, G., and Wyrzykowski, L. (2004). Search for Low-Mass Exoplanets by Gravitational Microlensing at High Magnification. *Science*, 305:1264–1267.
- Aigrain, S., Collier Cameron, A., Ollivier, M., Pont, F., Jorda, L., Almenara, J. M., Alonso, R., Barge, P., Bordé, P., Bouchy, F., Deeg, H., de La Reza, R., Deleuil, M., Dvorak, R., Erikson, A., Fridlund, M., Gondoin, P., Gillon, M., Guillot, T., Hatzes, A., Lammer, H., Lanza, A. F., Léger, A., Llebaria, A., Magain, P., Mazeh, T., Moutou, C., Paetzold, M., Pinte, C., Queloz, D., Rauer, H., Rouan, D., Schneider, J., Wuchter, G., and Zucker, S. (2008). Transiting exoplanets from the CoRoT space mission. IV. CoRoT-Exo-4b: a transiting planet in a 9.2 day synchronous orbit. *Astronomy & Astrophysics*, 488:L43–L46.
- Aigrain, S., Hodgkin, S. T., Irwin, M. J., Lewis, J. R., and Roberts, S. J. (2015a). Precise time series photometry for the Kepler-2.0 mission. *Monthly Notices of the Royal Astronomical Society*, 447:2880–2893.
- Aigrain, S., Llama, J., Ceillier, T., Chagas, M. L. d., Davenport, J. R. A., García, R. A., Hay, K. L., Lanza, A. F., McQuillan, A., Mazeh, T., de Medeiros, J. R., Nielsen, M. B., and Reinhold, T. (2015b). Testing the recovery of stellar rotation signals from Kepler light curves using a blind hare-and-hounds exercise. *Monthly Notices of the Royal Astronomical Society*, 450:3211–3226.
- Aigrain, S., Parviainen, H., and Pope, B. J. S. (2016). K2SC: Flexible systematics correction and detrending of K2 light curves using Gaussian Process regression. *Monthly Notices of the Royal Astronomical Society*.
- Aigrain, S., Pont, F., and Zucker, S. (2012a). A simple method to estimate radial velocity variations due to stellar activity using photometry. *Monthly Notices of the Royal Astronomical Society*, 419:3147–3158.

- Aigrain, S., Pont, F., and Zucker, S. (2012b). A simple method to estimate radial velocity variations due to stellar activity using photometry. *Monthly Notices of the Royal Astronomical Society*, 419:3147–3158.
- Ambikasaran, S., Foreman-Mackey, D., Greengard, L., Hogg, D. W., and O’Neil, M. (2014). Fast Direct Methods for Gaussian Processes. *ArXiv e-prints*.
- Andrews, J. J., Price-Whelan, A. M., and Agüeros, M. A. (2014). The Mass Distribution of Companions to Low-mass White Dwarfs. *Astrophysical Journal, Letters*, 797:L32.
- Angus, R., Aigrain, S., and Foreman-Mackey, D. (2015a). Probabilistic stellar rotation periods with Gaussian processes. *IAU General Assembly*, 22:2258396.
- Angus, R., Aigrain, S., Foreman-Mackey, D., and McQuillan, A. (2015b). Calibrating gyrochronology using Kepler asteroseismic targets. *Monthly Notices of the Royal Astronomical Society*, 450:1787–1798.
- Barban, C., Michel, E., Martic, M., Schmitt, J., Lebrun, J. C., Baglin, A., and Bertaux, J. L. (1999). Solar-like oscillations of Procyon A: stellar models and time series simulations versus observations. *Astronomy & Astrophysics*, 350:617–625.
- Barnes, S. A. (2003). On the Rotational Evolution of Solar- and Late-Type Stars, Its Magnetic Origins, and the Possibility of Stellar Gyrochronology. *Astrophysical Journal*, 586:464–479.
- Barnes, S. A. (2007). Ages for Illustrative Field Stars Using Gyrochronology: Viability, Limitations, and Errors. *Astrophysical Journal*, 669:1167–1189.
- Barnes, S. A. and Kim, Y.-C. (2010). Angular Momentum Loss from Cool Stars: An Empirical Expression and Connection to Stellar Activity. *Astrophysical Journal*, 721:675–685.
- Basri, G., Marcy, G. W., and Graham, J. R. (1996). Lithium in Brown Dwarf Candidates: The Mass and Age of the Faintest Pleiades Stars. *Astrophysical Journal*, 458:600.
- Bastien, F. A., Stassun, K. G., Basri, G., and Pepper, J. (2013). An observational correlation between stellar brightness variations and surface gravity. *Nature*, 500:427–430.
- Bastien, F. A., Stassun, K. G., Basri, G., and Pepper, J. (2016). A Granulation “Flicker”-based Measure of Stellar Surface Gravity. *Astrophysical Journal*, 818:43.
- Bazot, M., Bourguignon, S., and Christensen-Dalsgaard, J. (2012). A Bayesian approach to the modelling of  $\alpha$  Cen A. *Monthly Notices of the Royal Astronomical Society*, 427:1847–1866.

- Becker, J. C., Vanderburg, A., Adams, F. C., Rappaport, S. A., and Schwengeler, H. M. (2015). WASP-47: A Hot Jupiter System with Two Additional Planets Discovered by K2. *Astrophysical Journal, Letters*, 812:L18.
- Bedding, T. R., Butler, R. P., Kjeldsen, H., Baldry, I. K., O'Toole, S. J., Tinney, C. G., Marcy, G. W., Kienzle, F., and Carrier, F. (2001). Evidence for Solar-like Oscillations in  $\beta$  Hydri. *Astrophysical Journal, Letters*, 549:L105–L108.
- Béky, B., Holman, M. J., Kipping, D. M., and Noyes, R. W. (2014). Stellar Rotation-Planetary Orbit Period Commensurability in the HAT-P-11 System. *Astrophysical Journal*, 788:1.
- Bodenheimer, P. (1965). Studies in Stellar Evolution. II. Lithium Depletion during the Pre-Main Contraction. *Astrophysical Journal*, 142:451.
- Borucki, W. J., Koch, D., Basri, G., Batalha, N., Brown, T., Caldwell, D., Caldwell, J., Christensen-Dalsgaard, J., Cochran, W. D., DeVore, E., Dunham, E. W., Dupree, A. K., Gautier, T. N., Geary, J. C., Gilliland, R., Gould, A., Howell, S. B., Jenkins, J. M., Kondo, Y., Latham, D. W., Marcy, G. W., Meibom, S., Kjeldsen, H., Lissauer, J. J., Monet, D. G., Morrison, D., Sasselov, D., Tarter, J., Boss, A., Brownlee, D., Owen, T., Buzasi, D., Charbonneau, D., Doyle, L., Fortney, J., Ford, E. B., Holman, M. J., Seager, S., Steffen, J. H., Welsh, W. F., Rowe, J., Anderson, H., Buchhave, L., Ciardi, D., Walkowicz, L., Sherry, W., Horch, E., Isaacson, H., Everett, M. E., Fischer, D., Torres, G., Johnson, J. A., Endl, M., MacQueen, P., Bryson, S. T., Dotson, J., Haas, M., Kolodziejczak, J., Van Cleve, J., Chandrasekaran, H., Twicken, J. D., Quintana, E. V., Clarke, B. D., Allen, C., Li, J., Wu, H., Tenenbaum, P., Verner, E., Bruhweiler, F., Barnes, J., and Prsa, A. (2010). Kepler Planet-Detection Mission: Introduction and First Results. *Science*, 327:977.
- Bouvier, A. and Wadhwa, M. (2010). The age of the Solar System redefined by the oldest Pb-Pb age of a meteoritic inclusion. *Nature Geoscience*, 3:637–641.
- Bressan, A., Marigo, P., Girardi, L., Salasnich, B., Dal Cero, C., Rubele, S., and Nanni, A. (2012). PARSEC: stellar tracks and isochrones with the PAdova and TRieste Stellar Evolution Code. *Monthly Notices of the Royal Astronomical Society*, 427:127–145.
- Brown, T. M. (2000). Progress in Observational Asteroseismology. In Teixeira, T. and Bedding, T., editors, *The Third MONS Workshop: Science Preparation and Target Selection*, page 1.
- Brown, T. M. and Gilliland, R. L. (1994). Asteroseismology. *Anual Review of Astron and Astrophys*, 32:37–82.
- Brown, T. M., Gilliland, R. L., Noyes, R. W., and Ramsey, L. W. (1991). Detection of possible p-mode oscillations on Procyon. *Astrophysical Journal*, 368:599–609.

