UNIVERSITÀ DEGLI STUDI DI CATANIA
FACOLTÀ DI SCIENZE MATEMATICHE FISICHE E NATURALI
DIPARTIMENTO DI FISICA E ASTRONOMIA - SEZIONE ASTROFISICA
DOTTORATO DI RICERCA IN FISICA - XXV CICLO

Enrico Maria Nicola Corsaro

Asteroseismology in the *Kepler* era and from ground-based observations

TESI DI DOTTORATO

Tutors:
Prof. Francesco Leone
Dr. Alfio Bonanno
Dr. Dennis Stello

Coordinator:
Prof. Francesco Riggi

ANNO ACCADEMICO 2011-2012

"... the ways by which men arrive at knowledge of the celestial things are hardly less wonderful than the nature of these things themselves"
Johannes Kepler

"Mystery creates wonder and wonder is the basis of man's desire to understand."
Neil Armstrong

Alla mia famiglia.

Contents

List of Figures

List of Tables

List of Peer-Reviewed Papers

- **Corsaro, E.**; Fröhlich, H.-E.; Bonanno, A.; Huber, D.; Bedding, T. R.; Benomar, O.; De Ridder, J.; Stello, D., *A Bayesian approach to scaling relations for amplitudes of solar-like oscillations in Kepler stars*, 2012, MNRAS (submitted), arXiv:1212.1156v1 [astro-ph.SR]

- **Corsaro, Enrico**; Stello, Dennis; Huber, Daniel; Bedding, Timothy R.; Bonanno, Alfio; Brogaard, Karsten; Kallinger, Thomas; Benomar, Othman; White, Timothy R.; Mosser, Benoit; Basu, Sarbani; Chaplin, William J.; Christensen-Dalsgaard, Jørgen; Elsworth, Yvonne P.; García, Rafael A.; Hekker, Saskia; Kjeldsen, Hans; Mathur, Savita; Meibom, Søren; Hall, Jennifer R.; Ibrahim, Khadeejah A.; Klaus, Todd C., *Asteroseismology of the Open Clusters NGC 6791, NGC 6811, and NGC 6819 from 19 Months of Kepler Photometry*, 2012, ApJ, ApJ, 757, 190

- Bonaca, Ana; Tanner, Joel D.; Basu, Sarbani; Chaplin, William J.; Metcalfe, Travis S.; Monteiro, Mário J. P. F. G.; Ballot, Jérôme; Bedding, Timothy R.; Bonanno, Alfio; Broomhall, Anne-Marie; Bruntt, Hans; Campante, Tiago L.; Christensen-Dalsgaard, Jørgen; **Corsaro, Enrico**; Elsworth, Yvonne; García, Rafael A.; Hekker, Saskia; Karoff, Christoffer; Kjeldsen, Hans; Mathur, Savita; Régulo, Clara; Roxburgh, Ian; Stello, Dennis; Trampedach, Regner; Barclay, Thomas; Burke, Christopher J.; Caldwell, Douglas A., *Calibrating Convective Properties of Solar-like Stars in the Kepler Field of View*, 2012, ApJ, 755, 12

- Fröhlich, H.-E.; Frasca, A.; Catanzaro, G.; Bonanno, A.; **Corsaro, E.**; Molenda-Żakowicz, J.; Klutsch, A.; Montes, D., *Magnetic activity and differential rotation in the young Sun-like stars KIC 7985370 and KIC 7765135*, 2012, A&A, 543, 146

- **Corsaro, E.**; Grundahl, F.; Leccia, S.; Bonanno, A.; Kjeldsen, H.; Paternò, L., *Solar-like oscillations in the G9.5 subgiant β Aquilae*, 2012, A&A, 537, 9

- Huber, D.; Bedding, T. R.; Stello, D.; Hekker, S.; Mathur, S.; Mosser, B.; Verner, G. A.; Bonanno, A.; Buzasi, D. L.; Campante, T. L.; Elsworth, Y. P.; Hale, S. J.; Kallinger, T.; Silva Aguirre, V.; Chaplin, W. J.; De Ridder, J.; García, R. A.; Appourchaux, T.; Frandsen, S.; Houdek, G.; Molenda-Żakowicz, J.; Monteiro, M. J. P. F. G.; Christensen-Dalsgaard, J.; Gilliland, R. L.; Kawaler, S. D.; Kjeldsen, H.; Broomhall, A. M.; **Corsaro, E.**; Salabert, D.; Sanderfer, D. T.; Seader, S. E.; Smith, J. C., *Testing Scaling Relations for Solar-like Oscillations from the Main Sequence to Red Giants Using Kepler Data*, 2011, ApJ, 743, 143

- White, Timothy R.; Bedding, Timothy R.; Stello, Dennis; Appourchaux, Thierry; Ballot, Jérôme; Benomar, Othman; Bonanno, Alfio; Broomhall, Anne-Marie; Campante, Tiago L.; Chaplin, William J.; Christensen-Dalsgaard, Jørgen; **Corsaro, Enrico**; Doğan, Gülnur; Elsworth, Yvonne P.; Fletcher, Stephen T.; García, Rafael A.; Gaulme, Patrick; Handberg, Rasmus; Hekker, Saskia; Huber, Daniel; Karoff, Christoffer; Kjeldsen, Hans;

Mathur, Savita; Mosser, Benoit; Monteiro, Mario J. P. F. G.; Régulo, Clara; Salabert, David; Silva Aguirre, Victor; Thompson, Michael J.; Verner, Graham; Morris, Robert L.; Sanderfer, Dwight T.; Seader, Shawn E., *Asteroseismic Diagrams from a Survey of Solar-like Oscillations with Kepler*, 2011, ApJ, 742, 3

- Silva Aguirre, V.; Chaplin, W. J.; Ballot, J.; Basu, S.; Bedding, T. R.; Serenelli, A. M.; Verner, G. A.; Miglio, A.; Monteiro, M. J. P. F. G.; Weiss, A.; Appourchaux, T.; Bonanno, A.; Broomhall, A. M.; Bruntt, H.; Campante, T. L.; Casagrande, L.; **Corsaro, E.**; Elsworth, Y.; García, R. A.; Gaulme, P.; Handberg, R.; Hekker, S.; Huber, D.; Karoff, C.; Mathur, S.; Mosser, B.; Salabert, D.; Schönrich, R.; Sousa, S. G.; Stello, D.; White, T. R.; Christensen-Dalsgaard, J.; Gilliland, R. L.; Kawaler, S. D.; Kjeldsen, H.; Houdek, G.; Metcalfe, T. S.; Molenda-Żakowicz, J.; Thompson, M. J.; Caldwell, D. A.; Christiansen, J. L.; Wohler, B., *Constructing a One-solar-mass Evolutionary Sequence Using Asteroseismic Data from Kepler*, 2011, ApJ, 740, 2

- Campante, T. L.; Handberg, R.; Mathur, S.; Appourchaux, T.; Bedding, T. R.; Chaplin, W. J.; García, R. A.; Mosser, B.; Benomar, O.; Bonanno, A.; **Corsaro, E.**; Fletcher, S. T.; Gaulme, P.; Hekker, S.; Karoff, C.; Régulo, C.; Salabert, D.; Verner, G. A.; White, T. R.; Houdek, G.; Brandão, I. M.; Creevey, O. L.; Doğan, G.; Bazot, M.; Christensen-Dalsgaard, J.; Cunha, M. S.; Elsworth, Y.; Huber, D.; Kjeldsen, H.; Lundkvist, M.; Molenda-Żakowicz, J.; Monteiro, M. J. P. F. G.; Stello, D.; Clarke, B. D.; Girouard, F. R.; Hall, J. R., *Asteroseismology from multi-month Kepler photometry: the evolved Sun-like stars KIC 10273246 and KIC 10920273*, 2011, A&A, 534, 6

- Verner, G. A.; Elsworth, Y.; Chaplin, W. J.; Campante, T. L.; **Corsaro, E.**; Gaulme, P.; Hekker, S.; Huber, D.; Karoff, C.; Mathur, S.; Mosser, B.; Appourchaux, T.; Ballot, J.; Bedding, T. R.; Bonanno, A.; Broomhall, A.-M.; García, R. A.; Handberg, R.; New, R.; Stello, D.; Régulo, C.; Roxburgh, I. W.; Salabert, D.; White, T. R.; Caldwell, D. A.; Christiansen, J. L.; Fanelli, M. N., *Global asteroseismic properties of solar-like oscillations observed by Kepler: a comparison of complementary analysis methods*, 2011, MNRAS, 415, 3539

- Mathur, S.; Handberg, R.; Campante, T. L.; García, R. A.; Appourchaux, T.; Bedding, T. R.; Mosser, B.; Chaplin, W. J.; Ballot, J.; Benomar, O.; Bonanno, A.; **Corsaro, E.**; Gaulme, P.; Hekker, S.; Régulo, C.; Salabert, D.; Verner, G.; White, T. R.; Brandão, I. M.; Creevey, O. L.; Doğan, G.; Elsworth, Y.; Huber, D.; Hale, S. J.; Houdek, G.; Karoff, C.; Metcalfe, T. S.; Molenda-Żakowicz, J.; Monteiro, M. J. P. F. G.; Thompson, M. J.; Christensen-Dalsgaard, J.; Gilliland, R. L.; Kawaler, S. D.; Kjeldsen, H.; Quintana, E. V.; Sanderfer, D. T.; Seader, S. E., *Solar-like Oscillations in KIC 11395018 and KIC 11234888 from 8 Months of Kepler Data*, 2011, ApJ, 733, 95

- Chaplin, W. J.; Bedding, T. R.; Bonanno, A.; Broomhall, A.-M.; García, R. A.; Hekker, S.; Huber, D.; Verner, G. A.; Basu, S.; Elsworth, Y.; Houdek, G.; Mathur, S.; Mosser, B.; New, R.; Stevens, I. R.; Appourchaux, T.; Karoff, C.; Metcalfe, T. S.; Molenda-Żakowicz, J.; Monteiro, M. J. P. F. G.; Thompson, M. J.; Christensen-Dalsgaard, J.; Gilliland, R. L.; Kawaler, S. D.; Kjeldsen, H.; Ballot, J.; Benomar, O.; **Corsaro, E.**; Campante, T. L.; Gaulme, P.; Hale, S. J.; Handberg, R.; Jarvis, E.; Régulo, C.; Roxburgh, I. W.; Salabert, D.; Stello, D.; Mullally, F.; Li, J.; Wohler, W., *Evidence for the Impact of Stellar Activity on the Detectability of Solar-like Oscillations Observed by Kepler*, 2011, ApJ, 732, 5

- 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.; Suran, M. D., *Predicting the Detectability of Oscillations in Solar-type Stars Observed by Kepler*, 2011, ApJ, 732, 54

- Chaplin, W. J.; Kjeldsen, H.; Christensen-Dalsgaard, J.; Basu, S.; Miglio, A.; Appourchaux, T.; Bedding, T. R.; Elsworth, Y.; García, R. A.; Gilliland, R. L.; Girardi, L.; Houdek, G.; Karoff, C.; Kawaler, S. D.; Metcalfe, T. S.; Molenda-Żakowicz, J.; Monteiro, M. J. P. F. G.; Thompson, M. J.; Verner, G. A.; Ballot, J.; Bonanno, A.; Brandão, I. M.; Broomhall, A.-M.; Bruntt, H.; Campante, T. L.; **Corsaro, E.**; Creevey, O. L.; Doğan, G.; Esch, L.; Gai, N.; Gaulme, P.; Hale, S. J.; Handberg, R.; Hekker, S.; Huber, D.; Jiménez, A.; Mathur, S.; Mazumdar, A.; Mosser, B.; New, R.; Pinsonneault, M. H.; Pricopi, D.; Quirion, P.-O.; Régulo, C.; Salabert, D.; Serenelli, A. M.; Silva Aguirre, V.; Sousa, S. G.; Stello, D.; Stevens, I. R.; Suran, M. D.; Uytterhoeven, K.; White, T. R.; Borucki, W. J.; Brown, T. M.; Jenkins, J. M.; Kinemuchi, K.; Van Cleve, J.; Klaus, T. C., *Ensemble Asteroseismology of Solar-Type Stars with the NASA Kepler Mission*, 2011, Science, 332, 213

Abstract

In this dissertation we present and investigate new results attained in the field of Asteroseismology, a branch of stellar physics that aims at deriving the fundamental stellar properties and at describing the internal structure of the stars by means of the oscillations observed at their surface. The stochastically excited and intrinsically damped acoustic modes, also known as *solar-like* oscillations, represent the main type of pulsations discussed in this thesis.

The document is sectioned into four parts, where the last one includes a supplementary material for the reader. Part I deals with ground-based observations of solar-like oscillations detected through the measurement of Doppler shifts from line profiles. Here, Chapter 1 provides a general and comprehensive introduction to Asteroseismology, with a focus on the class of solar-like oscillations. Chapter 2 readily describes the main techniques, tools and features of the asteroseismic analysis of solar-like oscillations, and subsequently presents a detailed study of the solar-like subgiant stars β Aquilae and μ Herculis, both observed by means of the SARG spectrograph operating @TNG. This study was done by using a preliminary version of the iSONG code, the new software to be used within the SONG project of stellar oscillations for radial velocity extractions.

Part II gives a particular focus to the space-based NASA's *Kepler* Mission, whose photometric observations represent the main source of data used for the research presented in this work. An overview of the NASA's *Kepler* Mission is given in Chapter 3, together with a brief presentation of some of the main results derived from large ensemble of stars observed with *Kepler*. The results discussed in this chapter were derived by means of the ORK pipeline developed @OACT. Chapter 4 instead, is related to a thorough investigation of the ensemble seismic properties of a sample of red giant stars belonging to the open clusters NGC 6791, NGC 6811, and NGC 6819 observed by *Kepler*. The entire analysis was done by means of the AARG code. The excellent quality of the 19 months-long observations carried out by *Kepler* made this analysis possible, despite of the low signal-to-noise ratio of these stars.

Last but not least, Part III debates with the use of Bayesian statistics applied to *Kepler* light curves for studying both the seismic properties of solar-like oscillations and the stellar differential rotation in active stars. In particular, an exhaustive introduction to the Bayesian concept of probability, the main statistical tools, and the Markov Chain Monte Carlo method are provided in Chapter 5. Chapter 6 presents a detailed Bayesian inference and model comparison applied to the scaling relations adopted for predicting the amplitudes of solar-like oscillations in stars spanning from main sequence to the late red giant phase of the stellar evolution. Finally, Chapter 7 is related to a detailed study of the differential rotation in the *Kepler* targets KIC 7985370 and KIC 7765135, two young and active Sun-like stars that show an evident photometric modulation caused by star spots at their surfaces.

The appendices are given in Part IV. Appendix A describes in detail the AARG code developed for the analysis discussed in Chapter 4. All the results derived with AARG are listed in Appendix B for each star investigated. Lastly, Appendix C deals with a description about some useful numerical methods adopted for the analysis described in Chapter 6, while Appendix D discusses the theory of Principal Component Analysis.

Preface

In the first period of my PhD project I was interested in the analysis of solar-like oscillations in the subgiants β Aquilae and μ Herculis. This study involved the data reduction of spectra collected at the TNG Italian telescope and the derivation of radial velocity time-series by means of the iSONG code [142], which I developed in part during my period of collaboration with Prof. Frank Grundahl and Dr. Hans Kjeldsen at the Department of Physics and Astronomy of Aarhus Univeristy. The analysis of these time-series was published in A&A [86].

Afterward I focused on the analysis of light curves acquired by the NASA's *Kepler* Mission. My research activity within the WG1 of the KASC (Kepler Asteroseismic Science Consortium) was related the investigation of the global asteroseismic parameters of $\sim$ 800 main sequence and early subgiant stars. In particular, I developed an automated pipeline together with my PhD tutor Dr. Alfio Bonanno, that allowed me to search for the global asteroseismic quantities of these stars. These results contributed to some publications in scientific journals such as Science, ApJ, A&A, MNRAS [71, 72, 73, 216, 310, 63, 275, 218, 317, 167, 45].

At the same time *Kepler* is allowing for detailed studies of differential rotation in active stars. During my visit at the AIP (Astrophysics Institute of Postdam), I collaborated with Dr. Hans-Erich Fröhlich in the analysis of active stars by means of a Bayesian spot modeling code. This also involved spectroscopic studies that were done by Dr. Antonio Frasca from OACT (Catania Astrophysical Observatory), for the measurement of fundamental stellar properties. The results of this work have been published in A&A [118]. The research conducted through these collaborations allowed me to gain knowledge and expertise about the use of Bayesian methods applied to the astrophysics field.

During my period of collaboration with Prof. Tim Bedding and my second PhD tutor Dr. Dennis Stello at the SIfA (Sydney Institute for Astronomy), I studied the asteroseismic properties of 115 red giant stars belonging to the open clusters NGC 6791, NGC 6811 and NGC 6819 observed by *Kepler*. This research has been accepted for publication by ApJ [87]. The code that I developed for this purpose (AARG - Asymptotic Analysis of Red Giant stars), allows for a detailed asteroseismic analysis of large samples of cluster red giant stars in a semi-automated manner. The AARG code is expected to produce even better results in the case of field red giant stars in the near future.

In conclusion to my PhD, my research has been focused to the study of amplitude scaling relations for solar-like oscillations by using a sample of 1640 targets, spanning from main sequence to red giant stars observed with *Kepler*. This involved the application of Bayesian statistics for both the parameter estimation and the model comparison of the scaling relations studied. The work sets the basis for a subsequent investigation of the methods used for measuring the amplitude uncertainties and possibly for further theoretical developments of new scaling relations. The paper produced with this research is currently under revision by MNRAS.

Enrico Corsaro
6$^{\text{th}}$ December 2012, Catania

Part I

Asteroseismology from ground-based Observations

Aerts, C., Christensen-Dalsgaard, J., Kurtz, D. W., *Asteroseismology*, First Edition, Springer Verlag, 2010 [1]
Christensen-Dalsgaard, J., *Lecture Notes on Stellar Oscillations. Fifth Edition*, University of Aarhus, 2003 [76]

1 Introduction

In the opening of *The Internal Constitution of the Stars* [105], Sir Arthur Eddington questioned about the possibility to study the interior of the stars without direct observations:

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

Nowadays, we are aware that from our knowledge of the basic laws of physics, and from the observable boundary conditions at the surface of a star, it is possible to investigate its interior structure, and also with a high level of confidence. In fact, many stars have sound waves traveling in their interior. Those waves cannot get out and not travel in a vacuum, but for many kinds of pulsating stars the sound waves make the star periodically swell and contract, get hotter and cooler. With our telescopes we can see the effects of this behavior: the periodic changes in the star's brightness; the periodic motion of its surface moving back-and- forth with respect to the observer. Thus we can detect the natural oscillations of the star and "hear" the sounds inside them.

It is the combination of the frequencies, amplitudes and phases of the harmonics that defines the character of the sound waves generated. A sound wave is in fact a pressure wave, which is basically a rarefaction and compression of the gas that propagates at the speed of sound. The propagation occurs indeed at the molecular level because the information that the high pressure wavefront is coming is transmitted by individual molecular collisions. In the adiabatic case, the speed of sound is $c = \sqrt{\Gamma_1 p/\rho}$, where Γ_1 is one of the adiabatic exponents (see below), p the pressure and ρ the density of the gas. Of course, for an ideal gas $p = \rho k_B T/\mu m_u$, where k_B is the Boltzmann's constant, μ is the mean molecular weight, and m_u the atomic mass unit; thus $c = \sqrt{\Gamma_1 k_B T/\mu m_u}$.

The changes in pressure are therefore accompanied by changes in density and temperature. In this case the speed of sound depends on the temperature and chemical composition of the gas. Thus, if the temperature is higher, and the molecules are moving more quickly, then the collisions become more frequent and the sound speed increases. Moreover, given a temperature of thermal equilibrium, in gases having a lighter molecular weight, molecules collide easier than for heavier gases. Therefore by measuring the speed of sound in a gas, we gain information about the pressure and density of that gas, and, with the additional use of the equation of state, we may also constrain the temperature and chemical composition.

Stars are made of gas and they have natural overtones (which, conversely to the harmonics of an instrument, would appear dissonant to our ears when played at audible frequencies),

and similarly as one can recognize the shape of the instrument producing a particular sound wave, we can use the frequencies, amplitudes and phases of the sound waves that we detect in the stars to "see" their interiors, i.e. to see their internal "shapes". One of the main goals of asteroseismology is in fact to measure the sound speed throughout a star so that we can derive the fundamental parameters of the stellar structure.

Asteroseismology uses astronomical observations in photometry and spectroscopy in order to extract the frequencies, amplitudes and phases of the sound waves at a star's surface [1]. Then, basic physics and mathematical models are used to infer the sound speed and density inside a star, throughout its interior, and hence derive a pressure profile. With reasonable assumptions about chemical composition and knowledge of appropriate equations of state, the temperature can then be derived.