- Bruntt, H. (2009). Accurate fundamental parameters of CoRoT asteroseismic targets. The solar-like stars HD 49933, HD 175726, HD 181420, and HD 181906. *Astronomy & Astrophysics*, 506:235–244.
- Bruntt, H., Basu, S., Smalley, B., et al. (2012). Accurate fundamental parameters and detailed abundance patterns from spectroscopy of 93 solar-type Kepler targets. *Monthly Notices of the Royal Astronomical Society*, 423:122–131.
- Burke, C. J., Christiansen, J. L., Mullally, F., Seader, S., Huber, D., Rowe, J. F., Coughlin, J. L., Thompson, S. E., Catanzarite, J., Clarke, B. D., Morton, T. D., Caldwell, D. A., Bryson, S. T., Haas, M. R., Batalha, N. M., Jenkins, J. M., Tenenbaum, P., Twicken, J. D., Li, J., Quintana, E., Barclay, T., Henze, C. E., Borucki, W. J., Howell, S. B., and Still, M. (2015). Terrestrial Planet Occurrence Rates for the Kepler GK Dwarf Sample. *Astrophysical Journal*, 809:8.
- Burke, C. J., Pinsonneault, M. H., and Sills, A. (2004). Theoretical Examination of the Lithium Depletion Boundary. *Astrophysical Journal*, 604:272–283.
- Carter, J. A. and Winn, J. N. (2009). Parameter Estimation from Time-series Data with Correlated Errors: A Wavelet-based Method and its Application to Transit Light Curves. *Astrophysical Journal*, 704:51–67.
- Carter, J. A. and Winn, J. N. (2010). Parameter Estimation from Time-Series Data with Correlated Errors: A Wavelet-Based Method and its Application to Transit Light Curves. *Astrophysics Source Code Library*.
- Casagrande, L., Ramírez, I., Meléndez, J., et al. (2010). An absolutely calibrated  $T_{eff}$  scale from the infrared flux method. Dwarfs and subgiants. *Astronomy & Astrophysics*, 512:A54.
- Cassan, A., Kubas, D., Beaulieu, J.-P., Dominik, M., Horne, K., Greenhill, J., Wambsganss, J., Menzies, J., Williams, A., Jørgensen, U. G., Udalski, A., Bennett, D. P., Albrow, M. D., Batista, V., Brillant, S., Caldwell, J. A. R., Cole, A., Coutures, C., Cook, K. H., Dieters, S., Prester, D. D., Donatowicz, J., Fouqué, P., Hill, K., Kains, N., Kane, S., Marquette, J.-B., Martin, R., Pollard, K. R., Sahu, K. C., Vinter, C., Warren, D., Watson, B., Zub, M., Sumi, T., Szymański, M. K., Kubiak, M., Poleski, R., Soszynski, I., Ulaczyk, K., Pietrzyński, G., and Wyrzykowski, Ł. (2012). One or more bound planets per Milky Way star from microlensing observations. *Nature*, 481:167–169.
- Cegla, H. M., Shelyag, S., Watson, C. A., and Mathioudakis, M. (2013). Stellar Surface Magneto-convection as a Source of Astrophysical Noise. I. Multi-component Parameterization of Absorption Line Profiles. *Astrophysical Journal*, 763:95.
- Chaplin, W. J., Basu, S., Huber, D., Serenelli, A., Casagrande, L., Silva Aguirre, V., Ball, W. H., Creevey, O. L., Gizon, L., Handberg, R., Karoff, C., Lutz, R., Marques, J. P., Miglio, A., Stello, D., Suran, M. D., Pricopi, D., Metcalfe, T. S.,

Monteiro, M. J. P. F. G., Molenda-Żakowicz, J., Appourchaux, T., Christensen-Dalsgaard, J., Elsworth, Y., García, R. A., Houdek, G., Kjeldsen, H., Bonanno, A., Campante, T. L., Corsaro, E., Gaulme, P., Hekker, S., Mathur, S., Mosser, B., Régulo, C., and Salabert, D. (2014). Asteroseismic Fundamental Properties of Solar-type Stars Observed by the NASA Kepler Mission. *Astrophysical Journal, Supplements*, 210:1.

Chaplin, W. J., Kjeldsen, H., Bedding, T. R., Christensen-Dalsgaard, J., Gilliland, R. L., Kawaler, S. D., Appourchaux, T., Elsworth, Y., García, R. A., Houdek, G., Karoff, C., Metcalfe, T. S., Molenda-Żakowicz, J., Monteiro, M. J. P. F. G., Thompson, M. J., Verner, G. A., Batalha, N., Borucki, W. J., Brown, T. M., Bryson, S. T., Christiansen, J. L., Clarke, B. D., Jenkins, J. M., Klaus, T. C., Koch, D., An, D., Ballot, J., Basu, S., Benomar, O., Bonanno, A., Broomhall, A.-M., Campante, T. L., Corsaro, E., Creevey, O. L., Esch, L., Gai, N., Gaulme, P., Hale, S. J., Handberg, R., Hekker, S., Huber, D., Mathur, S., Mosser, B., New, R., Pinsonneault, M. H., Pricopi, D., Quirion, P.-O., Régulo, C., Roxburgh, I. W., Salabert, D., Stello, D., and Suran, M. D. (2011). Predicting the Detectability of Oscillations in Solar-type Stars Observed by Kepler. *Astrophysical Journal*, 732:54.

Chaplin, W. J. and Miglio, A. (2013). Asteroseismology of Solar-Type and Red-Giant Stars. *Anual Review of Astron and Astrophys*, 51:353–392.

Charbonneau, D., Brown, T. M., Latham, D. W., and Mayor, M. (2000). Detection of Planetary Transits Across a Sun-like Star. *Astrophysical Journal, Letters*, 529:L45–L48.

Charbonneau, P. (2010). Dynamo Models of the Solar Cycle. *Living Reviews in Solar Physics*, 7.

Christiansen, J. L., Jenkins, J. M., Caldwell, D. A., et al. (2012). The Derivation, Properties, and Value of Kepler’s Combined Differential Photometric Precision. *Publications of the A. S. P.*, 124:1279–1287.

Clanton, C. and Gaudi, B. S. (2016). Synthesizing Exoplanet Demographics: A Single Population of Long-period Planetary Companions to M Dwarfs Consistent with Microlensing, Radial Velocity, and Direct Imaging Surveys. *Astrophysical Journal*, 819:125.

Coelho, H. R., Chaplin, W. J., Basu, S., Serenelli, A., Miglio, A., and Reese, D. R. (2015). A test of the asteroseismic  $\nu_{max}$  scaling relation for solar-like oscillations in main-sequence and subgiant stars. *Monthly Notices of the Royal Astronomical Society*, 451:3011–3020.

Collier Cameron, A., Davidson, V. A., Hebb, L., et al. (2009). The main-sequence rotation-colour relation in the Coma Berenices open cluster. *Monthly Notices of the Royal Astronomical Society*, 400:451–462.

- Collier Cameron, A. and Jianke, L. (1994). Magnetic Braking of G and K Dwarfs Without Core Envelope Decoupling. *Monthly Notices of the Royal Astronomical Society*, 269:1099.
- Cox, A. N. and Pilachowski, C. A. (2000). Allen's Astrophysical Quantities. *Physics Today*, 53(10):77.
- Cox, J. P. (1980). *Theory of stellar pulsation*.
- Cranmer, S. R., Bastien, F. A., Stassun, K. G., and Saar, S. H. (2014). Stellar Granulation as the Source of High-frequency Flicker in Kepler Light Curves. *Astrophysical Journal*, 781:124.
- Crossfield, I. J. M., Petigura, E., Schlieder, J. E., Howard, A. W., Fulton, B. J., Aller, K. M., Ciardi, D. R., Lépine, S., Barclay, T., de Pater, I., de Kleer, K., Quintana, E. V., Christiansen, J. L., Schlafly, E., Kaltenegger, L., Crepp, J. R., Henning, T., Obermeier, C., Deacon, N., Weiss, L. M., Isaacson, H. T., Hansen, B. M. S., Liu, M. C., Greene, T., Howell, S. B., Barman, T., and Mordasini, C. (2015). A Nearby M Star with Three Transiting Super-Earths Discovered by K2. *Astrophysical Journal*, 804:10.
- Davenport, J. R. A. (2015). The Shape of M Dwarf Flares in Kepler Light Curves. *ArXiv e-prints*.
- Davies, G. R., Chaplin, W. J., Farr, W. M., et al. (2014). Asteroseismic inference on rotation, gyrochronology and planetary system dynamics of 16 Cygni. *ArXiv e-prints*.
- Delorme, P., Cameron, A. C., Hebb, L., Rostron, J., Lister, T. A., Norton, A. J., Pollacco, D., and West, R. G. (2011a). Stellar Rotation in the Hyades and Praesepe: Gyrochronology and Braking Timescale. In Johns-Krull, C., Browning, M. K., and West, A. A., editors, *16th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*, volume 448 of *Astronomical Society of the Pacific Conference Series*, page 841.
- Delorme, P., Cameron, A. C., Hebb, L., Rostron, J., Lister, T. A., Norton, A. J., Pollacco, D., and West, R. G. (2011b). VizieR Online Data Catalog: Stellar rotation in Hyades and Praesepe (Delorme+, 2011). *VizieR Online Data Catalog*, 741.
- Demarque, P., Woo, J.-H., Kim, Y.-C., et al. (2004).  $Y^2$  Isochrones with an Improved Core Overshoot Treatment. *Astrophysical Journal Supplements*, 155:667–674.
- Demory, B.-O. (2014). The Albedos of Kepler's Close-in Super-Earths. *Astrophysical Journal Letters*, 789:L20.
- Dobbie, P. D., Casewell, S. L., Burleigh, M. R., et al. (2009). A massive white dwarf member of the Coma Berenices open cluster. *Monthly Notices of the Royal Astronomical Society*, 395:1591–1598.

- Donahue, R. A. (1993). *Surface differential rotation in a sample of cool dwarf stars*. PhD thesis, New Mexico State University, University Park.
- Donahue, R. A., Saar, S. H., and Baliunas, S. L. (1996). A Relationship between Mean Rotation Period in Lower Main-Sequence Stars and Its Observed Range. *Astrophysical Journal*, 466:384.
- Dotter, A., Chaboyer, B., Jevremović, D., Kostov, V., Baron, E., and Ferguson, J. W. (2008). The Dartmouth Stellar Evolution Database. *Astrophysical Journal, Supplements*, 178:89–101.
- Dressing, C. D. and Charbonneau, D. (2015). The Occurrence of Potentially Habitable Planets Orbiting M Dwarfs Estimated from the Full Kepler Dataset and an Empirical Measurement of the Detection Sensitivity. *Astrophysical Journal*, 807:45.
- Dumusque, X., Boisse, I., and Santos, N. C. (2014). SOAP 2.0: A Tool to Estimate the Photometric and Radial Velocity Variations Induced by Stellar Spots and Plages. *Astrophysical Journal*, 796:132.
- Dumusque, X., Pepe, F., Lovis, C., et al. (2012). An Earth-mass planet orbiting  $\alpha$  Centauri B. *Nature*, 491:207–211.
- Edelson, R. A. and Krolik, J. H. (1988). The discrete correlation function - A new method for analyzing unevenly sampled variability data. *Astrophysical Journal*, 333:646–659.
- Epstein, C. R. and Pinsonneault, M. H. (2014). How Good a Clock is Rotation? The Stellar Rotation-Mass-Age Relationship for Old Field Stars. *Astrophysical Journal*, 780:159.
- Evans, T. M., Aigrain, S., Gibson, N., Barstow, J. K., Amundsen, D. S., Tremblin, P., and Mourier, P. (2015). A uniform analysis of HD 209458b Spitzer/IRAC light curves with Gaussian process models. *Monthly Notices of the Royal Astronomical Society*, 451:680–694.
- Foreman-Mackey, D., Hogg, D. W., Lang, D., et al. (2013). emcee: The MCMC Hammer. *Publications of the A. S. P.*, 125:306–312.
- Foreman-Mackey, D., Hogg, D. W., and Morton, T. D. (2014a). Exoplanet Population Inference and the Abundance of Earth Analogs from Noisy, Incomplete Catalogs. *Astrophysical Journal*, 795:64.
- Foreman-Mackey, D., Hoyer, S., Bernhard, J., and Angus, R. (2014b). Fast Gaussian Processes for regression.
- Foreman-Mackey, D., Montet, B. T., Hogg, D. W., Morton, T. D., Wang, D., and Schölkopf, B. (2015). A Systematic Search for Transiting Planets in the K2 Data. *Astrophysical Journal*, 806:215.

- Foreman-Mackey, D., Price-Whelan, A., Ryan, G., et al. (2014). A citable release.
- Funk, B., Wuchterl, G., Schwarz, R., Pilat-Lohinger, E., and Eggl, S. (2010). The stability of ultra-compact planetary systems. *Astronomy & Astrophysics*, 516:A82.
- García, R. A., Ceillier, T., Salabert, D., Mathur, S., van Saders, J. L., Pinsonneault, M., Ballot, J., Beck, P. G., Bloemen, S., Campante, T. L., Davies, G. R., do Nascimento, Jr., J.-D., Mathis, S., Metcalfe, T. S., Nielsen, M. B., Suárez, J. C., Chaplin, W. J., Jiménez, A., and Karoff, C. (2014). Rotation and magnetism of Kepler pulsating solar-like stars. Towards asteroseismically calibrated age-rotation relations. *Astronomy & Astrophysics*, 572:A34.
- García, R. A., Hekker, S., Stello, D., Gutiérrez-Soto, J., Handberg, R., Huber, D., Karoff, C., Uytterhoeven, K., Appourchaux, T., Chaplin, W. J., Elsworth, Y., Mathur, S., Ballot, J., Christensen-Dalsgaard, J., Gilliland, R. L., Houdek, G., Jenkins, J. M., Kjeldsen, H., McCauliff, S., Metcalfe, T., Middour, C. K., Molenda-Zakowicz, J., Monteiro, M. J. P. F. G., Smith, J. C., and Thompson, M. J. (2011). Preparation of Kepler light curves for asteroseismic analyses. *Monthly Notices of the Royal Astronomical Society*, 414:L6–L10.
- Gaudi, B. S. (2012). Microlensing Surveys for Exoplanets. *Anual Review of Astron and Astrophys*, 50:411–453.
- Gelman, A. (2009). Bayes, Jeffreys, Prior Distributions and the Philosophy of Statistics. *Statistical Science*, 24:176–178.
- Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7:457–472.
- Gibson, N. P., Aigrain, S., Roberts, S., Evans, T. M., Osborne, M., and Pont, F. (2012). A Gaussian process framework for modelling instrumental systematics: application to transmission spectroscopy. *Monthly Notices of the Royal Astronomical Society*, 419:2683–2694.
- Girardi, L., Barbieri, M., Groenewegen, M. A. T., Marigo, P., Bressan, A., Rocha-Pinto, H. J., Santiago, B. X., Camargo, J. I. B., and da Costa, L. N. (2012). TRILEGAL, a TRIdimensional modeL of thE GALaxy: Status and Future. *Astrophysics and Space Science Proceedings*, 26:165.
- Girardi, L., Bertelli, G., Bressan, A., Chiosi, C., Groenewegen, M. A. T., Marigo, P., Salasnich, B., and Weiss, A. (2002). Theoretical isochrones in several photometric systems. I. Johnson-Cousins-Glass, HST/WFPC2, HST/NICMOS, Washington, and ESO Imaging Survey filter sets. *Astronomy & Astrophysics*, 391:195–212.
- Gould, A., Dong, S., Gaudi, B. S., Udalski, A., Bond, I. A., Greenhill, J., Street, R. A., Dominik, M., Sumi, T., Szymański, M. K., Han, C., Allen, W., Bolt, G., Bos, M., Christie, G. W., DePoy, D. L., Drummond, J., Eastman, J. D., Gal-Yam, A., Higgins, D., Janczak, J., Kaspi, S., Kozłowski, S., Lee, C.-U., Mallia, F.,