The acoustic waves propagating in the stellar interior are known as "pressure" modes, or p modes. Anyhow, pulsations in stars are more complicated than simple sound waves generated from an instrument. In fact, there we can find equally important "gravity" modes, or g modes, where the restoring force of the pulsation is not pressure, but buoyancy. Much of the studies of stellar pulsation that can be found in the literature offer a valid view of the presence of gravity modes, though they are not acoustic. These modes allow us to probe different regions of the interior of stars than those where pressure waves propagate, as we shall discuss later in this chapter.

The Sun is certainly the star that has been investigated mostly in this sense. In fact, several thousand individual modes have been identified so far by observing its surface, and it is expected that with more precise observations a number up to 10^6 modes can be determined accurately. The enormous amount of information about the solar interior sets the basis for *helioseismology* (e.g. see [67]), the science of learning about the Sun from the observed frequencies. Helioseismology has already led to a considerable amount of information about the structure and rotation of the solar interior, similarly to what asteroseismology is actually doing with other stars.

A general background of helioseismology can be found in [97, 204, 79, 206, 135, 80]. Since the Sun is considered to be a normal star, similarly rich spectra of oscillations would be expected in other similar stars. However, due to the distance of the other stars their observation has no or very limited, spatial resolution, which has the effect of reducing the number of modes that can be detected. In fact, most of the observed solar modes have relatively short horizontal wavelength on the solar surface, and hence would not be detected in stellar observations due to the poor spatial resolution. A second problem in trying to detect the expected solar-like oscillations in other stars is their very small amplitudes. On the Sun the maximum velocity amplitude in a single mode is about $15\,\mathrm{cm\,s^{-1}}$, whereas the luminosity amplitudes are of the order of 1 μmag. As a consequence, a particular care is required when observing such oscillations in other stars, hence a great deal of effort is being spent on developing new instrumentation with the required sensitivity (see Chapters 2, 3).

In this chapter we will describe how oscillations in stars are parametrized in order to be used for inferring the fundamental stellar properties. In addition we present the asteroseismic relations and quantities that are at the basis for the analyses presented in the forthcoming chapters, together with some qualitatively discussions about the main aspects of this science. A detailed derivation and description of the equations that characterize asteroseismology goes beyond the scope of this work, but can be found in [1, 76] for the interested reader.

1.1 Oscillations in Stars

For understanding how pulsations are visualized into a star we start by introducing some important basic concepts that are widely known in the field of asteroseismology. Clearly, stars are three-dimensional objects, so their natural oscillation modes have nodes in three orthogonal directions. Assuming to consider a frame in spherical coordinates and having origin in the center of the star, the oscillations are described by the distance r to the centre, the co-latitude θ and the longitude ϕ, where θ is measured from the pulsation pole, i.e. the axis of symmetry (θ is in fact co-latitude, since latitude is measured from the equator). The nodes are concentric shells at constant r, cones of constant θ and planes of constant ϕ. For a spherically symmetric star the solutions to the equations of motion have displacements in the (r, θ, ϕ) directions and are given by[1].

$$\xi_r(r,\theta,\phi,t) = a(r)Y_\ell^m(\theta,\phi)e^{-i2\pi\nu t} \tag{1.1}$$

$$\xi_\theta(r,\theta,\phi,t) = b(r)\frac{\partial Y_\ell^m(\theta,\phi)}{\partial\theta}e^{-i2\pi\nu t} \tag{1.2}$$

$$\xi_\phi(r,\theta,\phi,t) = \frac{b(r)}{\sin\theta}\frac{\partial Y_\ell^m(\theta,\phi)}{\partial\phi}e^{-i2\pi\nu t} \tag{1.3}$$

where ξ_r, ξ_θ and ξ_ϕ are the displacements, $a(r)$ and $b(r)$ are amplitudes, ν is the so called *cyclic frequency* and $Y_\ell^m(\theta,\phi)$ are the spherical harmonics

$$Y_\ell^m(\theta,\phi) = (-1)^m c_{\ell m} P_\ell^m(\cos\theta)e^{im\phi} \tag{1.4}$$

being $P_\ell^m(\cos\theta)$ the Legendre polynomials given by

$$P_\ell^m(\cos\theta) = \frac{1}{2^\ell \ell!}(1-\cos^2\theta)^{m/2}\frac{d^{\ell+m}}{d\cos^{\ell+m}\theta}(\cos^2\theta-1)^\ell\,, \tag{1.5}$$

and $c_{\ell m}$ the normalization coefficients given by

$$c_{\ell m} \equiv \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}}\,, \tag{1.6}$$

which let the spherical harmonics to satisfy the normalization condition $|Y_\ell^m|^2 = 1$. In most of the pulsating stars the pulsation axis coincides with the rotation axis. The main exceptions are the rapidly oscillating Ap stars where the axis of pulsational symmetry is the magnetic axis which is tilted with respect to their rotational axis. Thus, in stars there are three quantum numbers to specify these modes: n is related to the number of radial nodes and is called the *overtone* of the mode, or simply *radial order*; ℓ is the *angular degree* of the mode and specifies the number of surface nodes that are present; m is the azimuthal order of the mode, where $|m|$ specifies how many of the surface nodes are lines of longitude. It follows therefore that the number of surface nodes that are lines of co-latitude is equal to $\ell - |m|$. The values of m range from $-\ell$ to $+\ell$, so there are $2\ell+1$ modes for each degree ℓ. In the following we shall describe the two kinds of modes observable in pulsating stars, namely radial and nonradial modes.

[1] A detailed analytical derivation of the equations that describe the oscillations in stars goes beyond the scope of this thesis. Details can be found in, e.g., [1, 76]

1.1.1 Radial Modes

The simplest modes are the *radial* modes with $\ell = 0$, and the simplest of those is the fundamental radial mode. In this mode the star swells and contracts, heats and cools, spherically symmetrically with the core as a node and the surface as a displacement antinode. This is the typical mode of pulsation for Cepheid variables and for RR Lyrae stars, amongst others.

The first overtone radial mode has one radial node that is a concentric shell within the star. As we are thinking in terms of the radial displacement, that shell is a node that does not move; the motions above and below the node move in antiphase. The surface of the star is again an antinode. There are Cepheid variables, RR Lyrae stars and δ Sct stars that pulsate simultaneously in the fundamental and first overtone radial modes. The of the ratios between the overtone periods and the fundamental one reflect the sound speed gradient in the stars where we observe these modes, hence they provide information about temperature and (in some places) chemical composition gradients.

1.1.2 Nonradial Modes

The simplest of the nonradial modes is the axisymmetric *dipole* mode with $\ell = 1$, $m = 0$. For this mode the equator is a node; the northern hemisphere swells up while the southern hemisphere contracts, then vice versa; one hemisphere heats while the other cools, and so on and so forth, all with the simple cosine dependence of $P_1^0(\cos\theta) = \cos\theta$, where θ is again the co-latitude. From the observer's point of view, the star seems to oscillate up and down in space because the centre-of-mass of the star is not displaced during dipole oscillations due to the fact that stars are not incompressible bodies. Nonradial modes only occur for $n \geq 1$, so in the case of the $\ell = 1$ dipole mode, there is at least one radial node within the star. While the outer shell is displaced upwards from the point of view of the observer, the inner shell is displaced downwards and the centre of mass stays fixed. Dipole modes are observed in many kinds of pulsating variables.

Modes with two surface nodes ($\ell = 2$) are known as quadrupole modes. For the $\ell = 2$, $m = 0$ mode the nodes lie at latitudes $\pm 35°$, since $P_2^0(\cos\theta) = (3\cos^2\theta - 1)/2$ (see also [1] for further details). The poles of an $\ell = 2$, $m = 0$ mode swell up (and heat up, although not usually in phase with the swelling) while the equator contracts (and cools), and vice versa. Figure 1.1 represents and explains a set of octupole modes with $\ell = 3$, giving a mental picture of what the modes look like on the stellar surface which is generally inclined with respect to the line-of-sight.

Unfortunately, we are not yet at the stage where we can resolve stellar surfaces and detect the nodal lines directly from intensity or Doppler maps such as the ones shown in Figure 1.1. This can only be done for the Sun so far. As a consequence, for other stars we have to deal with observations representing integrated quantities over the stellar surface, such as the surface-averaged brightness or radial velocity. It is then intuitively clear that, for a fixed value of the amplitude of the oscillation, and for a particular value of the inclination of the symmetry axis of pulsation with respect to the line-of-sight, such observed quantities must be smaller for higher angular degree ℓ modes than for lower degree modes. Indeed, the higher ℓ, the more sectors and/or zones will divide the stellar surface, with neighbouring regions having opposite sign in intensity or velocity. Their influence on the integrated quantity therefore partially tends to cancel out. A rigorous mathematical expression for the partial cancellation can be found in [1], Chapter 6, or in [76]. To get a feel for the consequences of this effect, let us assume here the simplest case, which is the surface-integrated intensity of an axisymmetric mode over a stellar disc that does not suffer from limb darkening. In that simplest case, the partial cancellation is

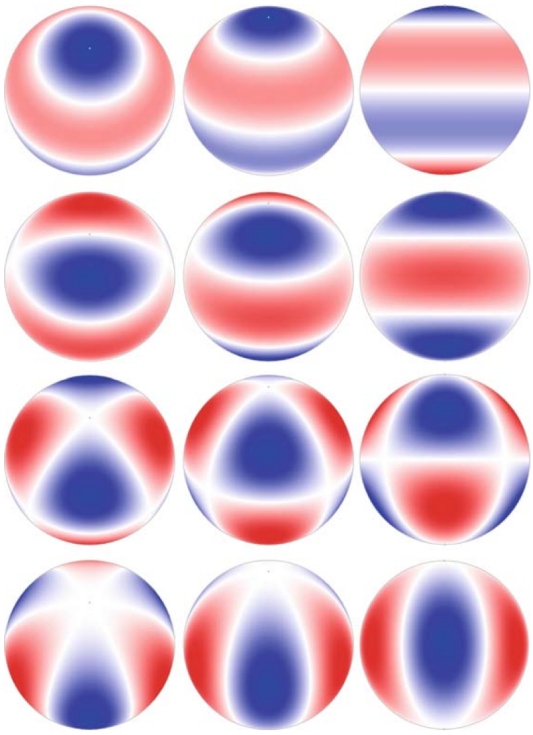

Figure 1.1: *Snapshot of the radial component of the $\ell = 3$ octupole modes. The columns show the modes from different viewing angles; the left column is for an inclination of the pulsation pole of 30°, the middle column is for 60°, and the right column is for 90°. The white bands represent the positions of the surface nodes; red and blue represent sections of the star that are moving in (out) and/or heating (cooling) at any given time, then vice versa. The top row shows the axisymmetric octupole mode ($\ell = 3$, $m = 0$) where the nodes lie at latitudes ±51° and 0°. The second row shows the tesseral (meaning $0 < |m| < \ell$) $\ell = 3$, $m = \pm 1$ mode with two nodes that are lines of latitude and one that is a line of longitude. The third row is the tesseral $\ell = 3$, $m = \pm 2$ mode, and the bottom row shows the sectoral mode (meaning $\ell = |m|$) with $\ell = 3$, $m = \pm 3$. Importantly, rotation distinguishes the sign of m. Image copied from [1].*

described well by an integral of the intensity eigenfunction over the visible stellar disc, i.e., it is proportional to

$$c_{\ell 0} \int_0^{\pi/2} P_\ell(\cos\theta) \sin\theta \cos\theta d\theta \,, \tag{1.7}$$

where $c_{\ell 0}$ is defined according to Eq. (1.6). This factor is shown for all axisymmetric modes with $\ell = 0, \ldots, 10$ in Figure 1.2. The radial mode does not suffer from partial cancellation and thus reaches value unity, which is about twice as high as a quadrupole $\ell = 2$ mode. It is very important to be aware that axisymmetric $\ell = 3$ modes are almost invisible in intensity measurements due to the partial cancelling, while they are still quite visible in velocity measurements because they allow for a slightly higher spatial resolution[2]. Partial cancellation for the $\ell = 4$ mode is a factor 10 greater than that of the dipole mode, and this factor increases as ℓ increases. While the inclusion of rotation and limb darkening complicates this simplistic description, and the effects are more complicated for velocity quantities than for the intensity, Figure 1.2 explains why even modes are much easier to detect in the data than odd modes,

[2]This is caused by the fact that Doppler shifts let us distinguish about different regions on the stellar's surface more than intensity variations do. This also represents an interesting advantage of spectroscopic measurements to the photometric ones.

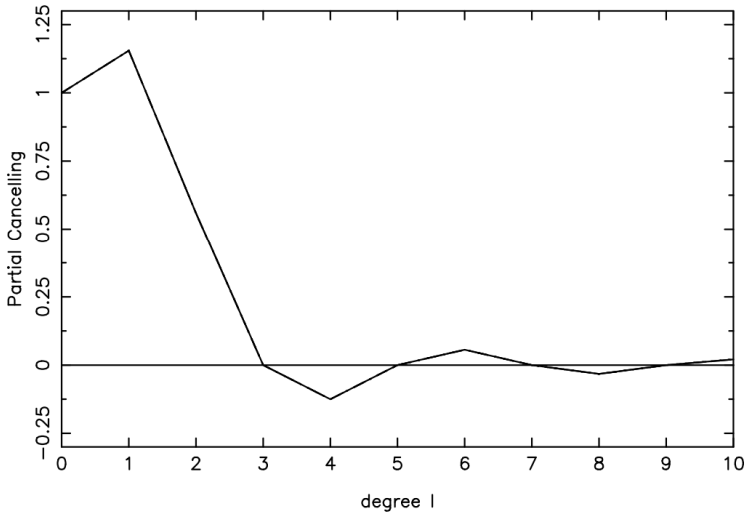

Figure 1.2: *The partial cancelling factor for the surface-integrated intensity in the case of axisymmetric modes for $\ell = 0, 1, \ldots, 10$, when ignoring the darkening at the limb of the stellar surface. Image copied from [1].*

except for the special case of the dipole mode. The modes become more difficult to detect as their degree increases, and in general no modes having $\ell > 3$ have been identified so far in any star different than the Sun.

1.1.3 Including the Rotation Effect

Although the rotation effect is not part of the asteroseismic studies presented in this dissertation, for the sake of completeness we shall discuss it briefly in the following. In Eqs. (1.1) and (1.4) it can be seen that for modes with $m \neq 0$ the exponentials in the two equations combine to give a time dependence that goes as $\exp[-i(2\pi\nu t - m\phi)]$. This phase factor in the time dependence means that the $m \neq 0$ modes are traveling waves, where our sign convention is that modes with positive m are traveling in the direction of rotation (prograde modes), and modes with negative m are traveling against the direction of rotation (retrograde modes).

For a spherically symmetric star the frequencies of all $2\ell + 1$ members of a multiplet (such as the octupole septuplet $\ell = 3$, $m = -3, -2, -1, 0, +1, +2, +3$) are the same. But deviations from spherical symmetry can lift this frequency degeneracy, and the most important physical cause of a star's departure from spherical symmetry is rotation. For example, in a rotating star the Coriolis force causes pulsational variations that would have been up-and-down to become circular with the direction of the Coriolis force being against the direction of rotation. Because of this effect (and others effects as well, e.g. see [1]), the prograde modes traveling in the direction of rotation have frequencies slightly lower than the $m = 0$ axisymmetric mode, and the retrograde modes going against the rotation have slightly higher frequencies, in the co-rotating reference frame of the star, thus the degeneracy of the frequencies of the multiplet is lifted. This was discussed by [205] in a study of the β Cep star β CMa (see [1], Chapter 2, for a definition of this class of stars). In the *observer's frame of reference* the Ledoux rotational splitting relation for a uniformly rotating star is

$$\nu_{n\ell m} = \nu_{n\ell 0} + m(1 - C_{n\ell})\Omega/2\pi\,, \tag{1.8}$$

where $\nu_{n\ell m}$ is the observed frequency of the mode having numbers n, ℓ and m, $\nu_{n\ell 0}$ is the unperturbed central frequency of the multiplet (for which $m = 0$) which is unaffected by

the rotation, $C_{n\ell}$ is a model-dependent quantity with value below 1 (see [1], Chapter 3 for a definition), and Ω is the angular velocity, corresponding to a rotation frequency of $\Omega/2\pi$. If we rewrite Eq. (1.8) as

$$\nu_{n\ell m} = \nu_{n\ell 0} - mC_{n\ell}\Omega/2\pi + m\Omega/2\pi\,, \tag{1.9}$$

then it is easy to see that the Coriolis force, contributing as the second term in the right-hand side of the equation, reduces the frequency of the prograde modes with positive m slightly in the co-rotating rest frame, but then the rotation frequency is added to that since the mode is going in the direction of rotation. Likewise the retrograde modes with negative m are traveling against the rotation so have their frequency in the observer's frame reduced by the rotation frequency.

In this way we end up with a multiplet with $2\ell+1$ components all separated by the rotational splitting $(1 - C_{n\ell})\Omega/2\pi$. In a real star rotation is not expected to be uniform and hence the rotational splitting would depend on the properties of the modes in a more complicated manner; also, the various components of the multiplet may be excited to different amplitudes, and some may not have any observable amplitude, so all members of the multiplet may not be present. The importance for asteroseismology is that in case such rotationally-split multiplets are observed, the ℓ and m of the corresponding modes may be identified and the splitting used to measure the rotation rate of the star. Where multiplets of modes of different degree or different overtone are observed, it is possible to gain knowledge of the *interior* rotation rate of the star – something that is not knowable by any other means.

In the case of the Sun, helioseismology has spectacularly measured the differential rotation rate of the Sun down to about half way to the core. Below the convection zone at $r/R_\odot \sim 0.7$ the Sun rotates approximately rigidly with a period close to the 27-d period seen at latitudes of about 35° on the surface. Within the convection zone the rotation is not simply dependent on distance from the solar rotation axis, as had been expected in the absence of any direct observation. It is certainly a great triumph of helioseismology that we can know the internal rotation behaviour of the Sun. In addition, rotationally-split frequencies have been recently identified also in red giant stars observed by *Kepler* [22, 23], providing the first remarkable measure of internal rotation rates in distant stars.

1.2 Using Asteroseismology

Since p modes are acoustic waves, for modes that are not directed at the centre of the star (i.e., the nonradial modes) the lower part of the wave is in a higher temperature environment than the upper part of the wave, thus in a region of higher sound speed according to the discussion given in the introduction of this chapter. As a consequence the wave is refracted back to the surface, where it is then reflected, since the acoustic energy is trapped in the star, as can be seen in Figure 1.4(a). For higher ℓ values, the modes have more reflection points. This means that high degree modes penetrate only to a shallow depth, while lower degree modes penetrate more deeply. The frequency of the mode observed at the surface depends on the sound travel time along its ray path, hence on the integral of the sound speed within its "acoustic cavity". Clearly, if many modes that penetrate to all possible depths can be observed on the surface, then it is possible to "invert" the observations to make a map of the sound speed throughout the star, and from that deduce the temperature profile, with reasonable assumptions about the chemical composition. In the Sun the sound speed is now known to a few parts per thousand over 90% of its radius. To do the same for other stars is an ultimate goal of asteroseismology.

Thus asteroseismology lets us literally see the insides of stars because different modes penetrate to different depths in the star. However, stellar oscillations are not so simple as just p

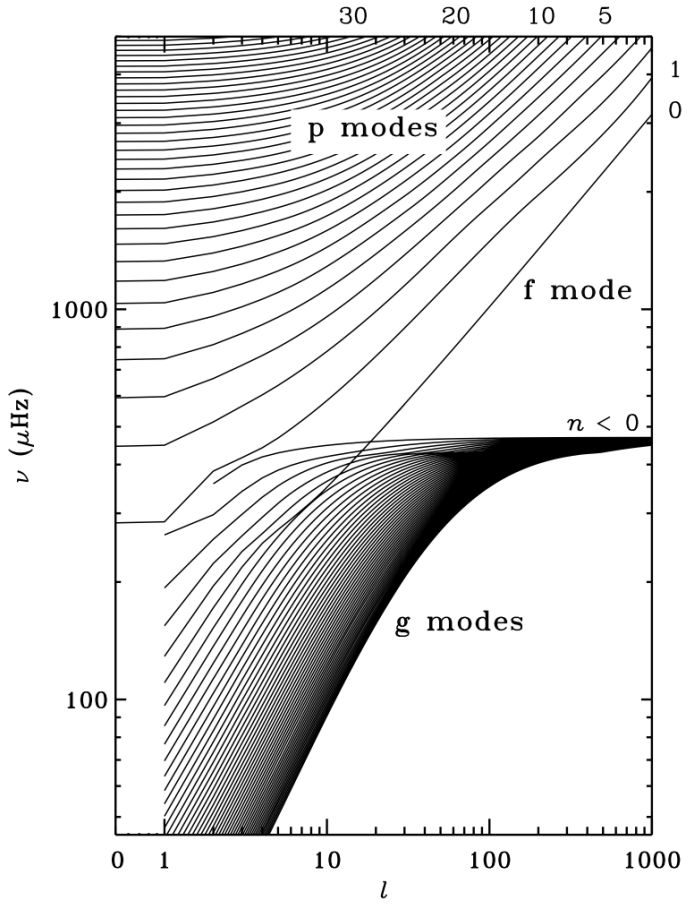

Figure 1.3: *The frequency of modes versus their degree ℓ for a solar model. The figure clearly illustrates the general property of* p *modes that frequency increases with overtone n and degree ℓ. For* g *modes frequency decreases with higher overtone, but increases with n if we use the convention that n is negative for g modes. Frequency still increases with degree ℓ for* g *modes, just as it does for* p *modes. Some values of the overtone n are given for the* p *modes lines in the upper right of the figure. Note that while continuous lines are shown for clarity, the individual modes are discrete points, corresponding to integer ℓ, which are not shown here. Image copied from [1].*

modes. We can also see inside the stars with g modes. In fact, for some stars, and for parts of others, we can only see with g modes (see also the recent findings for *Kepler* stars done by [21, 32] and the mixed mode analysis presented in Chapter 4).