Maury, A., Maoz, D., McCormick, J., Monard, L. A. G., Moorhouse, D., Morgan, N., Natusch, T., Ofek, E. O., Park, B.-G., Pogge, R. W., Polishook, D., Santallo, R., Shporer, A., Spector, O., Thornley, G., Yee, J. C.,  $\mu$ FUN Collaboration, Kubiak, M., Pietrzyński, G., Soszyński, I., Szewczyk, O., Wyrzykowski, Ł., Ulaczyk, K., Poleski, R., OGLE Collaboration, Abe, F., Bennett, D. P., Botzler, C. S., Douchin, D., Freeman, M., Fukui, A., Furusawa, K., Hearnshaw, J. B., Hosaka, S., Itow, Y., Kamiya, K., Kilmarin, P. M., Korpela, A., Lin, W., Ling, C. H., Makita, S., Masuda, K., Matsubara, Y., Miyake, N., Muraki, Y., Nagaya, M., Nishimoto, K., Ohnishi, K., Okumura, T., Perrott, Y. C., Philpott, L., Rattenbury, N., Saito, T., Sako, T., Sullivan, D. J., Sweatman, W. L., Tristram, P. J., von Seggern, E., Yock, P. C. M., MOA Collaboration, Albrow, M., Batista, V., Beaulieu, J. P., Brillant, S., Caldwell, J., Calitz, J. J., Cassan, A., Cole, A., Cook, K., Coutures, C., Dieters, S., Dominis Prester, D., Donatowicz, J., Fouqué, P., Hill, K., Hoffman, M., Jablonski, F., Kane, S. R., Kains, N., Kubas, D., Marquette, J.-B., Martin, R., Martioli, E., Meintjes, P., Menzies, J., Pedretti, E., Pollard, K., Sahu, K. C., Vinter, C., Wambsganss, J., Watson, R., Williams, A., Zub, M., PLANET Collaboration, Allan, A., Bode, M. F., Bramich, D. M., Burgdorf, M. J., Clay, N., Fraser, S., Hawkins, E., Horne, K., Kerins, E., Lister, T. A., Mottram, C., Saunders, E. S., Snodgrass, C., Steele, I. A., Tsapras, Y., RoboNet Collaboration, Jørgensen, U. G., Anguita, T., Bozza, V., Calchi Novati, S., Harpsøe, K., Hinse, T. C., Hundertmark, M., Kjærgaard, P., Liebig, C., Mancini, L., Masi, G., Mathiasen, M., Rahvar, S., Ricci, D., Scarpetta, G., Southworth, J., Surdej, J., Thöne, C. C., and MiNDSTEp Consortium (2010). Frequency of Solar-like Systems and of Ice and Gas Giants Beyond the Snow Line from High-magnification Microlensing Events in 2005-2008. *Astrophysical Journal*, 720:1073–1089.

Hallam, K. L., Altner, B., and Endal, A. S. (1991). Rotational modulation in IUE spectra of cool dwarf chromospheres. I - Data analysis and period search techniques. *Astrophysical Journal*, 372:610–632.

Hawley, S. L., Davenport, J. R. A., Kowalski, A. F., Wisniewski, J. P., Hebb, L., Deitrick, R., and Hilton, E. J. (2014). Kepler Flares. I. Active and Inactive M Dwarfs. *Astrophysical Journal*, 797:121.

Haywood, R. D. (2015). *Hide and Seek: Radial-Velocity Searches for Planets around Active Stars*. PhD thesis, University of St Andrews.

Haywood, R. D., Collier Cameron, A., Queloz, D., Barros, S. C. C., Deleuil, M., Fares, R., Gillon, M., Lanza, A. F., Lovis, C., Moutou, C., Pepe, F., Pollacco, D., Santerne, A., Ségransan, D., and Unruh, Y. C. (2014). Planets and stellar activity: hide and seek in the CoRoT-7 system. *Monthly Notices of the Royal Astronomical Society*, 443:2517–2531.

Hogg, D. W., Angus, R., Barclay, T., Dawson, R., Fergus, R., Foreman-Mackey, D., Harmeling, S., Hirsch, M., Lang, D., Montet, B. T., Schiminovich, D., and Schölkopf, B. (2013). Maximizing Kepler science return per telemetered pixel: Detailed models of the focal plane in the two-wheel era. *ArXiv e-prints*.

- Hogg, D. W., Bovy, J., and Lang, D. (2010a). Data analysis recipes: Fitting a model to data. *ArXiv e-prints*.
- Hogg, D. W., Myers, A. D., and Bovy, J. (2010b). Inferring the Eccentricity Distribution. *Astrophysical Journal*, 725:2166–2175.
- Howell, S. B., Sobeck, C., Haas, M., Still, M., Barclay, T., Mullally, F., Troeltzsch, J., Aigrain, S., Bryson, S. T., Caldwell, D., Chaplin, W. J., Cochran, W. D., Huber, D., Marcy, G. W., Miglio, A., Najita, J. R., Smith, M., Twicken, J. D., and Fortney, J. J. (2014). The K2 Mission: Characterization and Early Results. *Publications of the A. S. P.*, 126:398–408.
- Huber, D., Chaplin, W. J., Christensen-Dalsgaard, J., Gilliland, R. L., Kjeldsen, H., Buchhave, L. A., Fischer, D. A., Lissauer, J. J., Rowe, J. F., Sanchis-Ojeda, R., Basu, S., Handberg, R., Hekker, S., Howard, A. W., Isaacson, H., Karoff, C., Latham, D. W., Lund, M. N., Lundkvist, M., Marcy, G. W., Miglio, A., Silva Aguirre, V., Stello, D., Arentoft, T., Barclay, T., Bedding, T. R., Burke, C. J., Christiansen, J. L., Elsworth, Y. P., Haas, M. R., Kawaler, S. D., Metcalfe, T. S., Mullally, F., and Thompson, S. E. (2013). Fundamental Properties of Kepler Planet-candidate Host Stars using Asteroseismology. *Astrophysical Journal*, 767:127.
- Huber, D., Stello, D., Bedding, T. R., et al. (2009). Automated extraction of oscillation parameters for Kepler observations of solar-type stars. *Communications in Asteroseismology*, 160:74.
- Irwin, J. and Bouvier, J. (2009). The rotational evolution of low-mass stars. In Mamajek, E. E., Soderblom, D. R., and Wyse, R. F. G., editors, *IAU Symposium*, volume 258 of *IAU Symposium*, pages 363–374.
- Ivezic, Z., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., AlSayyad, Y., Anderson, S. F., Andrew, J., Angel, R., Angeli, G., Ansari, R., Antilogus, P., Arndt, K. T., Astier, P., Aubourg, E., Axelrod, T., Bard, D. J., Barr, J. D., Barrau, A., Bartlett, J. G., Bauman, B. J., Beaumont, S., Becker, A. C., Becla, J., Beldica, C., Bellavia, S., Blanc, G., Blandford, R. D., Bloom, J. S., Bogart, J., Borne, K., Bosch, J. F., Boutigny, D., Brandt, W. N., Brown, M. E., Bullock, J. S., Burchat, P., Burke, D. L., Cagnoli, G., Calabrese, D., Chandrasekharan, S., Chesley, S., Cheu, E. C., Chiang, J., Claver, C. F., Connolly, A. J., Cook, K. H., Cooray, A., Covey, K. R., Cribbs, C., Cui, W., Cutri, R., Daubard, G., Daves, G., Delgado, F., Digel, S., Doherty, P., Dubois, R., Dubois-Felsmann, G. P., Durech, J., Eracleous, M., Ferguson, H., Frank, J., Freeman, M., Gangler, E., Gawiser, E., Geary, J. C., Gee, P., Geha, M., Gibson, R. R., Gilmore, D. K., Glanzman, T., Goodenow, I., Gressler, W. J., Gris, P., Guyonnet, A., Hascall, P. A., Haupt, J., Hernandez, F., Hogan, C., Huang, D., Huffer, M. E., Innes, W. R., Jacoby, S. H., Jain, B., Jee, J., Jernigan, J. G., Jevremovic, D., Johns, K., Jones, R. L., Juramy-Gilles, C., Juric, M., Kahn, S. M., Kalirai, J. S., Kallivayalil, N., Kalmbach, B., Kantor, J. P., Kasliwal, M. M., Kessler, R., Kirkby, D., Knox, L., Kotov, I., Krabbendam, V. L., Krughoff, S.,

Kubanek, P., Kuczewski, J., Kulkarni, S., Lambert, R., Le Guillou, L., Levine, D., Liang, M., Lim, K., Lintott, C., Lupton, R. H., Mahabal, A., Marshall, P., Marshall, S., May, M., McKercher, R., Migliore, M., Miller, M., Mills, D. J., Monet, D. G., Moniez, M., Neill, D. R., Nief, J., Nomerotski, A., Nordby, M., O'Connor, P., Oliver, J., Olivier, S. S., Olsen, K., Ortiz, S., Owen, R. E., Pain, R., Peterson, J. R., Petry, C. E., Pierfederici, F., Pietrowicz, S., Pike, R., Pinto, P. A., Plante, R., Plate, S., Price, P. A., Prouza, M., Radeka, V., Rajagopal, J., Rasmussen, A., Regnault, N., Ridgway, S. T., Ritz, S., Rosing, W., Roucelle, C., Rumore, M. R., Russo, S., Saha, A., Sassolas, B., Schalk, T. L., Schindler, R. H., Schneider, D. P., Schumacher, G., Sebag, J., Sembroski, G. H., Seppala, L. G., Shipsey, I., Silvestri, N., Smith, J. A., Smith, R. C., Strauss, M. A., Stubbs, C. W., Sweeney, D., Szalay, A., Takacs, P., Thaler, J. J., Van Berg, R., Vanden Berk, D., Vetter, K., Virieux, F., Xin, B., Walkowicz, L., Walter, C. W., Wang, D. L., Warner, M., Willman, B., Wittman, D., Wolff, S. C., Wood-Vasey, W. M., Yoachim, P., Zhan, H., and for the LSST Collaboration (2008). LSST: from Science Drivers to Reference Design and Anticipated Data Products. *ArXiv e-prints*.

Jacklin, S., Lund, M. B., Pepper, J., and Stassun, K. G. (2015). Transiting Planets with LSST. II. Period Detection of Planets Orbiting 1 M Hosts. *Astronomical Journal*, 150:34.

Jeffers, S. V. and Keller, C. U. (2009). An analytical model to demonstrate the reliability of reconstructed ‘active longitudes’. In Stempels, E., editor, *15th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun*, volume 1094 of *American Institute of Physics Conference Series*, pages 664–667.

Kalas, P., Graham, J. R., Chiang, E., Fitzgerald, M. P., Clampin, M., Kite, E. S., Stapelfeldt, K., Marois, C., and Krist, J. (2008). Optical Images of an Exosolar Planet 25 Light-Years from Earth. *Science*, 322:1345.

Kallinger, T. and Matthews, J. M. (2010). Evidence for Granulation in Early A-Type Stars. *Astrophysical Journal, Letters*, 711:L35–L39.

Kass, R. E. and Wasserman, L. (1996). The Selection of Prior Distributions by Formal Rules. *Journal of the American Statistical Association*, 91:1343–1370.

Kawaler, S. D. (1988). Angular momentum loss in low-mass stars. *Astrophysical Journal*, 333:236–247.

Kawaler, S. D. (1989). Rotational dating of middle-aged stars. *Astrophysical Journal, Letters*, 343:L65–L68.

Kipping, D. M., Bastien, F. A., Stassun, K. G., Chaplin, W. J., Huber, D., and Buchhave, L. A. (2014). Flicker as a Tool for Characterizing Planets Through Asterodensity Profiling. *Astrophysical Journal, Letters*, 785:L32.

Kjeldsen, H. and Bedding, T. R. (2001). Current status of asteroseismology. In Wilson, A. and Pallé, P. L., editors, *SOHO 10/GONG 2000 Workshop: Helio-*

*and Asteroseismology at the Dawn of the Millennium*, volume 464 of *ESA Special Publication*, pages 361–366.

Kjeldsen, H., Bedding, T. R., and Christensen-Dalsgaard, J. (2008). Correcting Stellar Oscillation Frequencies for Near-Surface Effects. *Astrophysical Journal, Letters*, 683:L175.

Kjeldsen, H., Bedding, T. R., Frandsen, S., and dall, T. H. (1999). A search for solar-like oscillations and granulation in alpha CEN A. *Monthly Notices of the Royal Astronomical Society*, 303:579–587.

Kjeldsen, H., Bedding, T. R., Viskum, M., and Frandsen, S. (1995). Solarlike oscillations in eta Boo. *Astronomical Journal*, 109:1313–1319.

Kraft, R. P. (1967). Studies of Stellar Rotation. V. The Dependence of Rotation on Age among Solar-Type Stars. *Astrophysical Journal*, 150:551.

Krishnamurthi, A., Terndrup, D. M., Pinsonneault, M. H., et al. (1998). New Rotation Periods in the Pleiades: Interpreting Activity Indicators. *Astrophysical Journal*, 493:914–925.

Lachaume, R., Dominik, C., Lanz, T., and Habing, H. J. (1999). Age determinations of main-sequence stars: combining different methods. *Astronomy & Astrophysics*, 348:897–909.

Lanza, A. F., Das Chagas, M. L., and De Medeiros, J. R. (2014). Measuring stellar differential rotation with high-precision space-borne photometry. *Astronomy & Astrophysics*, 564:A50.

Lebreton, Y. and Goupil, M. J. (2014). Asteroseismology for "à la carte" stellar age-dating and weighing. Age and mass of the CoRoT exoplanet host HD 52265. *Astronomy & Astrophysics*, 569:A21.

Li, T. D., Bi, S. L., Liu, K., et al. (2012). Stellar parameters and seismological analysis of the star 18 Scorpii. *Astronomy & Astrophysics*, 546:A83.

Lomb, N. R. (1976). Least-squares frequency analysis of unequally spaced data. *Astrophysics and Space Science*, 39:447–462.

Lund, M. N., Handberg, R., Davies, G. R., Chaplin, W. J., and Jones, C. D. (2015). K2P\$^2\\$ \$-\$ A photometry pipeline for the K2 mission. *ArXiv e-prints*.

Mamajek, E. E. and Hillenbrand, L. A. (2008). Improved Age Estimation for Solar-Type Dwarfs Using Activity-Rotation Diagnostics. *Astrophysical Journal*, 687:1264–1293.

Martić, M., Lebrun, J. C., Schmitt, J., Appourchaux, T., and Bertaux, J. L. (2001). Observing solar-like oscillations:  $\alpha$  CMi,  $\eta$  Cas A and  $\zeta$  Her A. In Wilson, A. and

Pallé, P. L., editors, *SOHO 10/GONG 2000 Workshop: Helio- and Asteroseismology at the Dawn of the Millennium*, volume 464 of *ESA Special Publication*, pages 431–434.

Martić, M., Schmitt, J., Lebrun, J.-C., Barban, C., Connes, P., Bouchy, F., Michel, E., Baglin, A., Appourchaux, T., and Bertaux, J.-L. (1999). Evidence for global pressure oscillations on Procyon. *Astronomy & Astrophysics*, 351:993–1002.

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

Mathur, S., Metcalfe, T. S., Woitaszek, M., et al. (2012). A Uniform Asteroseismic Analysis of 22 Solar-type Stars Observed by Kepler. *Astrophysical Journal*, 749:152.