1.2.1 p Modes and g Modes

There are two main sets of solutions to the equation of motion for a pulsating star, and these lead to two types of pulsation modes: p modes and g modes. As already mentioned at the beginning of our discussion, for the p modes, or pressure modes, pressure is the primary restoring force for a star perturbed from equilibrium. These p modes are acoustic waves and have gas motions that are primarily vertical. For the g modes, or gravity modes, buoyancy is the restoring force and the gas motions are primarily horizontal. There is also an f mode situated between the p mode of radial order 1 and the g mode of radial order 1 for all $\ell \geq 2$.

Both p and g modes of high order can be described in terms of the propagation of rays. This provides illuminating graphical representations of their properties; examples are shown

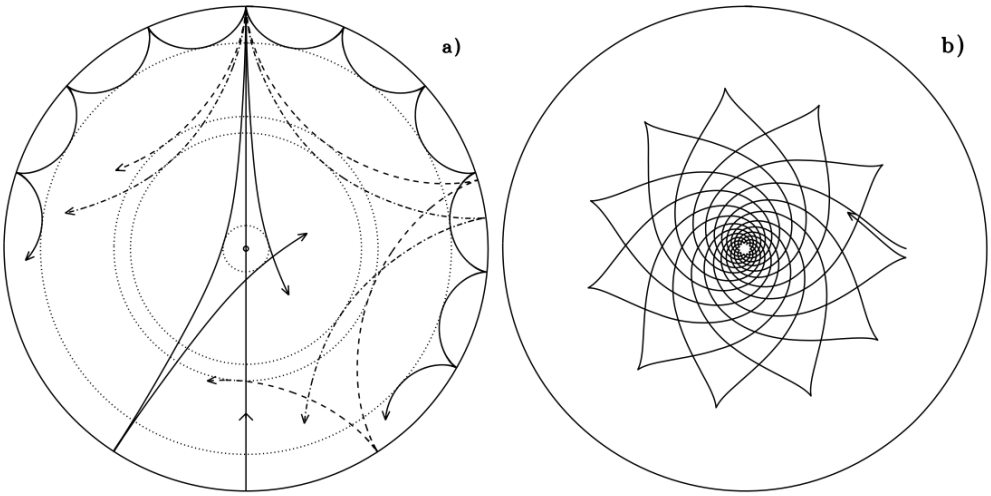

Figure 1.4: *Propagation of rays of sound or gravity waves in a cross-section of a Sun-like star. The acoustic ray paths (panel a) are bend by the increase in sound speed with depth until they reach the inner turning point (indicated by the dotted circles) where they undergo total internal refraction. At the surface the acoustic waves are reflected by the rapid decrease in density. Shown are rays corresponding to modes of frequency* $3000\,\mu Hz$ *and degrees (in order of increasing penetration depth)* $\ell = 75, 25, 20$ *and* 2*; the line passing through the centre schematically illustrates the behavior of a radial mode. The g-mode ray path (panel b) corresponds to a mode of frequency* $190\,\mu Hz$ *and angular degree 5 and is trapped in the interior. In this example, it does not propagate in the convective outer part.* g *modes are observed at the surface of other types of pulsators. This figure illustrates that the* g *modes are sensitive to the conditions in the very core of the star. Image copied from [1].*

in Figures 1.4 and 1.5. Also, this representation forms the basis for powerful asymptotic descriptions of the modes [1, 76]. There are three other important properties of p modes and g modes:

1. as the number of radial nodes increases the frequencies of the p modes increase, but the frequencies of the g modes decrease, as is shown in Figure 1.3

2. for stars having an internal structure similar to that of the Sun, the p modes are most sensitive to conditions in the outer part of the star, whereas g modes are most sensitive to conditions in the deep interior of the star (except in white dwarfs where the g modes are sensitive mainly to conditions in the stellar envelope), as is shown in Figure 1.4

3. for $n \gg \ell$ an *asymptotic relation* for p shows that they are approximately equally spaced in frequency, while another asymptotic relation for g modes points out that they are approximately equally spaced in period.

As illustrated in Figure 1.4(b), g modes in solar-like stars are trapped beneath the convective envelope, when viewed as rays. In reality the modes have finite amplitudes also in the outer parts of the star and hence, at least in principle, can be observed on the surface; this is in fact the case in the γ Dor stars which have convective envelopes. In more massive main-sequence stars, such as illustrated in Figure 1.5, the g-mode rays are confined outside the convective core.

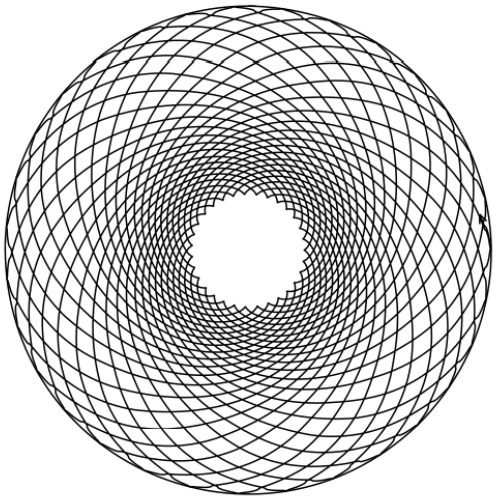

Figure 1.5: *Propagation of rays of gravity waves in a cross-section of an $8M_\odot$ ZAMS star. The ray path corresponds to a mode of frequency $50\,\mu Hz$ and $\ell = 5$. It is trapped outside the convective core of the star. Image credit by [1].*

1.2.2 The Asymptotic Relations

The asymptotic relations are very important in many pulsating stars. From [299] they show that for the p modes, the frequencies are approximately given by

$$\nu_{n\ell} = \Delta\nu \left(n + \frac{\ell}{2} + \epsilon \right) + \delta\nu_{n\ell}\,, \tag{1.10}$$

where n and ℓ are the overtone (radial order) and angular degree of the mode, ϵ is a phase shift sensitive to the properties of the near surface layers of the star [81], and $\delta\nu_{n\ell}$ is a small correction. $\Delta\nu$ is known as the *large frequency separation*, or simply the *large separation*, and is the inverse of the sound travel time for a sound wave from the surface of the star to the core and back again [304], given by

$$\Delta\nu = \left(2 \int_0^R \frac{dr}{c(r)} \right)^{-1}\,, \tag{1.11}$$

where $c(r)$ is the sound speed as a function of the radial coordinate. The large separation is obviously sensitive to the radius of the star, hence near the main sequence it is a good measure of the mass of the star. The term $\delta\nu_{n\ell}$ is the *small frequency separation*, or simply the *small separation* and it is sensitive to the core condensation, hence the age of the star (e.g. see [317] for the case of solar-like oscillations). Different small separations are adopted when analyzing the observations, and their investigation in relation to $\Delta\nu$ will be discussed in more detail in Chapter 4 for cool stars observed in open clusters.

The periods of g modes, asymptotically given by

$$\Pi_{n\ell} = \frac{\Pi_0}{\sqrt{\ell(\ell+1)}}(n + \alpha)\,, \tag{1.12}$$

are nearly uniformly spaced; here n and ℓ are again the radial order and angular degree of the mode, α is a small constant, and Π_0 is given by

$$\Pi_0 = 2\pi^2 \left(\int \frac{N}{r} dr \right)^{-1}\,, \tag{1.13}$$

where N is the Brunt-Väisälä frequency given by

$$N^2 = g\left(\frac{1}{\Gamma_1 p}\frac{dp}{dr} - \frac{1}{\rho}\frac{d\rho}{dr}\right) \tag{1.14}$$

Γ_1 being the adiabatic exponent

$$\Gamma_1 = \left(\frac{\partial \ln p}{\partial ln\rho}\right)_{\rm ad}, \tag{1.15}$$

and the integral is over the cavity in which the g mode propagates, as in Figure 1.4(b). Deviations of the period spacing for g modes are used to diagnose stratification in stars, since strong mean molecular weight gradients trap modes and cause deviations from the simple asymptotic relation given in Eq. (1.12). This technique has been particularly successful in measuring the stratification in white dwarf atmospheres with carbon-oxygen cores and layers of helium and hydrogen above (see [1], Chapter 7). Period spacings observed in red giant stars are also extremely useful for distinguishing between H-shell and He-core burning phases (see e.g. [32, 245, 87] and the mixed modes analysis presented in Chapter 4).

1.3 Asteroseismology across the HR Diagram

Figure 1.6 shows a power spectrum of the radial velocity variations (see Chapter 2 for a definition) observed over a time span of 9.5 years for the Sun by BiSON, the **Bi**rmingham **S**olar **O**scillation **N**etwork[3]. This shows the set of frequencies expected from Eq. (1.10) for high overtone, low degree ($n \gg \ell$) p modes. It is noteworthy that the comb of frequencies consists of alternating even and odd ℓ-modes, as expected from Eq. (1.10), where it can be seen that (to first order) modes of (n, ℓ) and $(n-1, \ell+2)$ have the same frequency. It is the small separation, $\delta\nu$, that lifts this degeneracy. This effect can be seen in Figure 1.8 which is a portion of an amplitude spectrum of the radial velocity variations of the Sun seen as a star made by the GOLF (**G**lobal **O**scillation at **L**ow **F**requencies[4]) experiment on SOHO (**SO**lar and **H**eliospheric **O**bservatory[5]) orbiting at the Earth-Sun L_1 Lagrangian point. Here it can be seen that while the large separations for even and odd ℓ-modes (reported as $\Delta\nu_0$, $\Delta\nu_1$) are very similar, the small separation lifts the degeneracy between paris of modes having (n, ℓ) and $(n-1, \ell+2)$. We also notice that there is a substantial difference between the small separations for even and odd ℓ-modes (i.e. $\delta\nu_{13}, \delta\nu_{02}$).

Figure 1.7 shows a spectrum similar to that of Figure 1.6 but corresponding to 30 days of photometric observations taken by the VIRGO instrument on the SOHO spacecraft. Again the regularity of the peaks corresponding to p-mode oscillations is evident, though they are rising above a sloping background which is caused by granulation effects in the surface layers of the star. In Section 1.3.3 we will discuss more about the differences between photometric and spectroscopic measurements of the oscillation signal.

Ultimately, it is the goal of asteroseismology for any star to detect enough frequencies over ranges in n, ℓ and possibly m that the interior sound speed may be mapped with precision, so that deductions can be made about interior temperature, pressure, density, chemical composition and rotation. A fundamental step along the way is to resolve a sufficient number of frequencies in a star, and to identify the modes associated with them unambiguously such that

[3] The official BiSON website can be found at http://bison.ph.bham.ac.uk/.
[4] The official website of the GONG experiment can be found at http://golfwww.medoc-ias.u-psud.fr/.
[5] The NASA's SOHO space mission can be found at http://sohowww.nascom.nasa.gov/.

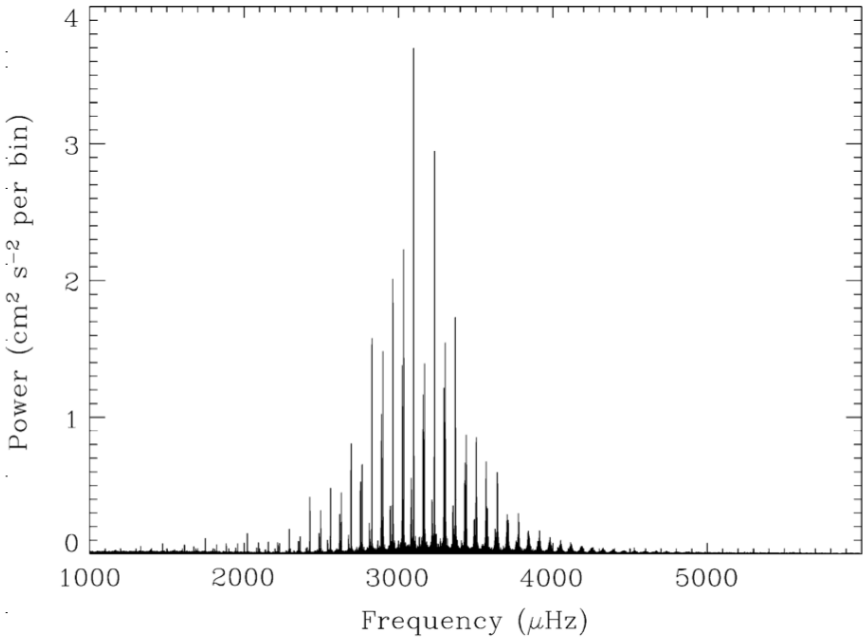

Figure 1.6: *A power spectrum of radial velocity variations in the Sun seen as a star for 9.5 years of data taken with the Birmingham Solar Oscillation Network (BiSON) telescopes. The equivalent amplitude noise level in this diagram is 0.5 mm s^{-1}. Image copied from [1].*

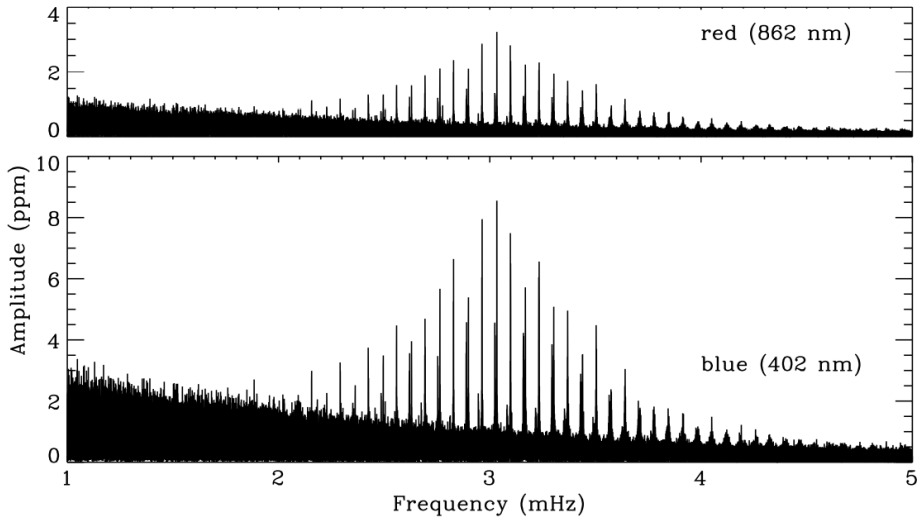

Figure 1.7: *Amplitude spectra of full-disk solar oscillations measured in intensity by the VIRGO instrument on the SOHO spacecraft. The observations are smoothed and rescaled here to show the spectrum corresponding to 30 days. Individual oscillation modes appear as strong peaks rising above a sloping background, which arises from random convective motion on the solar surface. Image copied from [1].*

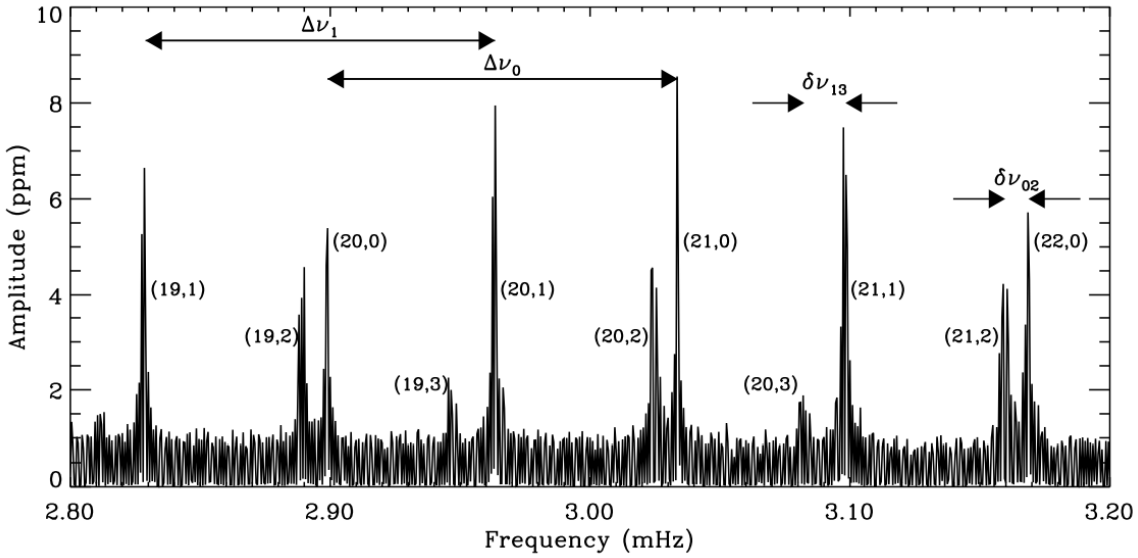

Figure 1.8: *This amplitude spectrum of radial velocity variations observed with the GOLF instrument on SOHO clearly shows the large and small separations in the p modes of the Sun. The radial order and angular degree of the modes are also included. Image copied from [1].*

the large and small separations may be deduced with confidence[6]. That step alone leads to determinations of the fundamental parameters of mass and age for some kinds of stars. A practical example of this method is given in Chapter 2 where the asteroseismic analysis of two solar-like subgiant stars is presented.

Figure 1.9 shows an "asteroseismic HR Diagram" [75] where the large separation clearly is a measure of mass (largely because of the relationship between mass and radius), and the small separation is most sensitive to the central mass fraction of hydrogen, hence age. Now that many solar-type oscillators have been found, it is possible to begin to model them using the large and small separations (e.g. see [316, 317]). The pattern of high overtone even and odd ℓ modes is also observed in some roAp stars, although their interpretation for those stars is more complex because of the strong effects of their global magnetic fields on the frequency separations.

Figure 1.10 shows a black-and-white version of the so-called "pulsation HR Diagram". The g-mode pulsators are common amongst other types of stars — even some, the γ Dor stars, that are not very much hotter than the Sun and are overlapping with the solar-like oscillators, keeping hope alive that g modes may eventually be detected with confidence in the Sun. There are three places in Figure 1.10 where there are p-mode and g-mode pulsators of similar stellar structure: for the β Cep (p-mode) and Slowly Pulsating B (SPB; g-mode) stars on the upper main sequence; for the δ Sct (p-mode) and γ Dor (g-mode) stars of the middle main sequence; and for the EC 14026 subdwarf B variables (p-mode) and the PG 1716+426 stars (g-mode). Stars pulsating in both p modes and g modes promise particularly rich asteroseismic views of their interiors.

[6] A remarkable case of ambiguous mode identification is that concerning the hot F type stars that show solar-like oscillations (the star Procyon among the most famous). This issue has been recently solved by adopting correlations between the asymptotic quantities appearing in Eq. (1.10) and the effective temperature of the stars, see [318].

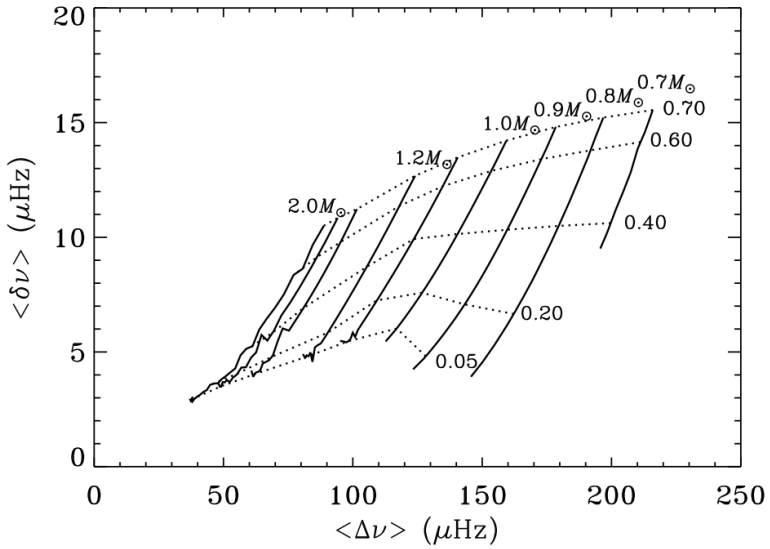

Figure 1.9: *An asteroseismic HR Diagram in which the large separation $\Delta\nu$ is most sensitive to mass, and the small separation $\delta\nu$ is most sensitive to age. The solid, nearly vertical lines are lines of constant mass, and the nearly horizontal dashed lines are isopleths of constant hydrogen mass fraction in the core, at the values indicated in the figure. Image copied from [1].*

1.3.1 The Driving Mechanisms

It is first worth to say that not all stars pulsate. It would be interesting to know whether all stars would be observed to pulsate at some level, if only we had the precision to detect those pulsations. At the present time, and with the level of the precision of our observations of μmag in photometry and $\mathrm{cm\,s^{-1}}$ in radial velocity, we can say that some stars do not pulsate.

In the longest known case of a pulsating star, that of *o* Ceti (Mira), it has been pulsating for hundreds of years, at least. In many other cases there are good light curves going back over a century, so that we know that stellar pulsation is a relatively stable phenomenon in many stars. That means that energy must be fed into the pulsation via what are known as *driving mechanisms*.

For most of the interior of the star, energy is lost in each pulsation cycle, i.e., most of the volume of the star *damps* the pulsation. The observed pulsation can only continue, therefore, if there is some part of the interior of the star where the energy fed into the pulsation is as much as that damped throughout the rest of the bulk of the star. In the following we briefly illustrate the three main driving mechanisms responsible of generating oscillations in all the main classes of pulsators.

There is a region in the star, usually a radial layer, that gains heat during the compression phase of the pulsation cycle and is able to drive the pulsation. All other layers that lose heat during the compression phase are responsible of damping the pulsation instead. If this layer succeeds in driving the oscillation, then the star functions as a heat engine, converting thermal energy into mechanical one; thus we refer to this type of driving as a heat-engine mechanism. For Cepheid variables, RR Lyrae stars, δ Sct stars, β Cep stars — as for most of the pulsating variables seen in Figure 1.10 — the driving mechanism is connected with the opacity, thus it is known as the κ mechanism. Simplistically, in the ionization layers for H and He opacity blocks radiation, the gas heats and the pressure increases causing the star to swell. But the ionization of the gas reduces the opacity, radiation flows through, the gas cools and can no longer support

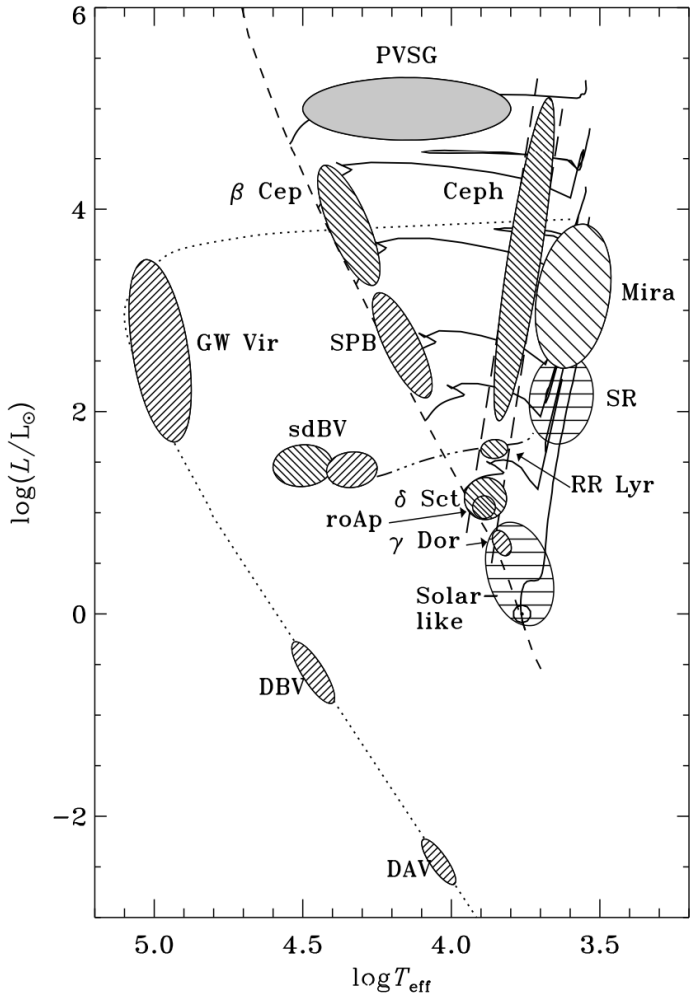

Figure 1.10: *A pulsation HR Diagram showing many classes of pulsating stars for which asteroseismology is possible. Tilted lines are oriented according to the nature of the modes observed (left orientation for* p *modes, right orientation for* g *modes). Image copied from [1].*

the weight of the overlying layers, so the star contracts. On contraction the H or He recombines and flux is once more absorbed as opacity is increased again, hence the condition for a heat engine is present: the layer gains heat on compression.

The other major driving mechanism that operates in the Sun and solar-like oscillators, as well as some pulsating red giant stars, is the *stochastic* driving, which identifies the class of pulsators studied in this dissertation. Indeed, nonadiabatic calculations taking convection into account generally find that modes in stars on the cool side of the instability strip are stable, hence in this case the heat-engine mechanism is not able to drive the oscillations. In these stars the convective motion near the surface likely reaches speeds close to that of sound. Such turbulent motion with near-sonic speed is an efficient source of acoustic radiation for making the star resonate in some of its natural oscillation frequencies. This is due to the stochastic noise that is transferred to energy of global oscillation. Since the excitation is caused by a very large number of convective elements, the driving is essentially random.

As the stochastically excited modes are also intrinsically damped (see Chapter 2 for more details). The outcome is that the average power spectrum is therefore a Lorentzian spectrum, with a width determined by the linear damping rate of the mode [76]. If a single realization,

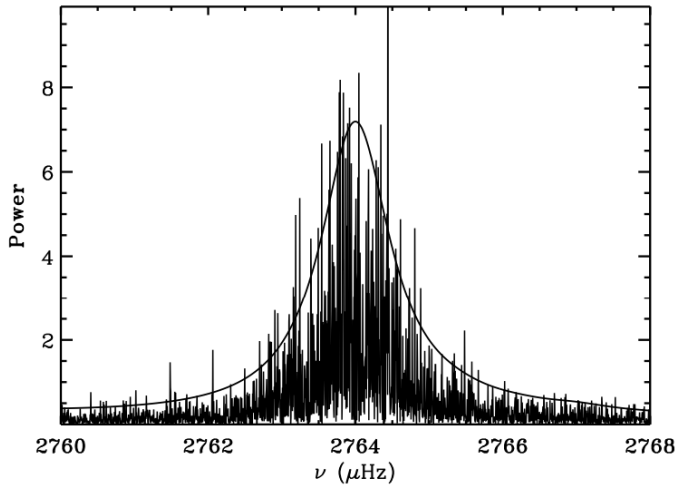

Figure 1.11: *Observed spectrum, from Doppler observations with the BiSON network, of a single radial mode of solar oscillations. The smooth curve shows the fitted Lorentzian profile, multiplied by three for clarity. Image copied from [76].*

rather than the average, of the spectrum is considered, as is generally the case for observations of stellar oscillations, the result is a random function with a Lorentzian envelope. An example is shown in Figure 1.11, based on observations of solar oscillations with the BiSON network. Such Lorentzian profiles are often assumed in the fits carried out to determine the frequency and other properties of the modes.

Lastly, the third major theoretical driving mechanism is the ϵ mechanism, where in this case ϵ is used to refer to the energy generation rate in the core of the star. Potentially, variations in ϵ could drive global pulsations. This has been discussed as a possible driving mechanism in some cases of evolved very massive stars, but there is no known class of pulsating stars at present that are thought to be driven by the ϵ mechanism alone.

1.3.2 Solar-like Oscillations in Cool Stars

It is fairly well established that the solar five-minute oscillations are intrinsically stable, and excited by the vigorous near-surface convection [13], that is the stochastic driving introduced above. Oscillations excited in this manner can therefore be called *solar-like*; since near-surface convection with speed approaching or exceeding the speed of sound is generally found in stars cooler than the Cepheid instability strip (e.g. see Figure 1.10), such stars may be expected to exhibit solar-like oscillations, unless other intrinsic excitation mechanisms are at work. This obviously includes stars which are otherwise quite different from the Sun, such as subgiants and giants. It should be noted that, even though the excitation mechanism is in principle the same as in the Sun, differences in internal structure can lead to substantial differences in the properties of the oscillation spectra, particularly for highly evolved stars (we refer to Chapters 3 and 4 for more details on solar-like oscillations in red giant stars and to Chapter 6 for a discussion about the oscillation amplitudes).

A major difficulty in observing solar-like oscillations, at least in main-sequence stars, relies on their expected very small amplitudes: in the solar case the largest amplitude per mode for radial-velocity observations is around $20\,\mathrm{cm\,s^{-1}}$, or around 5 ppm (parts per million) in broad-band intensity observations. Thus definite detection of oscillations analogous to the solar five-minute oscillations has been an elusive goal for a long time. In the last few years, however,

major breakthroughs have been achieved, largely thanks to the development of very stable techniques for radial-velocity observations and of space-based photometric missions aimed at detection of extra-solar planets (see Chapters 2 and 3).

Solar-like oscillations are expected in all cool stars. An important feature of these pulsations is that the envelope of the power excess in the acoustic spectrum moves to lower frequencies as the star evolves from main sequence to the red giant phase. The main asteroseismic parameter that is related to this effect is the *frequency of maximum power* $\nu_{\max}$. As demonstrated by [58], $\nu_{\max}$ scales with the cut-off frequency of the star, hence, $\nu_{\max} \propto g/\sqrt{T_{\rm eff}}$. This correlates $\nu_{\max}$ to the fundamental stellar properties of mass, radius and effective temperature. In Chapter 3 it will be shows how the (asteroseismic) mass and radius of a star can be derived by simply combining the measurements of $\nu_{\max}$ and $\Delta\nu$.

It is however important to note that the properties of solar-like oscillations change substantially with the changes in stellar internal structure, particularly following hydrogen exhaustion in the core, i.e. the beginning of the subgiant phase. The resulting core contraction greatly increases the local gravitational acceleration in the deep interior of the star. Furthermore, gradients in hydrogen abundance, associated with a hydrogen shell-burning source and possibly the presence of a shrinking convective core in the main-sequence phase, give rise to an increase of the buoyancy frequency N, Eq. (1.14). As a result, the acoustic oscillations may have a g-mode character in the deep interior, which arises when p and g modes undergo so-called *avoided crossings* (see Chapter 4 for further details). The analysis of these *mixed modes* is of great importance for gaining knowledge about the very deep interior of red giant stars, which also makes these targets one of the most interesting class of pulsators for testing of stellar evolution and stellar oscillation theory.

1.3.3 Observational Techniques

For completing this general introduction to asteroseismology, we present a summary concerning the observational techniques used to attain data of stellar variability. As discussed by [185, 25], stellar oscillations can be observed in the stellar atmosphere in three ways: (i) velocity shifts of the spectral lines (spectroscopy), (ii) variations in the total intensity (photometry), (iii) variations in the equivalent widths of temperature-sensitive lines (spectroscopy). All these techniques require coverages of the targets as continuous as possible in order to reduce the confusion from aliases in the spectral window (see Chapter 2 for more discussion). In the following we briefly describe the main features, advantages and disadvantages of each of these observational techniques. We also stress that the first two techniques are those used for the entire analysis presented in this thesis.

Velocity Measurements

Before the advent of the space-based missions *CoRoT* and *Kepler* (see below), most of the results of the asteroseismology were based on velocity measurements obtained using high-dispersion spectrographs with stable reference sources. The great improvement in Doppler precision over the past years was a direct result of programs aimed at detecting planets around other stars. Asteroseismology has benefited considerably from these advances. In a near future, the *SONG* project [142] is expected to improve dramatically the quality of spectroscopic observations for asteroseismology by exploiting a ground-based network of telescopes that allows for continuous observations of targets.

Velocity measurements have the important advantage of being more sensitive to modes having angular degree $\ell \geq 2$ (e.g. see [76]). The reason is that the observations have some

spatial resolution of the stellar disk, and so the tendency for high-degree modes to cancel is reduced. Thus, we measure velocities projected onto the line of sight, which gives more sensitivity to the centre of the disk relative to the limb. In addition, Doppler shift observations are far less sensitive to the stellar background (mainly originated from the granulation effects) than photometric observations, which is a very important advantage.

However, velocity observations have the disadvantage of suffering from pronounced effects of stellar rotation, which prevent to attain high precision due to line broadening, hence they cannot be used for stars that show a high rotational velocity (e.g. see the case of η Bootis studied by [189]). The observations also require high-resolution spectroscopy with extremely high precision, which is only achievable on a handful of instruments from ground. This clearly introduces further problems related mainly to Earth's atmosphere conditions and to the day and night alternation due to Earth's rotation, which limits significantly the continuity of the observations from a single site (see Chapter 2 for more details). Lastly, spectroscopy observations are, in general, single target dedicated, i.e. only one target per time is usually observed from a single operating site.

Intensity Measurements

Observing the brightness variations on the stellar's surface shows three important advantages: it can be done with an extremely simple instrumentation; the acquired photons span a wide range of wavelengths; it can be made simultaneously on many stars (although to some extent, multi-object spectroscopy also offers this possibility). Unfortunately, scintillation from the Earth's atmosphere severely limits the precision of ground-based photometry, but thanks to the simplicity of the instrumentations, space-based missions can operate as well. In fact, space is the ideal place to make intensity measurements, given the absence of atmospheric scintillation and the possibility for long periods of uninterrupted observations. The most successful space missions designed specifically for asteroseismology are *MOST* [221] (launch date 2003), *CoRoT* [11, 229] (launch date 2006) and *Kepler* [50, 193] (launch date 2009). The latter in particular is providing an enormous amount of high-precision light curves that span up to over more than 3 years of continuous observations (see Chapter 3 for an introduction and first results of the NASA's *Kepler* Mission).

Intensity measurements detect temperature variations, as the brightness follows a law $\propto T_{\rm eff}^4$, and so are much more sensitive to the stellar background that arises from granulation. This is seen in intensity observations of the Sun (see Figure 1.7) as background power rising towards low frequencies. Although this power contains information about stellar convection that might usefully be compared with models (see e.g. [301, 217]), for asteroseismology it represents an unwanted and fundamental noise source. This constitutes a significant disadvantage with respect to the velocity measurements.

Equivalent Width Measurements

This method for detecting oscillations was suggested by [189] and used for the first time in the study of η Bootis. It involves monitoring changes in spectral lines whose equivalent widths (EWs) are temperature sensitive, most notably the hydrogen Balmer lines. Similarly to the velocity measurements, EW measurements are more sensitive to modes with $\ell \geq 2$ than intensity measurements. Conversely to velocity measurements, at least for the Balmer lines, the centre-to-limb variation of line strengths gives a similar effect to EW measurements.

As for the photometric case, EW measurements are sensitive to temperature changes, hence they suffer from the granulation noise at low frequencies and are, in general, more difficult to be

obtained than photometric observations because they require spectroscopic techniques. Lastly, this technique has the advantage of being more useful for detecting oscillations in rapid rotators, where instead the velocity measurements become less precise.

2 Solar-like Oscillations in Subgiant Stars

The search for solar-like oscillations (see [77, 107] for a summary) in main-sequence and subgiant stars showed a tremendous growth in the last decade (for reviews see e.g. [25, 26]), especially by means of the photometric space-based missions *CoRoT* [11, 229] and *Kepler* [50, 193]. The latter in particular is presently providing an enormous amount of high-quality asteroseismic data (e.g. see [127] for more details on the KAI - *Kepler* Asteroseismic Investigation and Chapter 3). Photometric studies of a large number of solar-type stars are fundamental for statistical investigations of intrinsic stellar properties and for testing theories of stellar evolution [71]. However, as is known from the theory of solar-like oscillations, high-precision Doppler shift measurements are more effective for detecting p modes of higher angular degrees (see Chapter 1). At the present time the échelle spectrometers such as CORALIE, HARPS, UCLES, UVES, SOPHIE and SARG, attaining high-precision radial velocity (RV) measurements [211], offer a way to detect solar-like oscillations in bright asteroseismic targets (e.g. see [27] for a review on ground-based campaigns). The spectroscopic approach to the detection of solar-like oscillations will also be used in the ground-based *SONG* project [142] in the near future.

In the first part of this chapter we provide a general introduction to the Fourier analysis method adopted for analyzing the RV time-series, together with some definitions and techniques that are helpful for the classical asteroseismic study conducted in this dissertation.

Afterwards we describe the detection of excess of power and a detailed analysis of the solar-like oscillations observed in the evolved subgiant star β Aquilae (HR 7602, HD 188512, HIP 98036) [86], by exploiting RV measurements and a measure of its parallax obtained by Hipparcos. In particular, in Section 2.3.1 we describe the observations, the data reduction done by means of the standard IRAF[1] package facilities and the RV extraction developed for the first time with the iSONG code, an IDL[2] based software to be used within the *SONG* project [142] for an automated extraction of RV measurements from star's spectra. The Fourier analysis and the mode identification process that led us to the global asteroseismic quantities are presented in Sections from 2.3.3 to 2.3.7. In Section 2.4 we discuss the scaled mass, mode amplitudes, mode lifetimes, and frequency of maximum power of the star, where a comparison with expectations and a global list of stellar parameters also derived by means of the SEEK package [260] is presented.

In the last part of the chapter (Section 2.5) we show the new results derived from a revised analysis of the solar-like oscillations in the G5 subgiant μ Herculis [48] (HR 6623). We show

[1]IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of the Universities for Research in Astronomy, inc. (AURA) under cooperative agreement with the National Science Foundation.

[2]IDL (Interactive Data Language) is a registered trademark of ITT Visual Information Solutions.

how the RV time-series derived by means of the iSONG code from the same spectra used by [48], suffer from less systematic effects than the old dataset derived by means of the AUSTRAL code [110]. For this case, the raw spectra were reduced with the REDUCE IDL package [258].

The preliminary analyses described in this chapter show the interesting potential of the iSONG code for extractions of RV measurements from high-precision wavelength-calibrated stellar spectra. The conclusions are drawn in Section 2.6.

2.1 Fourier Analysis

In this section we introduce the Fourier analysis for the calculation of the so-called power spectrum, the main tool for the asteroseismic analysis discussed here. In fact, seismic analyses usually deal with time-series in the time domain, then transformed into a signal in the frequency domain to be used for deriving the oscillation properties[3]. The simple Fourier transform in the frequency domain is termed acoustic spectrum[4]. Here we plot the frequency in abscissa and the oscillation amplitudes in ordinate. This amplitude spectrum can be computed in different manners (see also [181]), the Fast-Fourier Transform (FFT) [123, 259] and the Discrete Fourier Transform (DFT, mostly known as the Lomb-Scargle periodigram [208, 272]) being the most known methods. These algorithms converge to the same spectrum for evenly sampled data without gaps, but keeping in mind that the scale of the amplitudes is not always the same: this aspect represents a problem when comparing data that have been analyzed with different methods.

For the analysis presented in this chapter (and also in Chapter 3 concerning the ensemble studies of *Kepler* targets) we used the so called *Least-Squares Spectrum* (LSS), which is defined so that a sine wave with an unitary amplitude in the time domain shows a peak having an unitary amplitude when transformed to the amplitude spectrum, and it has unitary area in the power-density spectrum (see Section 2.1.4 for a definition). The LSS differs from the amplitude spectrum derived by either the FFT or the Lomb-Scargle periodigram [208, 272] by a normalization factor only. In the following we describe the LSS method and present its main advantages compared to the other techniques mentioned above.

2.1.1 The Least-Squares Spectrum

Let us assume we have a set of N observations $x_0, x_1, \ldots, x_{N-1}$, having zero mean and sampled at times $t_0, t_1, \ldots, t_{N-1}$, respectively[5]. The observations that can be used can be in the form of either light curves or RV time-series, the latter being the case of the analysis presented in this chapter. Thus, one can set up a model of the observations as follows

$$x_n^{\mathrm{th}} = \alpha \cos(2\pi\nu t_n) + \beta \sin(2\pi\nu t_n)\,, \tag{2.1}$$

which assumes that our time-series consists of a single sine wave. By means of simple trigonometry formula, the model can be rewritten as

$$x_n^{\mathrm{th}} = A \sin(2\pi\nu t_n + \delta)\,, \tag{2.2}$$

[3]Studying the asteroseismic properties of a star directly from its time-series is however possible, see e.g. [54]. Nonetheless, description of the methods that do not involve the Fourier analysis techniques goes beyond the scope of this work.

[4]It is often common referring to the power spectrum, which is derived as the squared modulo of the amplitude spectrum.