Mayor, M. and Queloz, D. (1995). A Jupiter-mass companion to a solar-type star. *Nature*, 378:355–359.

Mazeh, T., Perets, H. B., McQuillan, A., and Goldstein, E. S. (2015). Photometric Amplitude Distribution of Stellar Rotation of KOIs—Indication for Spin-Orbit Alignment of Cool Stars and High Obliquity for Hot Stars. *Astrophysical Journal*, 801:3.

McQuillan, A., Aigrain, S., and Mazeh, T. (2013a). Measuring the rotation period distribution of field M dwarfs with Kepler. *Monthly Notices of the Royal Astronomical Society*, 432:1203–1216.

McQuillan, A., Mazeh, T., and Aigrain, S. (2013b). Stellar Rotation Periods of the Kepler Objects of Interest: A Dearth of Close-in Planets around Fast Rotators. *Astrophysical Journal, Letters*, 775:L11.

McQuillan, A., Mazeh, T., and Aigrain, S. (2014). Rotation Periods of 34,030 Kepler Main-sequence Stars: The Full Autocorrelation Sample. *Astrophysical Journal, Supplements*, 211:24.

Meibom, S., Barnes, S. A., Latham, D. W., Batalha, N., Borucki, W. J., Koch, D. G., Basri, G., Walkowicz, L. M., Janes, K. A., Jenkins, J., Van Cleve, J., Haas, M. R., Bryson, S. T., Dupree, A. K., Furesz, G., Szentgyorgyi, A. H., Buchhave, L. A., Clarke, B. D., Twicken, J. D., and Quintana, E. V. (2011a). The Kepler Cluster Study: Stellar Rotation in NGC 6811. *Astrophysical Journal, Letters*, 733:L9.

Meibom, S., Mathieu, R. D., Stassun, K. G., et al. (2011b). The Color-period Diagram and Stellar Rotational Evolution — New Rotation Period Measurements in the Open Cluster M34. *Astrophysical Journal*, 733:115.

- Mermilliod, J.-C. (1986). Compilation of Eggen's UBV data, transformed to UBV (unpublished). *Catalogue of Eggen's UBV data.*, 0 (1986), page 0.
- Mestel, L. (1984). Angular Momentum Loss During Pre-Main Sequence Contraction. In Baliunas, S. L. and Hartmann, L., editors, *Cool Stars, Stellar Systems, and the Sun*, volume 193 of *Lecture Notes in Physics*, Berlin Springer Verlag, page 49.
- Mestel, L. and Spruit, H. C. (1987). On magnetic braking of late-type stars. *Monthly Notices of the Royal Astronomical Society*, 226:57–66.
- Metcalfe, T. S., Chaplin, W. J., Appourchaux, T., et al. (2012). Asteroseismology of the Solar Analogs 16 Cyg A and B from Kepler Observations. *Astrophysical Journal, Letters*, 748:L10.
- Metcalfe, T. S., Creevey, O. L., and Christensen-Dalsgaard, J. (2009). A Stellar Model-fitting Pipeline for Asteroseismic Data from the Kepler Mission. *Astrophysical Journal*, 699:373–382.
- Metcalfe, T. S., Creevey, O. L., Dogan, G., et al. (2014). Properties of 42 Solar-type Kepler Targets from the Asteroseismic Modeling Portal. *ArXiv e-prints*.
- Metcalfe, T. S., Monteiro, M. J. P. F. G., Thompson, M. J., Molenda-Żakowicz, J., Appourchaux, T., Chaplin, W. J., Doğan, G., Eggenberger, P., Bedding, T. R., Bruntt, H., Creevey, O. L., Quirion, P.-O., Stello, D., Bonanno, A., Silva Aguirre, V., Basu, S., Esch, L., Gai, N., Di Mauro, M. P., Kosovichev, A. G., Kitiashvili, I. N., Suárez, J. C., Moya, A., Piau, L., García, R. A., Marques, J. P., Frasca, A., Biazzo, K., Sousa, S. G., Dreizler, S., Bazot, M., Karoff, C., Frandsen, S., Wilson, P. A., Brown, T. M., Christensen-Dalsgaard, J., Gilliland, R. L., Kjeldsen, H., Campante, T. L., Fletcher, S. T., Handberg, R., Régulo, C., Salabert, D., Schou, J., Verner, G. A., Ballot, J., Broomhall, A.-M., Elsworth, Y., Hekker, S., Huber, D., Mathur, S., New, R., Roxburgh, I. W., Sato, K. H., White, T. R., Borucki, W. J., Koch, D. G., and Jenkins, J. M. (2010). A Precise Asteroseismic Age and Radius for the Evolved Sun-like Star KIC 11026764. *Astrophysical Journal*, 723:1583–1598.
- Michel, E., Baglin, A., Weiss, W. W., Auvergne, M., Catala, C., Aerts, C., Appourchaux, T., Barban, C., Baudin, F., Briquet, M.-L., Carrier, F., Debosscher, J., De Ridder, J., Garcia, R. A., Garrido, R., Gutierrez, J., Kallinger, T., Lefevre, L., Neiner, C., Poretti, E., Samadi, R., Sarro, L., and COROT Team (2008). First asteroseismic results from CoRoT. *Communications in Asteroseismology*, 157:69.
- Moffett, T. J. and Barnes, III, T. G. (1979). Secondary standards for BVRI photometry. *Publications of the A. S. P.*, 91:180–192.
- Montet, B. T., Morton, T. D., Foreman-Mackey, D., Johnson, J. A., Hogg, D. W., Bowler, B. P., Latham, D. W., Bieryla, A., and Mann, A. W. (2015). Stellar and Planetary Properties of K2 Campaign 1 Candidates and Validation of 17 Planets, Including a Planet Receiving Earth-like Insolation. *Astrophysical Journal*, 809:25.

- Morton, T. D. and Winn, J. N. (2014). Obliquities of Kepler Stars: Comparison of Single- and Multiple-Transit Systems. *ArXiv e-prints*.
- Mottola, S., Sears, W. D., Erikson, A., Harris, A. W., Young, J. W., Hahn, G., Dahlgren, M., Mueller, B. E. A., Owen, B., Ricardo, G.-H., Licandro, J., Barucci, M. A., Angeli, C., Neukum, G., Lagerkvist, C.-I., and Felix Lahulla, J. (1995). The slow rotation of 253 Mathilde. *Planetary Space Science*, 43:1609–1613.
- Newton, E. R., Irwin, J., Charbonneau, D., Berta-Thompson, Z. K., Dittmann, J. A., and West, A. A. (2015). The rotation and Galactic kinematics of mid M dwarfs in the Solar Neighborhood. *ArXiv e-prints*.
- Nielsen, E. L., Liu, M. C., Wahhaj, Z., Biller, B. A., Hayward, T. L., Close, L. M., Males, J. R., Skemer, A. J., Chun, M., Ftacclas, C., Alencar, S. H. P., Artymowicz, P., Boss, A., Clarke, F., de Gouveia Dal Pino, E., Gregorio-Hetem, J., Hartung, M., Ida, S., Kuchner, M., Lin, D. N. C., Reid, I. N., Shkolnik, E. L., Tecza, M., Thatte, N., and Toomey, D. W. (2013). The Gemini NICI Planet-Finding Campaign: The Frequency of Giant Planets around Young B and A Stars. *Astrophysical Journal*, 776:4.
- Nissen, P. E. and Schuster, W. J. (1991). Uvby-beta photometry of high-velocity and metal-poor stars. V - Distances, kinematics and ages of halo and disk stars. *Astronomy & Astrophysics*, 251:457–468.
- Noyes, R. W., Hartmann, L. W., Baliunas, S. L., Duncan, D. K., and Vaughan, A. H. (1984a). Rotation, convection, and magnetic activity in lower main-sequence stars. *Astrophysical Journal*, 279:763–777.
- Noyes, R. W., Weiss, N. O., and Vaughan, A. H. (1984b). The relation between stellar rotation rate and activity cycle periods. *Astrophysical Journal*, 287:769–773.
- Pallavicini, R., Golub, L., Rosner, R., Vaiana, G. S., Ayres, T., and Linsky, J. L. (1981). Relations among stellar X-ray emission observed from Einstein, stellar rotation and bolometric luminosity. *Astrophysical Journal*, 248:279–290.
- Parker, E. N. (1970). The Generation of Magnetic Fields in Astrophysical Bodies. I. The Dynamo Equations. *Astrophysical Journal*, 162:665.
- Paulson, D. B., Cochran, W. D., and Hatzes, A. P. (2004). Searching for Planets in the Hyades. V. Limits on Planet Detection in the Presence of Stellar Activity. *Astronomical Journal*, 127:3579–3586.
- Perryman, M. A. C., Brown, A. G. A., Lebreton, Y., et al. (1998). The Hyades: distance, structure, dynamics, and age. *Astronomy & Astrophysics*, 331:81–120.
- Petigura, E. A., Howard, A. W., and Marcy, G. W. (2013). Prevalence of Earth-size planets orbiting Sun-like stars. *Proceedings of the National Academy of Science*, 110:19273–19278.