[5]The zero mean dataset is given as an example for simplifying the calculations and it can be applied to any time-series in general by simply removing any offset given by the position of its mean value. A more general introduction to the LSS that considers the offset as additional parameter can be found in [1].

where $A = \sqrt{\alpha^2 + \beta^2}$ and $\delta = \tan^{-1}(\alpha/\beta)$. In a LSS the coefficients α and β are calculated by minimizing the following equation

$$\begin{aligned} R(\nu_k) &= \sum_{n=0}^{N-1} \left(x_n - x_n^{\text{th}}\right)^2 \\ &= \sum_{n=0}^{N-1} \left\{x_n - [\alpha \cos(2\pi\nu_k t_n) + \beta \sin(2\pi\nu_k t_n)]\right\}^2 , \end{aligned} \tag{2.3}$$

which basically represents a least-squares method applied to the difference between model and observations. This consists in solving the system of two equations

$$\frac{\partial R}{\partial \alpha} = 0 , \quad \frac{\partial R}{\partial \beta} = 0 . \tag{2.4}$$

This yields the system of equations represented by

$$\begin{bmatrix} c_{cc}(\nu_k) & c_{cs}(\nu_k) \\ c_{cs}(\nu_k) & c_{ss}(\nu_k) \end{bmatrix} \begin{bmatrix} \alpha(\nu_k) \\ \beta(\nu_k) \end{bmatrix} = \begin{bmatrix} y_c(\nu_k) \\ y_s(\nu_k) \end{bmatrix} \tag{2.5}$$

where, following [181], we used the notations:

$$\begin{aligned} y_c(\nu_k) &= \sum_{n=0}^{N-1} x_n \cos(2\pi\nu_k t_n) , \\ y_s(\nu_k) &= \sum_{n=0}^{N-1} x_n \sin(2\pi\nu_k t_n) , \\ c_{cc}(\nu_k) &= \sum_{n=0}^{N-1} \cos^2(2\pi\nu_k t_n) , \\ c_{ss}(\nu_k) &= \sum_{n=0}^{N-1} \sin^2(2\pi\nu_k t_n) , \\ c_{cs}(\nu_k) &= \sum_{n=0}^{N-1} \cos(2\pi\nu_k t_n) \sin(2\pi\nu_k t_n) . \end{aligned} \tag{2.6}$$

Thus, the system of Eq. (2.5) can be denoted as

$$\mathbf{C}\mathbf{a} = \mathbf{y} , \tag{2.7}$$

$\mathbf{C}$ being the matrix of the terms c_{cc}, c_{ss}, c_{cs} given by Eq. (2.6), $\mathbf{a}$ the vector of the coefficients α, β to be determined, and $\mathbf{y}$ the vector of the terms y_c, y_s, Eq. (2.6). Therefore, the solution to the system (2.5) is given by

$$\mathbf{a} = \mathbf{C}^{-1}\mathbf{y} , \tag{2.8}$$

namely

$$\begin{aligned} \begin{bmatrix} \alpha(\nu_k) \\ \beta(\nu_k) \end{bmatrix} &= \begin{bmatrix} c_{cc}(\nu_k) & c_{cs}(\nu_k) \\ c_{cs}(\nu_k) & c_{ss}(\nu_k) \end{bmatrix}^{-1} \begin{bmatrix} y_c(\nu_k) \\ y_s(\nu_k) \end{bmatrix} \\ &= \begin{bmatrix} c_{ss}(\nu_k)/\Delta(\nu_k) & -c_{cs}(\nu_k)/\Delta(\nu_k) \\ -c_{cs}(\nu_k)/\Delta(\nu_k) & c_{cc}(\nu_k)/\Delta(\nu_k) \end{bmatrix} \begin{bmatrix} y_c(\nu_k) \\ y_s(\nu_k) \end{bmatrix} , \end{aligned} \tag{2.9}$$

where $\Delta(\nu_k) = c_{cc}(\nu_k) \cdot c_{ss}(\nu_k) - c_{cs}^2(\nu_k)$. The coefficients can be then written as

$$\alpha(\nu_k) = \frac{c_{ss}(\nu_k) y_c(\nu_k) - c_{cs}(\nu_k) y_s(\nu_k)}{c_{cc}(\nu_k) c_{ss}(\nu_k) - c_{cs}^2(\nu_k)} \,, \tag{2.10}$$

$$\beta(\nu_k) = \frac{c_{cc}(\nu_k) y_s(\nu_k) - c_{cs}(\nu_k) y_c(\nu_k)}{c_{cc}(\nu_k) c_{ss}(\nu_k) - c_{cs}^2(\nu_k)} \,. \tag{2.11}$$

Hence, given a frequency value ν_k, the amplitude in the LSS is given by

$$A(\nu_k) = \sqrt{\alpha^2(\nu_k) + \beta^2(\nu_k)} \,. \tag{2.12}$$

A great advantage of the LSS method with respect to the FFT one is that the spectrum can be calculated using statistical weights. By assigning a statistical weight w_n to each data point x_n, usually $w_n = \sigma_n^{-2}$ with σ_n the corresponding uncertainty (see Section 2.3.1 for more discussion), Eq. (2.3) can be rewritten as

$$\begin{aligned} R(\nu_k) &= \frac{1}{W} \sum_{n=0}^{N-1} w_n \left(x_n - x_n^{\mathrm{th}} \right)^2 \\ &= \frac{1}{W} \sum_{n=0}^{N-1} w_n \left\{ x_n - [\alpha \cos(2\pi\nu_k t_n) + \beta \sin(2\pi\nu_k t_n)] \right\}^2 \,, \end{aligned} \tag{2.13}$$

where

$$W = \sum_{n=0}^{N-1} w_n \,. \tag{2.14}$$

The solution to Eq. (2.13) is obtained in the same way as for Eq. (2.3) but adopting the terms

$$\begin{aligned} y_c(\nu_k) &= \sum_{n=0}^{N-1} w_n x_n \cos(2\pi\nu_k t_n) \,, \\ y_s(\nu_k) &= \sum_{n=0}^{N-1} w_n x_n \sin(2\pi\nu_k t_n) \,, \\ c_{cc}(\nu_k) &= \sum_{n=0}^{N-1} w_n \cos^2(2\pi\nu_k t_n) \,, \\ c_{ss}(\nu_k) &= \sum_{n=0}^{N-1} w_n \sin^2(2\pi\nu_k t_n) \,, \\ c_{cs}(\nu_k) &= \sum_{n=0}^{N-1} w_n \cos(2\pi\nu_k t_n) \sin(2\pi\nu_k t_n) \,, \end{aligned} \tag{2.15}$$

instead of those given by Eq. (2.6) (see [112]). Another important and very useful feature of the LSS method is that it can be calculated also for unevenly sampled data, as happens in most of the cases, especially for spectroscopic observations from ground sites.

At this point, we briefly enhance the main differences between the LSS method and the other methods mentioned above. First, in the normalization of the Lomb-Scargle periodigram the significance of a given peak in the spectrum has a central importance, whereas for the LSS method it is important to ensure that a sine wave with a unitary amplitude in the time domain has a corresponding unitary amplitude in the frequency domain. In order to evaluate when the

significance of a peak in the power spectrum is relevant, it is common to adopt the 4σ criteria given by [52]. This empirical criteria has been tested over many ground-based datasets [52].

Second, as also stressed in [181], highlighting the difference between a typical Fourier transform and the LSS is of importance for a deeper understanding of the method. The former technique allows one to derive a series of values $X_0, X_1, \ldots, X_{N-1}$ in the frequency domain by transforming a series of observations $x_0, x_1, \ldots, x_{N-1}$ having zero mean in the time domain, according to the relation

$$X_k = \sum_{n=0}^{N-1} x_n \exp\left(-\frac{2\pi i}{N} kn\right) \tag{2.16}$$

with $X_k = X(\nu_k) \in \mathbb{R}$, and ν_k the corresponding frequency. The Fourier transform fulfills the Parseval's theorem by definition, namely

$$\sum_{n=0}^{N-1} |x_n|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X_k|^2 \,, \tag{2.17}$$

which means that the total information contained in the time domain is preserved when passing to the frequency domain. This conservation follows from the fact that the possible natural frequencies ν_k are only in a well defined number, i.e.

$$\nu_k = \frac{2\pi n}{T} \,, \qquad k = -N/2, \ldots, +N/2 \,, \tag{2.18}$$

with N again the total number of data points, and T the total duration of the observing run. Therefore, it follows that the natural frequencies are only given for evenly sampled data and the Parseval's theorem is not fulfilled for the case of unevenly sampled observations, which also include time-series with gaps. In fact, as stated by [108], unevenly sampled data have neither a well defined Nyquist frequency nor a set of natural frequencies. The main consequence is that particular care must be paid when using statistics for the modeling of the observations, because the resulting acoustic spectra may not be χ^2 distributed with two degrees of freedom, according to the statistic of an acoustic spectrum [319].

2.1.2 Aliasing

One of the main problems of time-limited and/or unevenly sampled observations is that given a frequency of a real oscillation signal, there will be always more than a single peak (sidebands or sidelobes) in the amplitude spectrum that corresponds to the same sinusoidal contribution. This effect is the well known aliasing phenomenon of the theory of signal processing.

To understand the problem, let us assume to have a single sinusoidal oscillation with frequency ν_0, phase δ_0, and unitary amplitude, given by

$$W(t) = \sin(2\pi\nu_0 t - \delta_0) \,. \tag{2.19}$$

Its corresponding amplitude spectrum will be hereafter denoted as the *window function*, or equivalently *spectral window* if one refers to the squared modulus [190]. As calculated by [76], the functional form of the window function approximately reads

$$W(\nu) \propto T \mathrm{sinc}\left[\pi T(\nu - \nu_0)\right] \,, \tag{2.20}$$

where T is the total observing time and

$$\mathrm{sinc}(x) = \frac{\sin(x)}{x} \,, \tag{2.21}$$

is the sinc function (see Figure 2.1). Thus, the spectral window $S(\nu)$ is expressed as

$$S(\nu) \propto T^2 \mathrm{sinc}^2 \left[\pi T(\nu - \nu_0)\right] . \tag{2.22}$$

In practice the peak corresponding to the single frequency ν_0 has a more complex structure than the one described by the window function due to observational noise and fluctuations in the oscillation amplitude.

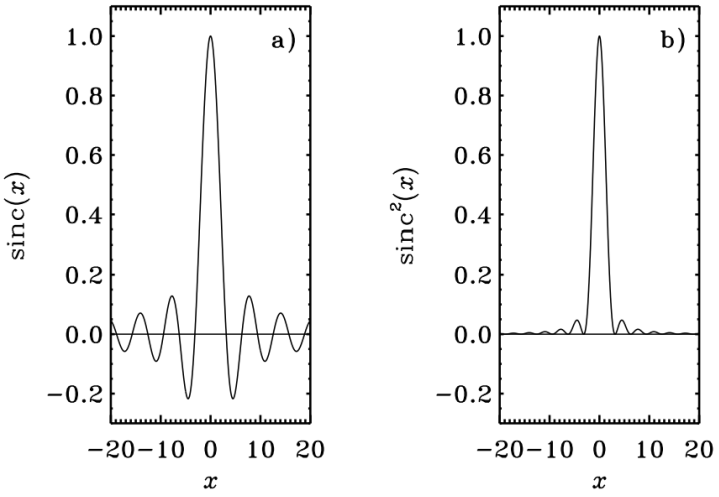

Figure 2.1: *The sinc function (a) and the sinc^2 function (b). Image copied from [76].*

In addition data can from gaps caused by the day-night alternation (corresponding to a frequency range of $11.57\,\mu$Hz) as it happens for ground-based observations from single site. Then, following [76], the window function is approximately given by

$$W(\nu) \propto T \cos\left[\pi\tau(\nu - \nu_0)\right] \mathrm{sinc}\left[\pi T(\nu - \nu_0)\right] , \tag{2.23}$$

where τ is the duration of the gap in the data, while the spectral window is again expressed as $S(\nu) = |W(\nu)|^2$. Both Eqs. (2.20) and (2.23) reduce to a delta of Dirac for $T \to \infty$.

The integral of the spectral window, that is the integral in power of the window function, represents the reciprocal of the "effective" observing time of a time-series with either unevenly sampling or gaps, or both of them. This integral is known as the *spectral resolution* of the time-series. Clearly, the effective observing time will be always lower (or equal at most) to the total observing time. In conclusion, the presence of big gaps can distort the real amplitude of a sinusoidal signal as measured in the amplitude spectrum because, according to Eq. (2.23), the area under the peak profile would be spread into a wider range of frequencies.

2.1.3 Damping of the Modes

As stressed in [76], a further complication to the case described in Section 2.1.2 is represented by an additional damping term appearing in the single sinusoidal component. The new signal we consider is

$$A(t) = A_0 sin(2\pi\nu_0 t - \delta_0) e^{-\eta t} , \tag{2.24}$$

where η is the *damping rate*. The reciprocal of the damping rate is known as the *mode lifetime*, and can be found in literature as the quantity τ expressed in days. The resulting power

spectrum as a function of the angular frequency $\omega = 2\pi\nu$, where ν is the so-called *cyclic frequency*, assuming to observe the signal for a time $T \to \infty$, is expressed as

$$P(\omega) \propto \frac{A_0^2}{(\omega - \omega_0)^2 + \eta^2}\,, \tag{2.25}$$

where $\omega_0 = 2\pi\nu_0$. Thus, $P(\omega)$ is a *Lorentzian profile* having HWHM $= \eta$. If we express $P(\omega)$ in terms of the cyclic frequency, we have

$$P(\nu) \propto \frac{A_0^2}{(\nu - \nu_0)^2 + \left(\frac{\eta}{2\pi}\right)^2}\,, \tag{2.26}$$

that is a *Lorentzian profile* having HWHM $= \eta/(2\pi)$. This let us introducing another important quantity that characterizes solar-like oscillations, namely the *mode linewidth*, defined as the FWHM of Eq. (2.26), that is $\Gamma \equiv \eta/\pi$. The mode linewidth is directly related to the mode lifetime according to the relation $\Gamma = (\pi\tau)^{-1}$, and is usually expressed in μHz. As we will describe in Chapter 4, measuring the mode linewidths is important for our understanding of the damping mechanism of solar-like oscillations and can be used as a probe for estimating the temperature of the stars [87].

In the real case of a finite observing time T, the resulting shape of the peak in the power spectrum will be intermediate between the sinc^2 function and the Lorentzian one. In particular, the profile will tend to the former for $\eta T \ll 1$, and toward the latter for $\eta T \gg 1$ (see [76] for further details). However, this is not yet a real case of oscillation because we are implicitly assuming a sudden excitation of the mode. For solar-like oscillations, as it appears for the Sun, the oscillations are excited stochastically (see Chapter 1) and it can be shown that this effect originates a number of sharp peaks, whose distribution surrounds the general Lorentzian profile of the mode peak (e.g. see [76]). Thus, this process, combined with the exponential decay, generates a spectrum that on average has a Lorentzian profile.

2.1.4 The Power Density Spectrum

At this stage, defining a spectrum that is independent of data sampling (both uneven sampling and data gaps) can be of great utility, especially when ground-based observations carried out from different sites need to be merged into a single dataset (see e.g. [29, 9, 31]). The so-called *power density spectrum* (PSD) fulfills this requirement and it is defined so that a sine wave with unitary amplitude in the time domain produces a window function with unitary area in the frequency domain. Intuitively, the PSD can be computed by normalizing the power spectrum to the value of the spectral resolution introduced above, namely

$$\mathrm{PSD}(\nu) = \frac{P(\nu)}{\int_0^{\nu_{\mathrm{Nyquist}}} S(\nu) d\nu}\,, \tag{2.27}$$

where $P(\nu)$ is the power spectrum and the spectral resolution is calculated as the integral of the spectral window $S(\nu) = |W(\nu)|^2$ from 0 to the Nyquist frequency ν_{Nyquist}. Although for unevenly sampled data the Nyquist frequency is not defined, we are still able to introduce a limit frequency for the integral of the spectral window, that is

$$\nu_{\mathrm{Limit}} = \frac{1}{2\langle t \rangle}\,, \tag{2.28}$$

with $\langle t \rangle$ the average sampling time of the time-series. In practice this method works properly also because the spectral window goes rapidly to zero as the frequency increases and the integral

can be estimated correctly within the range $[0, \nu_{\rm Limit}]$. In fact, the window function for ground-based observations (e.g. those used for the analysis presented in this chapter) is typically represented by a sinc function having a number of significant sidebands not larger than 8 [48, 86]. Therefore, the PSD at a given frequency ν_k can be expressed analytically as

$$\mathrm{PSD}(\nu_k) = \frac{|A(\nu_k)|^2}{\int_0^{\nu_{\rm Limit}} S(\nu) d\nu} \tag{2.29}$$

where $A(\nu_k)$ is the LSS computed according to Eq. (2.12) (either weighted or unweighted). Since the spectral resolution is a constant for a given dataset, Eq. (2.29) can be applied to any frequency ν_k chosen in the sampling, bearing in mind that the window function $W(\nu)$ has to be computed with the same frequency sampling of the acoustic spectrum $A(\nu)$, as suggested by [181].

The PSD has the advantage of having amplitudes that are insensitive to the sampling of the data and the duration of the observation. This means that amplitudes measured in datasets with different sampling and observing length can be compared directly. The PSD has been widely used in literature for estimating the oscillation amplitudes in stars showing solar-like pulsations by means of the method described by [190, 191]. Such method is also adopted for the analysis of the oscillation amplitudes in both β Aquilae and μ Herculis described below.

2.1.5 The CLEAN algorithm

The CLEAN algorithm described in this section is analogous to the original algorithm introduced by [153] for aperture synthesis in radio interferometry. It was afterwards adapted by [266] for cleaning up the spectral window pattern for frequency analysis. It is the widest adopted algorithm for frequency analysis of acoustic spectra and it will be used in the context of the study presented in this chapter.

The first step in CLEANing an acoustic spectrum is to construct the *dirty* (i.e. observed) spectrum $A(\nu)$, which is the Fourier transform of the data without any manipulation. Secondly, one deconvolves the observed spectrum with the window function $W(\nu)$ shifted to the position of the highest peak appearing in the dirty spectrum. The deconvolution is usually done by first applying a particular scaling factor to the window function according to the gain factor g, with $0 < g < 2$. Thus, one subtracts the g-scaled spectral window from the dirty spectrum in order to produce a residual spectrum, $R(\nu)$. We have

$$A(\nu) = R(\nu) \otimes W(\nu - \nu_0) \tag{2.30}$$

where ν_0 is in this case the frequency of the highest peak in the observed spectrum (first iteration) or the highest peak in the residual spectrum (starting from second iteration), $W(\nu - \nu_0)$ is the window function centered at the peak to be CLEANed, and '$\otimes$' represents convolution. Hence the residual spectrum is formally given as

$$R(\nu) = A(\nu) \otimes^{-1} W(\nu - \nu_0)\,. \tag{2.31}$$

Let now assume we computed the LSS, hence we have the coefficients $\alpha(\nu)$ and $\beta(\nu)$ introduced in Section 2.1.1, or alternatively the amplitude $A(\nu)$ and the phase $\delta(\nu)$. A simple way to remove a peak at a given frequency from the observed spectrum is to subtract a sinusoidal signal from the time-series. In fact, assuming to have an oscillation of frequency ν_0 at a given time t_n

$$\begin{aligned} x_n^{\rm th} &= A_0 \sin(2\pi\nu_0 t_n + \delta_0) \\ &= \alpha(\nu_0)\cos(2\pi\nu_0 t_n) + \beta(\nu_0)\sin(2\pi\nu_0 t_n)\,, \end{aligned} \tag{2.32}$$

where $A_0 = A(\nu_0)$ and $\delta_0 = \delta(\nu_0)$. When subtracting it from the data we obtain

$$\begin{aligned} c_n = x_n - x_n^{\text{th}} &= x_n - A_0 \sin(2\pi\nu_0 t_n + \delta_0) \\ &= x_n - \alpha(\nu_0)\cos(2\pi\nu_0 t_n) - \beta(\nu_0)\sin(2\pi\nu_0 t_n)\,, \end{aligned} \tag{2.33}$$

where c_n is n-th point of the residual time-series. This method is very similar to the CLEAN algorithm mentioned above and it is known as Iterative Sine Wave Fitting (ISWF) [181]. The great advantage of this method, as well as for the CLEAN algorithm, is that it can be done iteratively, i.e. for an arbitrary number of frequencies. The ISWF algorithm is used in this chapter for extracting the frequencies of oscillation from an acoustic spectrum (see Section 2.3.3). The algorithm can be summarized as follows:

1. Identify the frequency of the highest peak in the amplitude spectrum (this can be easily done by searching for the maximum of the signal in the frequency domain).

2. Subtract the corresponding sinusoid from the original data according to Eq. (2.33) — for our purposes they can be either radial velocity time-series or light curves — and ensuring that the computed sinusoid is sampled in the same manner of the data themselves.

3. Consider the new residuals time-series $\{c_n\}$, with $i = 0, 1, \ldots, N-1$ and recompute the amplitude spectrum with e.g. the method described in Section 2.1.1.

4. Repeat the procedure from point (1) for the new highest peak in the last residual spectrum computed in point (3), until the S/N of the peak to be removed is below a given threshold (usually S/N = 4, [52], but one can adopt even lower values, depending on the quality of the data and on the manner the noise level in the amplitude spectra is treated, as done e.g. by [48] and [86]).

This algorithm can be used to filter a desired region in the power spectrum. This is done by applying the ISWF to all the peaks whose amplitudes are above the threshold set by the white noise level, and being in a range of frequency we want to "clean" from any power. This method is usually denoted as *pre-whitening*.

2.2 The Échelle Diagram for p Modes

The échelle diagram is a fundamental tool for inferring the most reliable value of the mean large separation $\Delta\nu$ in asteroseismic data analysis, and for deriving a proper identification of the modes. An échelle diagram is commonly constructed by plotting the frequencies derived in the preliminary CLEANing phase against the frequencies modulo $\Delta\nu$. In fact, according to the regular pattern of p modes in the frequency domain given by Eq. (1.10), when the value of the large separation is correct the modes in the échelle diagram align into straight vertical ridges (e.g. see [293]). Thus, the échelle diagram can be used to estimate both the best value of $\Delta\nu$ and the average spacing between the $\ell = 2$ and $\ell = 0$ ridges, i.e. $\delta\nu_{02}$. In addition, this simple but fundamental tool allows us to identify the frequencies derived from the amplitude spectrum of a star in most of the cases.

Figure 2.2 shows a theoretical échelle diagram, which is built through several models evaluated at different values of the frequency of maximum power, $\nu_{\max}$. The plot was done by [293], with $\nu_{\max}$ ranging from 0 to 4500 μHz. From left to right we find $\ell = 2, 0, 3$ and 1 as the most detectable modes, while $\ell = 4, 5$ modes are too faint to be observed with short length time-series. As the star evolves the frequency of maximum power decreases, because it scales from the acoustic cut-off frequency of the star [58]. When $\nu_{\max} < 1500\,\mu$Hz, the star becomes a

subgiant, hence it stops to burn Hydrogen in its core. During this phase, acoustic modes (from the outer convective zone) and gravity modes (arising from the interior of the star) are, in some cases, able to couple together, undergoing so called avoided crossings, as already mentioned in Chapter 1. These mixed modes have the main feature of deviating from the position given by the asymptotic relation for p modes, which is conversely expected to be followed properly by pure pressure modes. As we shall discuss in more detail in Chapter 4, this behavior is more evident for dipole modes, as it is clear from Figure 2.2 where red symbols are spread along the entire range in the case of subgiant stars. Mixed modes are however visible during the entire red giant phase, though having a spread in the échelle diagram smaller than that of a subgiant. Clearly, the identification of the modes is rather difficult for subgiant stars because they could easily be confused with modes of different angular degree. In this case, further investigations deriving from both a theoretical modeling of the star and mode profile features such as the FWHM of the peak in the acoustic spectrum[6] are necessary for properly identifying them.

A modified version of the échelle diagram uses the amplitude spectrum as a background for the CLEANed frequencies. This is done by folding the observed spectrum of the star into slices of length $\Delta\nu$ and subsequently collapsing them along the entire range of frequency we intend to plot (see also Appendix A). This version of the échelle diagram can be very helpful especially for confusing cases in which a low number of modes is observed because the additional information of the mode amplitude is given (see below).

2.3 The G9.5 subgiant β Aquilae

Once that the main tools for the asteroseismic analysis of time-series are so provided, we can proceed by describing the asteroseismic study of the G9.5 subgiant star β Aquilae [86].

2.3.1 Observation and Data Reduction

The subgiant β Aquilae has $V = 3^m.699 \pm 0^m.016$ (UBV photometric measurements from [250]), spectral type G9.5 IV [139], distance 13.70 ± 0.04 pc derived from the *Hipparcos* parallax $\pi = 73.00 \pm 0.20$ mas [309], $T_{\rm eff} = 5160 \pm 100$ K [209], $\log g = 3.79 \pm 0.06$ [307] and [Fe/H] $= -0.17 \pm 0.07$ [121]. The radius $R = 3.05 \pm 0.13\ R_\odot$ was measured by means of Long Baseline Interferometry [249] that is listed in the CHARM2 catalogue for high angular resolution measurements [265]. An excess of power in the power spectrum (hereafter PS) of β Aquilae data acquired with HARPS was already found by [191], who estimated a mean large separation $\Delta\nu = 30\ \mu$Hz from stellar parameters. The data for β Aquilae used in this work were acquired with observations carried out during six nights (2009 August 5-11) by means of the high-resolution cross-dispersed échelle spectrograph SARG [137, 83], Figure 2.3, mounted on the 3.58 m Italian telescope TNG at the La Palma observing site. SARG operates in both single-object and long-slit (up to 26") observing modes and covers a spectral wavelength range from 370 nm up to about 1000 nm, with a resolution ranging from $R = 29000$ up to $R = 164000$. Our spectra were obtained at $R = 164000$ in the wavelength range 462-792 nm. The calibration iodine cell works only in the blue part of the spectrum (500-620 nm), which was used for measuring Doppler shifts. During the observing period we collected spectra with a signal-to-noise ratio (S/N) varying from 150 to 300, a typical exposure of ~ 150 s and a sampling time of ~ 190 s. A total of 828 spectra were collected with the following distribution over the six nights: 91, 133, 110, 146, 163, 185. The spectra were then reduced and calibrated in wavelength

[6] The width of the peak in the acoustic spectrum, i.e. the lifetime of the mode (see Section 2.1.3) can be an important parameter to distinguish between pure pressure modes and gravity-dominated mixed modes because gravity modes have lifetimes much longer than acoustic modes.

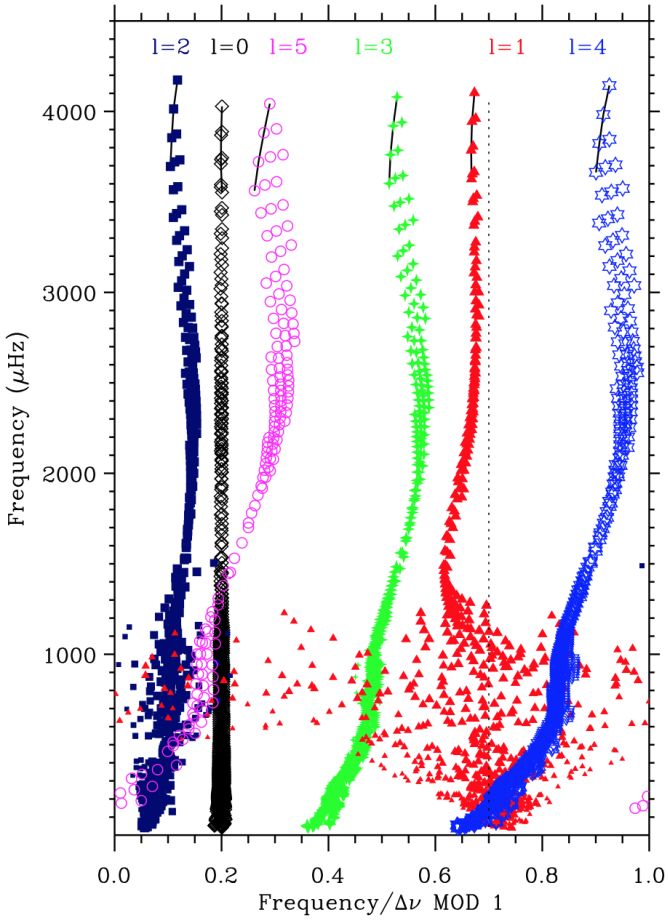

Figure 2.2: *Theoretical échelle diagram. Image copied from [293].*

with a Th-Ar lamp, using standard tasks of the IRAF package facilities (see e.g. [85]). No flat-fielding was applied to the spectra because of a degradation of the S/N level.

2.3.2 Radial Velocity Measurements with iSONG

The RV measurements were determined by means of the iSONG code, an IDL based software developed for the *SONG* project [142]. iSONG models the instrumental profile given by the point-spread function (PSF), stellar, and iodine cell spectra to measure Doppler shifts. The observed stellar spectrum, $I_{\mathrm{I}_2,\mathrm{obs}}$, was fitted with a reconstructed one by using a convolution between the oversampled stellar template, I_s, the very high-resolution iodine cell spectrum (derived from the transmission function of the iodine absorption cell, T_{I_2}), and the measured spectrograph instrumental profile. Essential to this process are the template spectra of β Aquilae taken with the iodine cell removed from the beam, and the iodine cell itself superimposed on a rapidly rotating B-type star, the same for all the measurements. The B-type star spectra allow us to construct the instrumental profile because they are essentially featureless, while the β Aquilae spectra acquired without the iodine cell create the original stellar template when deconvolved with the PSF (see [211, 61] for a detailed explanation of the method). The true stellar template, I_s, is found as

$$I_s = I_{\mathrm{obs}} \otimes^{-1} \mathrm{PSF}\,, \tag{2.34}$$

Figure 2.3: *SARG layout in the TNG Nasmyth room B. On the foreground, the SARG electronic rack. Behind it, the SARG, rigidly mounted on the telescope fork. Light reaches the spectrograph from a folding mirror. The rotator adaptor is kept fixed when using SARG. Field derotation is achieved by means of an optical derotator. Guiding is done at the slit, using a cooled CCD.*

where $I_{\rm obs}$ is the stellar spectrum taken without the iodine cell, while the PSF is obtained by using the B-star spectra taken with the absorption cell superimposed [306]. The deconvolution is done by means of a modified Jansson method [125]. Then, iSONG exploits the relation given by [211, 61] for attaining high precision Doppler shift measurements through iodine absorption cells, namely

$$I_{\rm I_2,obs}(\lambda; k, \Delta\lambda) = k\,[T_{I_2}(\lambda) I_s(\lambda + \Delta\lambda)] \otimes {\rm PSF} \tag{2.35}$$

where k is a normalization factor and $\Delta\lambda$ is the Doppler shift to be estimated, in order to construct a new observed stellar spectrum taken through the absorption cell, which is shifted by $\Delta\lambda$. The best $\Delta\lambda$, hence the RV we intend to measure, is found by minimizing the squared differences between the shifted observed spectrum and the original one. The velocities were corrected to the solar system barycenter [161, 162] and no other corrections, such as decorrelation or filtering by removing polynomial fits to the time-series, were applied.

iSONG provides also an estimate of the uncertainty in the velocity measurements, σ_i; these values were derived from the scatter of the velocities measured from many ($\simeq 650$), small ($\simeq 1.8$ Å) segments (chunks) of the échelle spectrum. To include the accuracy of the measurements in a weighted Fourier analysis of the data, we firstly verified that these σ_i reflected the noise properties of the RV measurements and their Fourier transform, following a slightly modified approach used by [62]. First we considered that the variance deduced from the noise level $\sigma_{\rm amp}$ in the amplitude spectrum has to satisfy the relationship

$$\sigma^2_{\rm amp} \sum_{i=1}^{N} \sigma_i^{-2} = \pi \tag{2.36}$$

according to [186] and [187]. This procedure yields new RV uncertainties scaled down by a factor of 1.51. Then we performed the following three steps. (i) The high-frequency noise

in the PS, well beyond the stellar signal ($> 1000\,\mu$Hz), reflects the properties of the noise in the RV data, and because we expected that the oscillation signal is the dominant cause of variations in the velocity time-series, we need to remove it to analyze the noise. This we made iteratively by finding the strongest peak in the PS of the velocity time-series and subtracting the corresponding sinusoid from the time-series (see also [112], Section 4.3, and Section 2.1.5 of this chapter). This procedure, namely the pre-whitening, was carried out for the strongest peaks in the oscillation spectrum in the frequency range 0-1.5 mHz, until the spectral leakage into high frequencies from the remaining power was negligible (see also [203, 48]). This left us with a time-series of residual velocities, r_i, that reflects the noise properties of the measurements. (ii) We then analyzed the ratio r_i/σ_i, expected to be Gaussian-distributed, so that the outliers correspond to the data points that exceed the given distribution. The cumulative histogram of the residuals is shown as a solid red curve in the upper panel of Figure 2.4, indicating the fraction of "good" data points. An excess of outliers is evident for $|r_i/\sigma_i| \gtrsim 1.5$. In this step we used the theoretical Gaussian function for the cumulative histogram, given by the expression

$$F(x_i) = \frac{N}{M}\left[1 - erf\left(x_i/x_0\right)\right], \tag{2.37}$$

where

$$erf(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{t^2}\,dt \tag{2.38}$$

is the error function, N is the total number of data points, $M = max\left[1 - erf\left(x_i/x_0\right)\right]$, and $x_i = |r_i/\sigma_i|$. The zero point fixed to $x_0 = 0.89$ allows us to adjust the fit for a noise-optimized distribution, which means that the chosen distribution minimizes the noise level in the weighted PS (see [29] for more details on different kinds of weight optimizations). The reason for this optimization relies on the possibility to improve the mode identification by increasing the S/N of the frequency peaks. (iii) The lower panel of Figure 2.4 shows the ratio f of the values of the observed points to the corresponding ones of the cumulative distribution function, i.e. the fraction of data points that could be considered as "good" observations, namely those that are close to unity. The quantities $w_i = 1/(\sigma_i^2 f)$ were adopted as statistical weights in the computation of the weighted PS according to the procedure described in Section 2.1.1 for the weighted LSS. In this way we get rid of the systematics that affect the gaussianity of the residuals and an improvement of the S/N ratio in the weighted PS is then expected.

The time-series of the whole dataset is presented in Figure 2.5 (lower panel) with the corresponding noise-scaled and outliers-corrected uncertainties σ_i (upper panel). Figure 2.6 shows the details of the oscillation observations during the fifth night, overlaid with a solid blue curve representing a smoothing of 3.30 min for enhancing the p-mode oscillations pattern. The final data point number of the full observation was reduced with respect to the number of observed spectra owing to a consistent improvement of the oscillation envelope, so that a total of 818 data points was taken into account for computing the time-series analysis.

We can now focus on the analysis concerning mode identification, amplitudes, lifetimes and mean global asteroseismic parameters of the subgiant star β Aquilae.

2.3.3 Time-Series Analysis

As already mentioned above, the amplitude spectrum of the velocity time-series was calculated as a weighted least-squares fit of sinusoids [112, 8, 28, 190], Section 2.1.1, with a weight assigned to each point according to its uncertainty estimate obtained from the RV measurement, explained in Section 2.3.2. The result is shown in Figure 2.7, where a clear excess of power around 420 μHz is visible, with the typical pattern for p-mode oscillations in a G9.5 subgiant

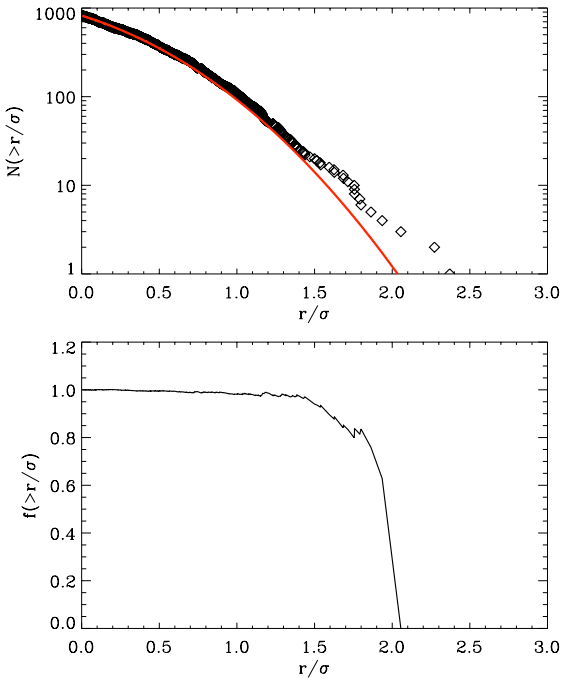

Figure 2.4: *Upper panel: cumulative histograms of $|r_i/\sigma_i|$ for SARG data. The diamonds show the observed data, and the solid curve shows the result expected for a Gaussian-distributed noise. Lower panel: ratio f of the observed to the expected histograms, indicating the fraction of "good" data points. An excess of outliers is evident for $|r_i/\sigma_i| \gtrsim 1.5$.*

star. This feature is apparent in the power spectra of individual nights, and its frequency agrees with theoretical expectations, as we will discuss in Section 2.4. To determine the S/N of the peaks in the PS, we measured the noise level $\sigma_{\rm amp}$ in the amplitude spectrum in the range 1200-1500 μHz, far from the excess of power. By means of the new weights introduced above, it was reduced from 14.4 cm s^{-1} to the final value of 12.8 cm s^{-1}, which corresponds to a noise level in the PS $\sigma_{\rm PS} = 0.02\,{\rm m^2\,s^{-2}}$, where $\sigma_{\rm PS} = (4/\pi)\sigma^2_{\rm amp}$ as explained by [187]. Since this is based on 818 measurements, we can deduce the velocity precision on the corresponding time-scales using the relation $\sigma_{\rm RMS} = \sigma_{\rm amp}\sqrt{N/\pi}$, as derived by [187], which gives a scatter per measurement of 2.1 m s^{-1}.

However, particular care has to be taken with the noise evaluation within the region of solar-like oscillations. Indeed, by evaluating the noise in the amplitude spectrum in the intervals 100-300 μHz and 600-800 μHz, just below and above the excess of power, we obtained the two noise levels $\sigma^{100-300}_{\rm amp} = 26.2\,{\rm cm\,s^{-1}}$ and $\sigma^{600-800}_{\rm amp} = 15.6\,{\rm cm\,s^{-1}}$, which appear to be quite different. We then adopted a noise decaying accordingly to a linear trend law, ranging between the two values within the region 200-700 μHz.

2.3.4 Search for a comb-like pattern

We remind that, as already discussed in Chapter 1, the frequencies for low-degree, high radial order p-mode oscillations in Sun-like stars are reasonably well approximated by the asymptotic relation [299], which can be written in a more useful form for the analysis presented here, that is

$$\nu_{n,\ell} = \Delta\nu\left(n + \frac{\ell}{2} + \epsilon\right) - \ell(\ell+1)\delta\nu_{02}/6\,, \tag{2.39}$$

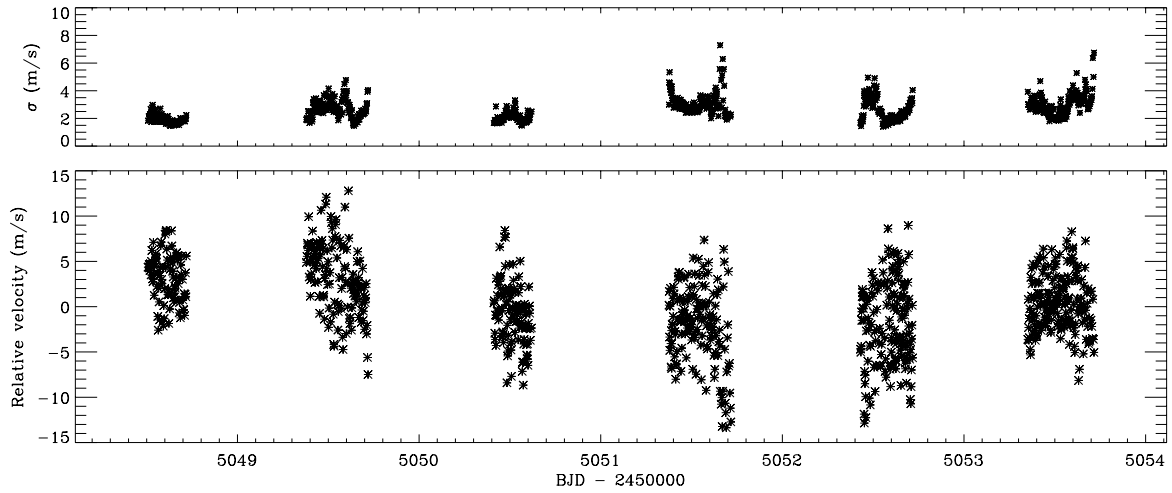

Figure 2.5: *Radial velocity measurements of β Aquilae for the entire time of observation, obtained with the iSONG code from the SARG spectra (lower panel) and their corresponding noise-scaled and outliers-corrected uncertainties σ_i (upper panel).*