- Petit, P., Dintrans, B., Solanki, S. K., et al. (2008). Toroidal versus poloidal magnetic fields in Sun-like stars: a rotation threshold. *Monthly Notices of the Royal Astronomical Society*, 388:80–88.
- Pinsonneault, M. H., An, D., Molenda-Zakowicz, J., et al. (2012). Effective temperature scale for KIC stars (Pinsonneault+, 2012). *VizieR Online Data Catalog*, 219:90030.
- Poppenhaeger, K. and Wolk, S. J. (2014). Indications for an influence of hot Jupiters on the rotation and activity of their host stars. *Astronomy & Astrophysics*, 565:L1.
- Prosser, C. F., Shetrone, M. D., Dasgupta, A., et al. (1995). Rotation periods of open-cluster stars, 3. *Publications of the A. S. P.*, 107:211–218.
- Pu, B. and Wu, Y. (2015). Spacing of Kepler Planets: Sculpting by Dynamical Instability. *Astrophysical Journal*, 807:44.
- Radick, R. R., Lockwood, G. W., Skiff, B. A., and Thompson, D. T. (1995). A 12 Year Photometric Study of Lower Main-Sequence Hyades Stars. *Astrophysical Journal*, 452:332.
- Radick, R. R., Thompson, D. T., Lockwood, G. W., et al. (1987). The activity, variability, and rotation of lower main-sequence Hyades stars. *Astrophysical Journal*, 321:459–472.
- Rajpaul, V., Aigrain, S., Osborne, M. A., Reece, S., and Roberts, S. (2015). A Gaussian process framework for modelling stellar activity signals in radial velocity data. *Monthly Notices of the Royal Astronomical Society*, 452:2269–2291.
- Rajpaul, V., Aigrain, S., and Roberts, S. (2016). Ghost in the time series: no planet for Alpha Cen B. *Monthly Notices of the Royal Astronomical Society*, 456:L6–L10.
- Rasmussen, C. E. and Williams, C. K. I. (2005). *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press.
- Reiners, A. and Mohanty, S. (2012). Radius-dependent Angular Momentum Evolution in Low-mass Stars. I. *Astrophysical Journal*, 746:43.
- Reiners, A., Schüssler, M., and Passegger, V. M. (2014). Generalized Investigation of the Rotation-Activity Relation: Favoring Rotation Period instead of Rossby Number. *Astrophysical Journal*, 794:144.
- Reinhold, T. and Reiners, A. (2013). Fast and reliable method for measuring stellar differential rotation from photometric data. *Astronomy & Astrophysics*, 557:A11.
- Reinhold, T., Reiners, A., and Basri, G. (2013). Rotation and differential rotation of active Kepler stars. *Astronomy & Astrophysics*, 560:A4.

- Ridgway, S. T., Chandrasekharan, S., Cook, K. H., Jones, R. L., Krughoff, K. S., Petry, C., and Ivezić, Ž. (2012). Measuring the effectiveness of simulated LSST observing programs. In *Observatory Operations: Strategies, Processes, and Systems IV*, volume 8448 of *Proceedings of the SPIE*, page 844810.
- Roberts, S., McQuillan, A., Reece, S., and Aigrain, S. (2013). Astrophysically robust systematics removal using variational inference: application to the first month of Kepler data. *Monthly Notices of the Royal Astronomical Society*, 435:3639–3653.
- Rogers, L. A. (2014). Most 1.6 Earth-Radius Planets are not Rocky. *ArXiv e-prints*.
- Rogers, L. A. (2015). Most 1.6 Earth-radius Planets are Not Rocky. *Astrophysical Journal*, 801:41.
- Russell, H. N. (1906). On the light variations of asteroids and satellites. *Astrophysical Journal*, 24:1–18.
- Saar, S. H. and Donahue, R. A. (1997). Activity-Related Radial Velocity Variation in Cool Stars. *Astrophysical Journal*, 485:319–327.
- Samadi, R., Belkacem, K., and Ludwig, H.-G. (2013a). Stellar granulation as seen in disk-integrated intensity. I. Simplified theoretical modeling. *Astronomy & Astrophysics*, 559:A39.
- Samadi, R., Belkacem, K., Ludwig, H.-G., Caffau, E., Campante, T. L., Davies, G. R., Kallinger, T., Lund, M. N., Mosser, B., Baglin, A., Mathur, S., and Garcia, R. A. (2013b). Stellar granulation as seen in disk-integrated intensity. II. Theoretical scaling relations compared with observations. *Astronomy & Astrophysics*, 559:A40.
- Sanchis-Ojeda, R., Rappaport, S., Pallé, E., et al. (2015). The K2-ESPRINT Project I: Discovery of the Disintegrating Rocky Planet with a Cometary Head and Tail EPIC 201637175b. *ArXiv e-prints*.
- Sanchis-Ojeda, R., Rappaport, S., Winn, J. N., et al. (2014). A Study of the Shortest-period Planets Found with Kepler. *Astrophysical Journal*, 787:47.
- Scargle, J. D. (1982). Studies in astronomical time series analysis. II - Statistical aspects of spectral analysis of unevenly spaced data. *Astrophysical Journal*, 263:835–853.
- Schaller, G., Schaerer, D., Meynet, G., and Maeder, A. (1992). New grids of stellar models from 0.8 to 120 solar masses at Z = 0.020 and Z = 0.001. *Astronomy and Astrophysics, Supplement*, 96:269–331.
- Schatzman, E. (1962). A theory of the role of magnetic activity during star formation. *Annales d'Astrophysique*, 25:18.
- Scott, T. R., Hearnshaw, J. B., Watson, R. D., Kilmartin, P. M., and Gilmore, A. C. (1992). Photometric analysis of the chromospherically active giant star HD 86005. *Journal of Astrophysics and Astronomy*, 13:279–286.