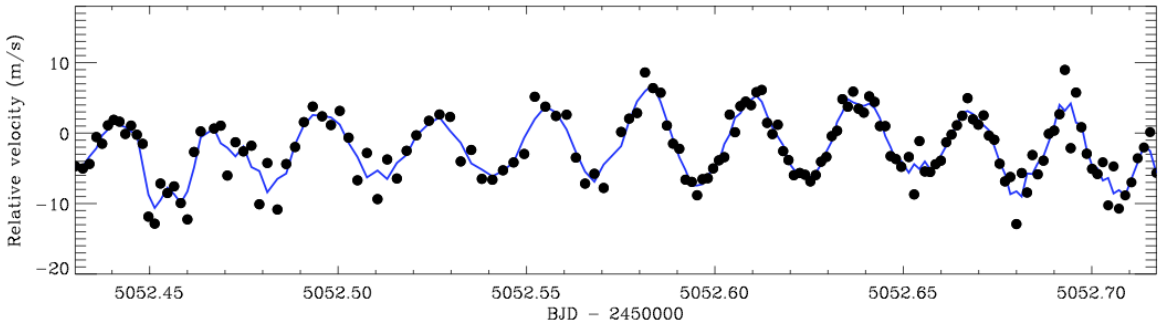

Figure 2.6: *Detail of the fifth night of observation of radial velocity measurements obtained with the iSONG code from the SARG spectra (filled black circles). The solid blue line represents a 3.30 min wide smoothing to enhance the p-mode oscillations pattern.*

where n and ℓ are integers that define the radial order and angular degree of the mode, respectively, $\Delta\nu = \langle \nu_{n,\ell} - \nu_{n-1,\ell} \rangle$ is termed mean large frequency separation, $\delta\nu_{02} = \langle \nu_{n,0} - \nu_{n-1,2} \rangle$ is the small frequency separation for the $\ell = 2, 0$ modes, a quantity sensitive to the sound speed gradient near the core, and ϵ is a quantity on the order of unity sensitive to the stratification of the surface layers (see also Chapter 1). On attempting to find the peaks in our power spectrum that match the asymptotic relation, we were severely hampered by the single-site window function, whose power is visible in normalized units in the inset of Figure 2.7, giving an effective observation time of ~ 1.80 days (i.e. the reciprocal of the integral in power of the window function). As is well known, daily gaps in a time-series produce aliases in the power spectrum at spacings $\pm 11.57\,\mu$Hz and multiples, which are difficult to disentangle from the genuine peaks. Various methods to search for regular series of peaks have been discussed in the literature, such as autocorrelation, comb response and histograms of frequencies. To find a starting value for the $\Delta\nu$ investigation we used the comb-response function method, where a comb-response function CR $(\Delta\nu)$ is calculated for all sensible values of $\Delta\nu$ (see [189] for details), representing a generalization of the PS of a PS and consequently allowing us to search for any regularity in the spectral pattern. In particular we used the generalized comb-response function discussed

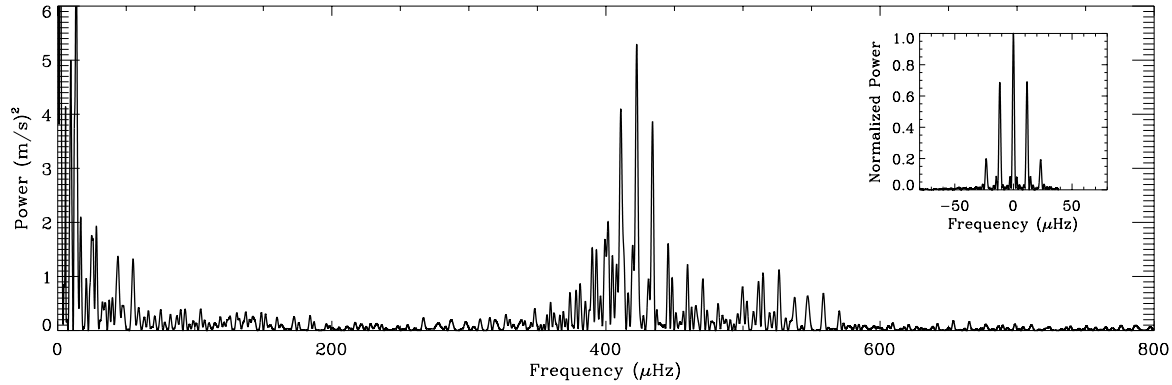

Figure 2.7: *Power spectrum of the weighted radial velocity measurements of β Aquilae extracted with the iSONG code from the SARG spectra. An excess of power is clearly visible, with a maximum centered at 422 μHz. The inset shows the normalized power spectrum of the window function for a sine-wave signal of amplitude 1 m s^{-1}, sampled in the same way as the observations.*

in [48]

$$\mathrm{CR}\,(\Delta\nu) = \prod_{n=1}^{N} \left[PS\left(\nu_0 \pm \frac{2n-1}{2}\Delta\nu\right) PS\,(\nu_0 \pm n\Delta\nu) \right]^{\frac{1}{2^{n-1}}}, \tag{2.40}$$

so that a peak in the CR at a particular value of $\Delta\nu$ indicates the presence of a regular series of peaks in the PS, centered at $\nu_{\max}$ and having a spacing of $\Delta\nu/2$. It differs from a correlation function because the product of individuals terms is considered rather than the sum. For actual calculations we used $N = 2$ but we checked that our result was stable for $N > 2$ as well.

To reduce the uncertainties caused by noise, only peaks above 300 μHz and with amplitude $> 0.5\,\mathrm{m\,s^{-1}}$ in the amplitude spectrum, corresponding to a S/N $\gtrsim 3.5$, were used to compute the CR. We determined the local maxima of the response function CR $(\Delta\nu)$ in the range $8 \leq \Delta\nu \leq 50\,\mu$Hz as first step. In this case the result was seriously affected by two very strong peaks at 11.57 μHz and 23.14 μHz, corresponding to once and twice the value of the daily gap, respectively, which consistently reduced the strength of the peak corresponding to $\Delta\nu$. We then decided to restrict the search range to $26 \leq \Delta\nu \leq 50\,\mu$Hz as second step to completely exclude the daily alias peaks. The resulting cumulative comb-response function for this range, obtained by summing the contributions of all the response functions, had the most prominent peak centered at $28.90 \pm 0.45\,\mu$Hz as shown in Figure 2.8, where the uncertainty was computed by considering the FWHM of the Gaussian used to fit the peak, and represented by the blue dot-dashed curve. The peak corresponding to three times the daily spacing is also visible in the right side of the plot, centered at 34.71 μHz.

2.3.5 Oscillation Frequencies

The comb-response function provided a guess of the large separation, which was then used as a starting point to investigate the most reliable value that could represent the observed data. The investigation involved a parallel checking and trade-off between two different methods, i.e. the folded PS and the échelle diagram (see Section 2.2). Firstly we computed the folded PS, namely the PS collapsed at the large separation value, for different values of $\Delta\nu$ subsequent to the guess number. The final result for the folding was computed for $\Delta\nu = 29.56\,\mu$Hz and is shown in Figure 2.9, where the peaks corresponding to $\ell = 0$ and $\ell = 2$ ridges are marked by a dashed and a dotted line, respectively. The daily side-lobes for $\pm 11.57\,\mu$Hz are also clearly visible and are marked with the same line-styles. The overlaid ridges represent the result of

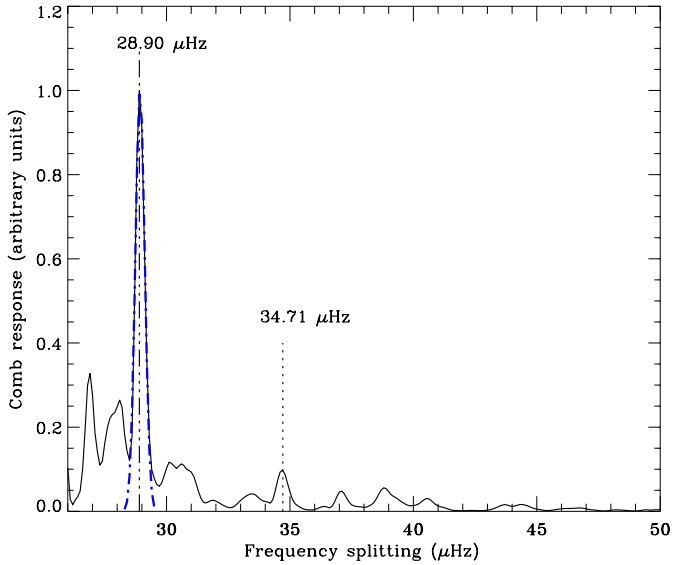

Figure 2.8: *Cumulative comb-response obtained as the sum of the individual comb-responses for each central frequency ν_0 with amplitude $> 0.5\,m\,s^{-1}$ ($S/N \gtrsim 3.5$) in the amplitude spectrum. The maximum peak is centered at $\Delta\nu = 28.90 \pm 0.45\,\mu Hz$ where the uncertainty is evaluated as the FWHM of the Gaussian used to fit the peak (blue dot-dashed curve). The second marked peak on the right, centered at $34.71\,\mu Hz$, represents three times the daily spacing.*

a least-squares fit to the asymptotic relation given by Eq. (2.39). It is noticeable how the $\ell = 1$ ridge, marked by a dot-dashed line, does not correspond to any peak in the folded PS. This could be caused by avoided crossings, although it could also be explained by a wrong identification of the modes, as we will discuss in more detail in Section 2.4. We note that the $\ell = 2$ ridge appears to be slightly shifted with respect to the position of the maximum of the relative peak, a result that can be explained by the limit of our frequency resolution.

The second method adopted for the investigation of the large separation is represented by the well known échelle diagram, an essential tool for frequency identification in asteroseismic data, as discussed in Section 2.2. The échelle diagram was computed for a list of 9 high S/N ($\gtrsim 3.5$) frequencies directly obtained from the PS by means of the CLEAN algorithm (see Section 2.1.5), although a further consideration regarding this aspect is required to explain the way the final list was attained. In fact, two different ways of CLEANing the PS were adopted for this work. In the first case the algorithm was applied to the weighted PS for a sufficiently large number of frequencies ($\simeq 50$ to find all the peaks with S/N > 3) in the range 0.01-700 μHz, then the resulting values were restricted to the interval 200-700 μHz, obtaining a list of 20 frequencies. For the second case the PS region below 200 μHz was at first completely pre-whitened. Then 20 frequencies were identified on the new resulting weighted PS in the range 200-700 μHz. Comparing both lists of frequencies, the most unstable ones, i.e. those that were not present in both lists, were rejected and only the first nine frequencies were considered. These frequencies, showing an amplitude $> 0.5\,\mathrm{m\,s^{-1}}$ (or S/N $\gtrsim 3.5$), were almost the same in both cases. However, the first frequency list was selected because the chance of the identification of one more $\ell = 2$ mode was possible for this case. In Figure 2.10 we show the final result for the échelle diagram superimposed on a colored scale background representing the amplitude spectrum, where the plot was computed in the same way as in [216]. The filled symbols (white and orange) represent the 10 identified modes for $\ell = 0$ (circles), $\ell = 1$ (triangles) and $\ell = 2$ (squares); the orange symbols are the corrected frequencies, i.e. those shifted for the daily

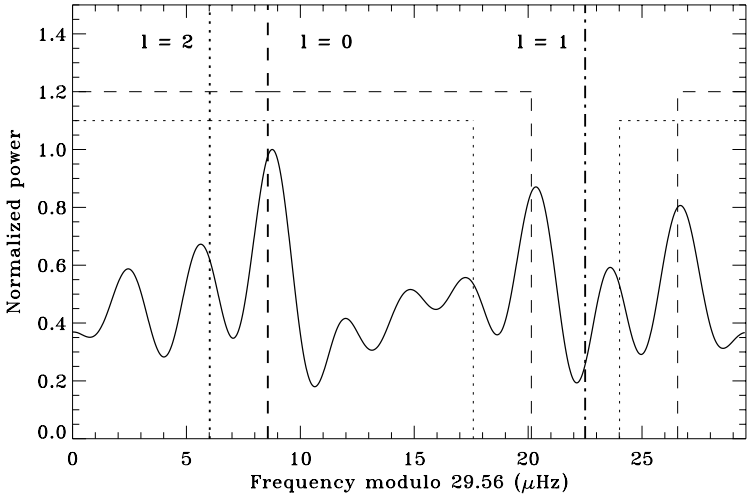

Figure 2.9: *Folded PS in normalized units in the case of $\Delta\nu = 29.56\,\mu$Hz. The overlaid ridges represent the result for $\ell = 0, 1, 2$ mode degrees as derived by a least-squares fit to the asymptotic relation, which are marked as dotted, dashed and dot-dashed lines, respectively. The daily side-lobes for $\pm 11.57\,\mu$Hz in the case of $\ell = 0, 2$ are also clearly visible and marked with the same line-style for each mode degree.*

gap of $\pm 11.57\,\mu$Hz from their original values (the corresponding open symbols) and reported in Table 2.4. The aliasing considerably affects the amplitude spectrum with the presence of several fictitious peaks, which appear as strong spots in the diagram without frequency symbols overlaid. In particular, the two daily side-lobes arising from the strongest mode ($\ell = 0$) are clearly visible as two red spots on the right-hand part of the diagram. The uncertainties on frequencies are listed in Table 2.4 and were evaluated by using the analytical relation provided by [237]

$$\sigma(\nu) = \sqrt{\frac{6}{N}}\frac{1}{\pi T}\frac{\langle\sigma_{\rm v}\rangle}{A} \simeq \frac{0.16}{A_{\rm m/s}}\ \mu{\rm Hz}\,, \tag{2.41}$$

where $N = 818$ is the total number of data points, $T = 5.21$ is the total duration of the run in days, $\langle\sigma_{\rm v}\rangle = 2.72\,{\rm m\,s^{-1}}$ is the average uncertainty on each RV data point and A is the amplitude per mode as derived in Section 2.3.6. This relation holds exactly for coherent oscillations, hence we remark that the estimated uncertainties represent only a lower limit to the real uncertainty value because these oscillations are not fully coherent. An upper limit to these uncertainties can be fixed to the formal frequency resolution, given as the reciprocal of the total duration of the run, which is $2.2\,\mu$Hz for this dataset. The complete identification of the p modes is reported in Table 2.4 together with their S/N, where only three frequencies out of nine were shifted by the daily alias. The $\ell = 1$ modes reported without any radial order number are potential candidates for mixed modes.

By means of a linear weighted least-squares fit to the asymptotic relation of the $\ell = 0$ frequencies, the final value of $\Delta\nu = 29.56 \pm 0.10\,\mu$Hz was obtained, together with the constant $\epsilon = 1.29 \pm 0.04$. The most likely value for the small separation was derived by using the definition from the asymptotic relation with the frequency pairs $(\nu_{12,0}, \nu_{11,2})$, $(\nu_{13,0}, \nu_{12,2})$ and $(\nu_{16,0}, \nu_{15,2})$, where $\nu_{15,2} = 508.47\,\mu$Hz was computed directly from $\nu_{16,2} = 538.03\,\mu$Hz by adopting our value of the large separation. This led to $\delta\nu_{02} = 2.55 \pm 0.71\,\mu$Hz, as reported in Table 2.2, but because it is comparable to the frequency resolution, its uncertainty is relatively high (Figure 2.10). As a consequence, the reliability of this result requires additional investigations and a longer dataset.

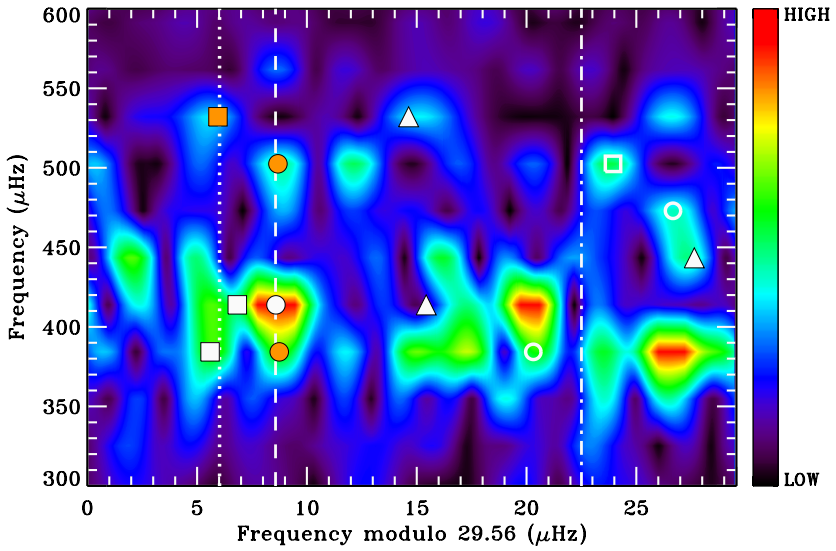

Figure 2.10: *Échelle diagram of β Aquilae overlaid on the amplitude spectrum with a colored background scale. The filled symbols (white and orange) represent the identified modes for $\ell = 0$ (circles), $\ell = 1$ (triangles) and $\ell = 2$ (squares). The orange symbols are the frequencies shifted for the daily gap of $\pm 11.57\,\mu$Hz while the open symbols correspond to the original unshifted values. The ridges derived from the fit to the asymptotic relation (2.39) are also marked.*

2.3.6 Mode Amplitudes

The evaluation of mode amplitudes is interesting for the discussion on the L/M scaling by extrapolating from the Sun, as we will discuss in Section 2.4. As is known from the theory of solar-like oscillations, the amplitudes of individual modes are affected by the stochastic nature of the excitation and damping (e.g. see [187]). To measure the oscillation amplitude per mode in a way that is independent of these effects, we followed the approach explained in [190, 191]. This involves the following steps: (i) heavy smoothing of the weighted PS by convolving it with a Gaussian whose FWHM is fixed to $4\Delta\nu$ (which is the value considered as a standard that allows comparisons since the amount of smoothing affects the exact height of the smoothed amplitude spectrum), which is large enough to produce a single hump of power that is insensitive to the discrete nature of the oscillation spectrum; (ii) conversion of the smoothed PS to PSD by multiplying by the effective length of the observing run (that is the reciprocal of the spectral resolution, Section 2.1.4, namely $6.40\,\mu$Hz for this dataset); (iii) subtraction of the background noise, which we computed as a linear trend in the interval 200-700 μHz, ranging from $26.2\,\mathrm{cm\,s^{-1}}$ to $15.6\,\mathrm{cm\,s^{-1}}$; (iv) multiplication of the result by $\Delta\nu/c$ where $c = 4.09$ (which represents the effective number of modes that fall in each segment of length $\Delta\nu$ as evaluated in the case of radial velocities) and taking the square root to convert to amplitude per oscillations mode. The result is shown in Figure 2.11 for the range 200-700 μHz, where $A_{\max} = 76 \pm 13\,\mathrm{cm\,s^{-1}}$ centered at $\nu_{\max} = 416\,\mu$Hz, which is assumed to be the frequency of maximum power and agrees with the result of [191]. The uncertainty on the amplitude is evaluated by means of the analytical estimation relation

$$\sigma(A) = \sqrt{\frac{2}{N}}\,\langle\sigma_{\mathrm{v}}\rangle \tag{2.42}$$

[237], where, as for the frequencies case, $N = 818$ is the total number of data points and $\langle\sigma_{\mathrm{v}}\rangle$ is once more the average uncertainty on each data point. Again, the uncertainty in amplitude

Table 2.1: *Mode identification for β Aquilae, in the frequency range 300-600 μHz. The modes listed show an amplitude $> 0.5\,m\,s^{-1}$ (or $S/N \gtrsim 3.5$). The corrected frequencies reported in the fourth column include the frequencies shifted for the daily gap of $\pm 11.57\,\mu$Hz. The fifth column represents the uncertainties as derived by means of the analytical relation (2.41). The $\ell = 1$ frequencies reported without any radial order are possible avoided crossings.*

ℓ	n	S/N	Corrected frequency[a] (μHz)	Uncertainty[b] (μHz)
0	12	6.4	393.02 (+11.57)	0.22
0	13	11.3	422.43	0.22
0	16	3.7	511.20 (-11.57)	0.34
1	-	4.0	429.27	0.16
1	-	4.7	471.04	0.20
1	-	3.5	546.71	0.38
2	11	3.5	389.87	0.22
2	12	3.5	420.67	0.21
2	16	5.1	538.03 (-11.57)	0.42

[a] The raw frequencies can be evaluated by adding the daily frequency reported in parentheses.

[b] The actual uncertainties can be several times larger up to the limit of the formal resolution of 2.2 μHz, because the modes are not coherent.

Table 2.2: *Asymptotic parameters for β Aquilae as derived by a linear weighted least-squares fit to the asymptotic relation given by Eq. (2.39).*

$\Delta\nu$ (μHz)	$\delta\nu_{02}$ (μHz)	ϵ
29.56 ± 0.10	2.55 ± 0.71^a	1.29 ± 0.04

[a] The most likely value for the small separation was derived by using the definition from the asymptotic relation with the couples of frequencies $(\nu_{12,0}, \nu_{11,2})$, $(\nu_{13,0}, \nu_{12,2})$ and $(\nu_{16,0}, \nu_{16,2} - \Delta\nu)$.

represents a lower limit to the real value. The amplitude distribution is evaluated for the radial modes only, but the calculation of the amplitude in the case of $\ell = 1$ and $\ell = 2$ modes is straightforward, namely it can be obtained by multiplying the result for a factor of 1.35 and 1.02 respectively, representing the relative strength given by the spatial response function (see [191] for more details). The result derived in this work is not far from the value obtained by [191].