- Seager, S. and Mallén-Ornelas, G. (2003). A Unique Solution of Planet and Star Parameters from an Extrasolar Planet Transit Light Curve. *Astrophysical Journal*, 585:1038–1055.
- Sekiguchi, M. and Fukugita, M. (2000). A Study of the B-V Color-Temperature Relation. *Astronomical Journal*, 120:1072–1084.
- Serenelli, A. M., Bergemann, M., Ruchti, G., and Casagrande, L. (2013). Bayesian analysis of ages, masses and distances to cool stars with non-LTE spectroscopic parameters. *Monthly Notices of the Royal Astronomical Society*, 429:3645–3657.
- Shevelev, Y. G., Marsakov, V. A., and Suchkov, A. A. (1989). Relation Between the Kinematics Chemical Composition and Age of F-G Stars of the Galactic Disk. *Soviet Astronomy*, 33:160.
- Silva Aguirre, V., Basu, S., Brandão, I. M., et al. (2013). Stellar Ages and Convective Cores in Field Main-sequence Stars: First Asteroseismic Application to Two Kepler Targets. *Astrophysical Journal*, 769:141.
- Silva Aguirre, V., Casagrande, L., Basu, S., Campante, T. L., Chaplin, W. J., Huber, D., Miglio, A., Serenelli, A. M., Ballot, J., Bedding, T. R., Christensen-Dalsgaard, J., Creevey, O. L., Elsworth, Y., García, R. A., Gilliland, R. L., Hekker, S., Kjeldsen, H., Mathur, S., Metcalfe, T. S., Monteiro, M. J. P. F. G., Mosser, B., Pinsonneault, M. H., Stello, D., Weiss, A., Tenenbaum, P., Twicken, J. D., and Uddin, K. (2012a). Verifying Asteroseismically Determined Parameters of Kepler Stars Using Hipparcos Parallaxes: Self-consistent Stellar Properties and Distances. *Astrophysical Journal*, 757:99.
- Silva Aguirre, V., Casagrande, L., Basu, S., et al. (2012b). Verifying Asteroseismically Determined Parameters of Kepler Stars Using Hipparcos Parallaxes: Self-consistent Stellar Properties and Distances. *Astrophysical Journal*, 757:99.
- Silva Aguirre, V., Chaplin, W. J., Ballot, J., et al. (2011). Constructing a One-solar-mass Evolutionary Sequence Using Asteroseismic Data from Kepler. *Astrophysical Journal, Letters*, 740:L2.
- Silva Aguirre, V., Davies, G. R., Basu, S., Christensen-Dalsgaard, J., Creevey, O., Metcalfe, T. S., Bedding, T. R., Casagrande, L., Handberg, R., Lund, M. N., Nissen, P. E., Chaplin, W. J., Huber, D., Serenelli, A. M., Stello, D., Van Eylen, V., Campante, T. L., Elsworth, Y., Gilliland, R. L., Hekker, S., Karoff, C., Kawaler, S. D., Kjeldsen, H., and Lundkvist, M. S. (2015). Ages and fundamental properties of Kepler exoplanet host stars from asteroseismology. *Monthly Notices of the Royal Astronomical Society*, 452:2127–2148.
- Silva Aguirre, V. and Serenelli, A. M. (2015). Asteroseismic age determination for dwarfs and giants. *ArXiv e-prints*.

- Skumanich, A. (1972). Time Scales for CA II Emission Decay, Rotational Braking, and Lithium Depletion. *Astrophysical Journal*, 171:565.
- Smith, A. W. and Lissauer, J. J. (2009). Orbital stability of systems of closely-spaced planets. *Icarus*, 201:381–394.
- Smith, J. C., Stumpe, M. C., Van Cleve, J. E., et al. (2012). Kepler Presearch Data Conditioning II - A Bayesian Approach to Systematic Error Correction. *Publications of the A. S. P.*, 124:1000–1014.
- Soderblom, D. R. (2010). The Ages of Stars. *Anual Review of Astron and Astrophys*, 48:581–629.
- Soderblom, D. R., Duncan, D. K., and Johnson, D. R. H. (1991). The chromospheric emission-age relation for stars of the lower main sequence and its implications for the star formation rate. *Astrophysical Journal*, 375:722–739.
- Spada, F., Demarque, P., Kim, Y.-C., and Sills, A. (2013). The Radius Discrepancy in Low-mass Stars: Single versus Binaries. *Astrophysical Journal*, 776:87.
- Stumpe, M. C., Smith, J. C., Van Cleve, J. E., et al. (2012). Kepler Presearch Data Conditioning I - Architecture and Algorithms for Error Correction in Kepler Light Curves. *Publications of the A. S. P.*, 124:985–999.
- Thompson, B., Frinchaboy, P., Kinemuchi, K., Sarajedini, A., and Cohen, R. (2014). WIYN Open Cluster Study. LXII. Comparison of Isochrone Systems using Deep Multi-band Photometry of M35. *Astronomical Journal*, 148:85.
- Toutain, T. and Froehlich, C. (1992). Characteristics of solar p-modes - Results from the IPHIR experiment. *Astronomy & Astrophysics*, 257:287–297.
- van Saders, J. L., Ceillier, T., Metcalfe, T. S., Silva Aguirre, V., Pinsonneault, M. H., García, R. A., Mathur, S., and Davies, G. R. (2016). Weakened magnetic braking as the origin of anomalously rapid rotation in old field stars. *Nature*, 529:181–184.
- van Saders, J. L. and Pinsonneault, M. H. (2013). Fast Star, Slow Star; Old Star, Young Star: Subgiant Rotation as a Population and Stellar Physics Diagnostic. *Astrophysical Journal*, 776:67.
- Vanderburg, A. and Johnson, J. A. (2014). A Technique for Extracting Highly Precise Photometry for the Two-Wheeled Kepler Mission. *Publications of the A. S. P.*, 126:948–958.
- Vanderburg, A., Latham, D. W., Buchhave, L. A., Bieryla, A., Berlind, P., Calkins, M. L., Esquerdo, G. A., Welsh, S., and Johnson, J. A. (2016). Planetary Candidates from the First Year of the K2 Mission. *Astrophysical Journal, Supplements*, 222:14.

- Vanderburg, A., Montet, B. T., Johnson, J. A., Buchhave, L. A., Zeng, L., Pepe, F., Collier Cameron, A., Latham, D. W., Molinari, E., Udry, S., Lovis, C., Matthews, J. M., Cameron, C., Law, N., Bowler, B. P., Angus, R., Baranec, C., Bieryla, A., Boschin, W., Charbonneau, D., Cosentino, R., Dumusque, X., Figueira, P., Guenther, D. B., Harutyunyan, A., Hellier, C., Kuschnig, R., Lopez-Morales, M., Mayor, M., Micela, G., Moffat, A. F. J., Pedani, M., Phillips, D. F., Piotto, G., Pollacco, D., Queloz, D., Rice, K., Riddle, R., Rowe, J. F., Rucinski, S. M., Sasselov, D., Ségransan, D., Sozzetti, A., Szentgyorgyi, A., Watson, C., and Weiss, W. W. (2015). Characterizing K2 Planet Discoveries: A Super-Earth Transiting the Bright K Dwarf HIP 116454. *Astrophysical Journal*, 800:59.
- VanderPlas, J. (2014). Frequentism and Bayesianism: A Python-driven Primer. *ArXiv e-prints*.
- Veras, D., Brown, D. J. A., Mustill, A. J., and Pollacco, D. (2015). Prospects for detecting decreasing exoplanet frequency with main-sequence age using PLATO. *Monthly Notices of the Royal Astronomical Society*, 453:67–72.
- Weber, E. J. and Davis, Jr., L. (1967). The Angular Momentum of the Solar Wind. *Astrophysical Journal*, 148:217–227.
- Woitaszek, M., Metcalfe, T., and Shorrock, I. (2010). AMP: A Science-driven Web-based Application for the TeraGrid. *ArXiv e-prints*.
- Wolfgang, A., Rogers, L. A., and Ford, E. B. (2015). Probabilistic Mass-Radius Relationship for Sub-Neptune-Sized Planets. *ArXiv e-prints*.
- Wolszczan, A. and Frail, D. A. (1992). A planetary system around the millisecond pulsar PSR1257 + 12. *Nature*, 355:145–147.
- Wright, J. T. (2005). Radial Velocity Jitter in Stars from the California and Carnegie Planet Search at Keck Observatory. *Publications of the A. S. P.*, 117:657–664.
- Yi, S., Demarque, P., Kim, Y.-C., Lee, Y.-W., Ree, C. H., Lejeune, T., and Barnes, S. (2001). Toward Better Age Estimates for Stellar Populations: The  $Y^2$  Isochrones for Solar Mixture. *Astrophysical Journal, Supplements*, 136:417–437.
- Yıldız, M. (2007). Models of  $\alpha$  Centauri A and B with and without seismic constraints: time dependence of the mixing-length parameter. *Monthly Notices of the Royal Astronomical Society*, 374:1264–1270.
- Zhou, J.-L., Lin, D. N. C., and Sun, Y.-S. (2007). Post-oligarchic Evolution of Proto-planetary Embryos and the Stability of Planetary Systems. *Astrophysical Journal*, 666:423–435.