2.3.7 Mode Lifetimes

For stars showing a large number of identified modes it is in principle possible to determine the mode lifetime by using the correlation between the mode lifetime and the scatter of the observed frequencies with respect to their asymptotic values (see [28, 190] for more details). Nevertheless this method is not applicable in our case since the number of identified modes is small and, being β Aquilae a quite evolved subgiant star, many avoided crossings crowding the échelle diagram are expected to be observed, a feature that would hamper considerably the regularity of the peaks and the modes identification. We then adopted the method presented in [203], where the modes amplitudes are supposed to increase with increasing lifetime. However, since there is not a simple relationship between amplitude and lifetime, we used simulated time-series in order to establish a calibrated relationship between the peak amplitude and the mode lifetime. With respect to the previous method, this one has the advantage of being independent of the mode identification but the disadvantage of being less accurate. The procedure adopted can be

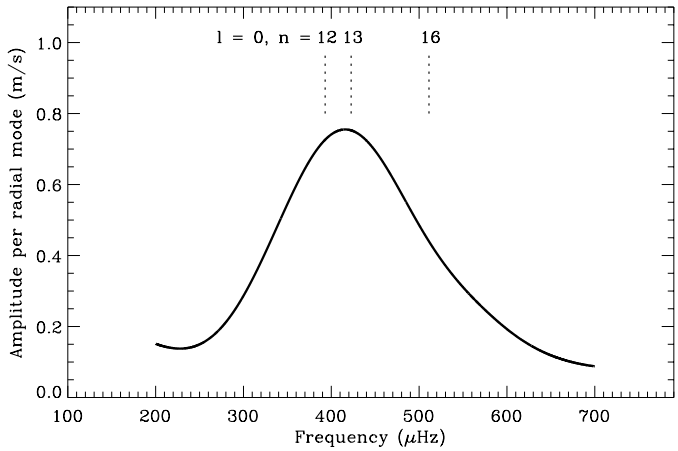

Figure 2.11: *Smoothed amplitude spectrum of β Aquilae showing the amplitude per radial mode computed in the range 200-700 μHz. The maximum amplitude $A_{\max} = 76 \pm 13\, cm\, s^{-1}$ occurs at $\nu_{\max} = 416\, \mu Hz$. The positions of the identified $\ell = 0$ frequencies as derived from the asymptotic relationship given by Eq. (2.39) are also marked.*

summarized as follows: (i) computation of one thousand simulated time-series for each value of mode lifetime (0.2, 0.5, 1, 2, 4, 8, 16, 32 and 64 days were the lifetimes chosen for this work, in order to cover a wide range of values), with a total of 9000 simulations computed; (ii) analysis of all the simulated PS in the way explained in Section 2.3.5, namely by CLEANing each PS with 50 peaks in the frequency range 0.01-700 μHz, in order to identify all the frequencies with S/N > 3; (iii) restriction of the number of frequencies to the interval 200-700 μHz; (iv) evaluation of the ratio $\langle A \rangle_{1-5} / \langle A \rangle_{1-20}$, i.e. the ratio of the average amplitude of the first 5 highest peaks with respect to the average amplitude of the first 20 peaks for each given simulated PS; (v) comparison of the ratio of the observed data with the simulated ones in order to derive the corresponding lifetime. The result is shown in Figure 2.12, where no clear correlation is visible. This can be explained by considering the fact that this star has a quite small value of the mean large separation $\Delta\nu$. This feature, combined with the complex structure of the spectral window caused by the daily gaps, and to the short observing time, makes each slice of PS having length $\approx 30\, \mu$Hz completely filled by frequency peaks and their corresponding sidelobes. The overlap of the line profiles occurring also for long lifetimes (i.e. 64 days) does not allow us to derive any information on their FWHM and consequently on the mode lifetimes.

2.4 Comparison with Theoretical Expectations

A complete discussion on the evolutionary state of this star goes beyond the scope of this work. Nevertheless, the identified modes provided a reliable estimate for the mean large separation, as described in Section 2.3.5, which can be used to derive the scaled mass for this star according to the fact that $\Delta\nu$ scales approximately with the square root of the mean density of the star [90]. From the scaling relation extrapolating from the solar case (e.g. see [29])

$$\frac{\Delta\nu}{\Delta\nu_\odot} = \left(\frac{M}{M_\odot}\right)^{0.5} \left(\frac{R}{R_\odot}\right)^{-1.5}, \tag{2.43}$$

where the radius is provided by [249] and $\Delta\nu_\odot = 134.9\, \mu$Hz, we obtained a mass of $M = 1.36 \pm 0.17\, M_\odot$, which agrees very well with the value found by [121]. By considering a luminosity of $L = 5.63 \pm 0.16\, L_\odot$, as derived by means of visual magnitude [250], bolometric correction

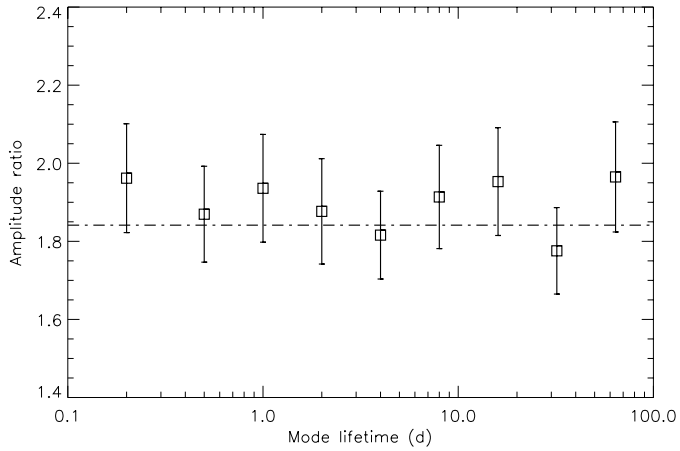

Figure 2.12: *Amplitude ratio $\langle A\rangle_{1-5}/\langle A\rangle_{1-20}$, i.e. the ratio of the average amplitude of the first 5 highest peaks respect to the average amplitude of the first 20 peaks, reported as a function of the mode lifetime in days. The error bars are computed as the standard deviation from one thousands points for each lifetime. The dot dashed line represents the same ratio level as measured on the observed dataset.*

[111], Sun bolometric magnitude [91] and HIPPARCOS parallax [309], and the scaled value for the mass, we were able to compute the value for the amplitude by using the relation provided by [187]:

$$A_{\rm osc} = \left(\frac{L}{L_\odot}\right)\left(\frac{M}{M_\odot}\right)^{-1}(23.4\pm1.4)\ {\rm cm\,s^{-1}}, \tag{2.44}$$

which gives $A_{\rm osc} = 97\pm6\,{\rm cm\,s^{-1}}$, which fairly agrees with the value obtained in Section 2.3.6. Moreover, the expected frequency of maximum power can be evaluated from scaling the acoustic cutoff frequency from the solar case. We computed this frequency by considering the relationship

$$\nu_{\rm max} = \left(\frac{M}{M_\odot}\right)\left(\frac{R}{R_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{-1.5} 3.05\cdot10^3\,\mu{\rm Hz} \tag{2.45}$$

where $T_{\rm eff,\odot} = 5777\,{\rm K}$ (see [187] and Chapter 6 for more details), whose result gives $\nu_{\rm max} = 472\pm72\,\mu$Hz, consistent with the value obtained on the smoothed amplitude spectrum.

Table 2.3: *Global list of stellar parameters for β Aquilae.*

Stellar parameter	Value	Source
M	$1.36\pm0.17\,M_\odot$	This work[a]
R	$3.05\pm0.13\,R_\odot$	[249]
$\langle\rho\rangle$	$0.0676\pm0.0004\,{\rm g\,cm^{-3}}$	This work
Z	0.015 ± 0.002	SEEK
X	0.70 ± 0.02	SEEK
Age	$2.43^{+3.56}_{-0.32}$ Gyr	SEEK
L	$5.63\pm0.16\,L_\odot$	This work
log g	$3.61^{+0.01}_{-0.02}$	SEEK
$T_{\rm eff}$	$5160\pm100\,{\rm K}$	[209]
$\Delta\nu$	$29.56\pm0.10\,\mu$Hz	This work
$\delta\nu_{02}$	$2.55\pm0.71\,\mu$Hz	This work
$A_{\rm max}$	$76\pm13\,{\rm cm\,s^{-1}}$	This work

[a] The corresponding parameter was derived directly in this work.

Concerning the values for the mean large separation and the frequency of maximum power

it is outstanding that they fit the power law relation of [286] quite well, which represents a considerable validation to the reliability of our results. Moreover, an independent measure for the large separation by using the Eq. (2.43), with the radius from [249] and the mass from [121] is also compatible with the value presented in Section 2.3.5.

Lastly, an important note regards the dipole frequencies reported in this work, which are expected to deviate strongly from their asymptotic values, especially for an evolved subgiant star like β Aquilae. Indeed, the presumed $\ell = 1$ modes identified here are possibly bumped because of avoided crossings. Although the lack of a clear $\ell = 1$ ridge in Figure 2.9 could also be explained by a wrong identification of the modes, we are confident that the large separation is correct. The reason of our belief is that different tools for its investigation, such as the comb-response function, the folded PS, the échelle diagram and the fit to the asymptotic relation, show consistent results. In addition, its compatibility with the independent estimation and the agreement with the $\Delta\nu$–$\nu_{\rm max}$ power law relation mentioned in the above paragraph ensure a robust derivation of the result presented. Nonetheless, we are talking about a very difficult star, such as ν Ind [64], because in the region of the HR diagram to which the star belongs, avoided crossings considerably hamper the p-mode identification (see [24] for a summary on ground-based observations across the HR diagram). Therefore, a theoretical confirmation is required before adopting the $\ell = 1$ frequencies reported in Table 2.4 as real frequencies of oscillations for mixed modes, and more observations by a multi-site ground-based project such as *SONG* are required to firmly solve the mode identification.

2.4.1 Stellar Parameters

The SEEK package [260] is developed for the analysis of asteroseismic data from the *Kepler* mission and is able to estimate stellar parameters in a form that is statistically well-defined. It is based on a large grid of stellar models computed with the Aarhus Stellar Evolution Code (ASTEC), which allow us to derive additional stellar parameters giving as input astrophysical quantities such as $T_{\rm eff}$ [209], $\log g$ [307] and [Fe/H] [121], and the asteroseismic values derived in this work, i.e. the large and small separation and the frequency of maximum power. The output list of parameters for β Aquilae is shown in Table 2.3, where the 1-σ error bars on the SEEK values are computed using the Bayesian evaluation of the posterior distributions. The mass, radius and luminosity computed by SEEK also agree with the values presented in this work.

2.5 The G5 subgiant μ Herculis

The G5 subgiant μ Herculis (HR 6623, $V = 3.417 \pm 0.014$), has a mass $M = 1.1\,M_{\odot}$, effective temperature $T_{\rm eff} = 5596 \pm 80\,$K, and $\log g = 3.93 \pm 0.10$ [120]. It has been studied already by [48], who found the star to have a clear excess of power centered at $\nu_{\rm max} = 1110\,\mu$Hz, a large frequency separation $\Delta\nu = 56.50 \pm 0.07\,\mu$Hz, a small frequency spacing $\delta\nu_{02} = 5.03 \pm 0.94\,\mu$Hz and $\epsilon = 1.44 \pm 0.03$.

In this section we provide a revised analysis of this star, made along the same lines of those used by [48], by exploiting a new set of RV measurements derived by means of the iSONG code. The data were acquired with observations carried out over six nights (2006 August 13-19) by means of the high resolution cross dispersed échelle spectrograph SARG [137, 83] mounted on the 3.58 m Italian telescope TNG at La Palma observing site. SARG operates in both single-object and long-slit (up to 26") observing modes and covers a spectral wavelength range from 370 nm up to about 1000 nm, with a resolution ranging from $R = 29000$ up to $R = 164000$. Our spectra were obtained at $R = 144000$ in the wavelength range 462-792 nm.

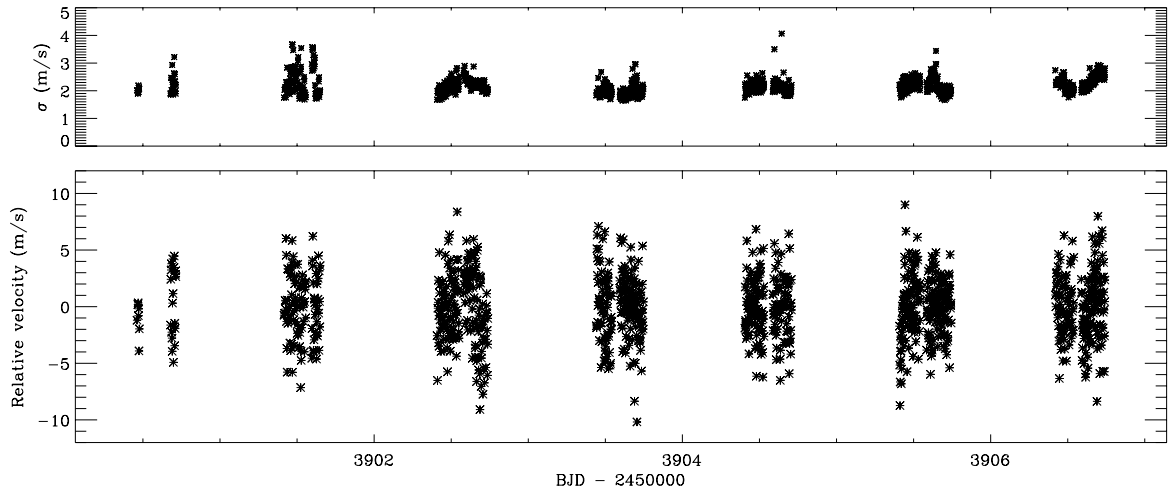

Figure 2.13: *Radial velocity measurements of μ Herculis for the entire time of observation, obtained with the* iSONG *code from the* SARG *spectra reduced with the* REDUCE *package (lower panel), and their corresponding noise-scaled and outliers-corrected uncertainties σ_i (upper panel).*

The calibration iodine cell works only in the blue part of the spectrum (500-620 nm) that has been used for measuring Doppler shifts. During the observing period we collected spectra with a signal-to-noise ratio (S/N) varying from 200 to 400, a typical exposure of 60 s with a dead-time 55 s between exposures due to the readout time. A total of 1106 spectra were collected with the following distribution over the seven nights: 27, 106, 184, 179, 227, 198. The spectra were then reduced and calibrated in wavelength with a Th-Ar lamp, using the recently released REDUCE package [258]. No flat-fielding was applied to all the spectra due to a degradation of the S/N level.

The RV measurements were determined by means of the iSONG code, Section 2.3.2. An offset caused by the inversion of the telescope within every night of observation was removed, in order to overcome the systematics introduced in the time-series. An example of such correction is provided for the 5^{th} night of observation, as shown in Figure 2.14 where the blue dots represent the original time-series, without any offset correction and the black ones are our final result. The final time-series is plotted in Figure 2.13, with a similar description to that adopted for Figure 2.5.

2.5.1 Comparing iSONG and AUSTRAL

Following the same procedure explained in Section 2.3.2 we computed the outliers distribution for the new time-series. The result is shown in Figure 2.15 (left) in comparison to the outliers distribution for the old dataset used by [48] (right). As it appears quite clear from both the upper and the lower panels, iSONG provides RV measurements that are suffering less from systematic effects than those given by AUSTRAL. In fact, the excess of outliers becomes evident for $|r_i/\sigma_i| \gtrsim 2$ in the case of iSONG, while it is already pronounced for $|r_i/\sigma_i| \gtrsim 1.5$ in the case of AUSTRAL.

The scaling factor for the uncertainties in the iSONG time-series is 1.53, while for the AUSTRAL one is 0.57. This shows that, according to Eq. (2.36), on one hand the uncertainties derived with the version of iSONG used for the analysis are overestimated, while on the other hand those of AUSTRAL used by [48] are underestimated. An improvement concerning the derivation of the uncertainties in radial velocity for the iSONG code is still undergoing.

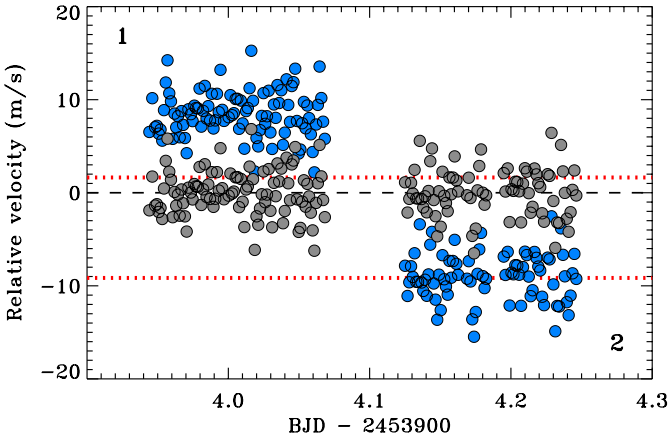

Figure 2.14: *Example of the 5$^{\text{th}}$ night of observation for μ Herculis, before (blue dots) and after (black dots) offset removing. Red dotted lines represent the offset level (one for each segment of the night, marked with numbers 1 and 2) subtracted with respect to the zero level corresponding to the black dashed line.*

Table 2.4: *Offsets of the original time-series μ Herculis for each night, according to Figure 2.14. Values are reported for each night of observation.*

Day	Offset 1 ($\mathrm{m\,s^{-1}}$)	Offset 2 ($\mathrm{m\,s^{-1}}$)
1	8.34	9.80
2	8.98	2.82
3	6.36	3.23
4	15.57	-11.24
5	-8.97	-8.97
6	1.56	-7.78
7	1.64	-9.15

2.5.2 Time-Series Analysis

The amplitude spectrum of the velocity time-series was calculated by means of the LSS method, as s explained in Section 2.3.3. The result is shown in Figure 2.16, where a clear excess of power around 1200 μHz is visible, with the typical pattern for p-modes oscillations in a G5 subgiant star. In order to determine the S/N of the peaks in the PS, we measured the noise level σ_{amp} in the amplitude spectrum in the range 3000-5000 μHz, far from the excess of power, as done by [48]. By means of the new statistical weights introduced above, it was reduced from $11.3\,\mathrm{cm\,s^{-1}}$ to the final value of $11.1\,\mathrm{cm\,s^{-1}}$, which corresponds to a noise level in the PS of $0.02\,\mathrm{m^2\,s^{-2}}$. Since this is based on 1106 measurements, we can deduce the velocity precision on the corresponding time-scales using the relation $\sigma_{\mathrm{RMS}} = \sigma_{\mathrm{amp}}\sqrt{N/\pi}$, as derived by [187], which gives a scatter per measurement of $2.1\,\mathrm{m\,s^{-1}}$. For this dataset it is evident from the values of σ_{amp} reported above that the noise reduction is not producing a significant result, which relies mainly on the fact that the excess of outliers for the time-series of μ Herculis is lower than that of β Aquilae.

2.5.3 Oscillation Frequencies

In order to derive the frequencies of the modes and the subsequent mode identification, we performed an identical approach to that exposed in Section 2.3.5 but by considering a frequency

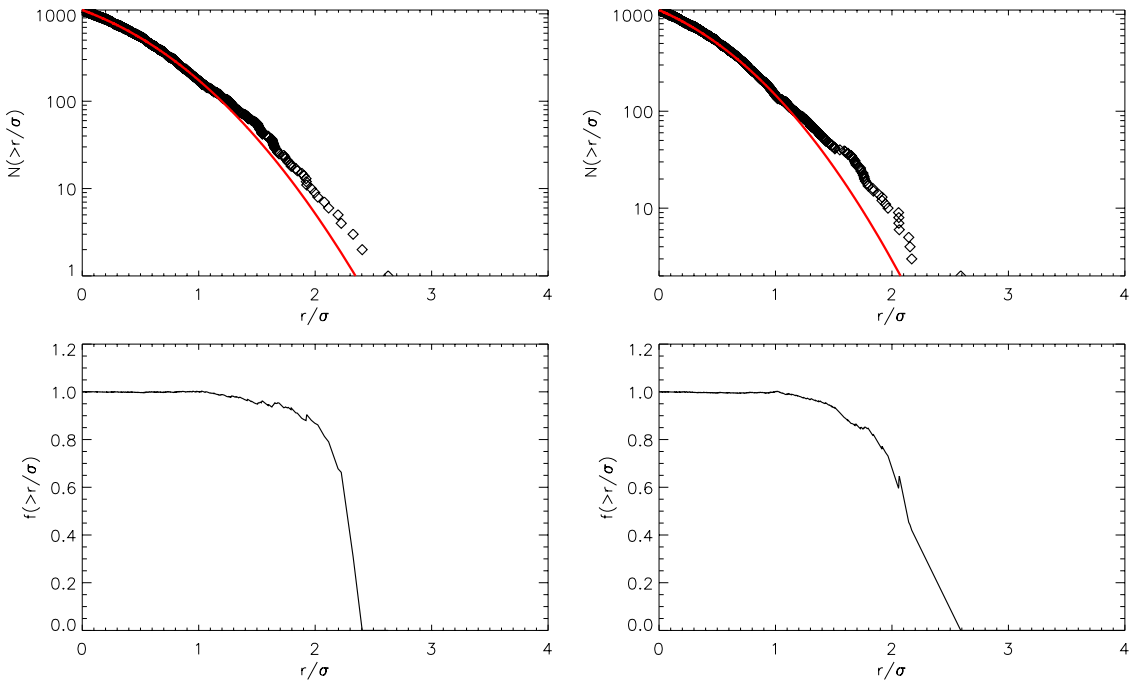

Figure 2.15: *Upper panels: cumulative histograms of $|r_i/\sigma_i|$ for μ Herculis (left for iSONG and right for AUSTRAL). The diamonds show the observed data, and the solid curve shows the result expected for a Gaussian-distributed noise. Lower panel: ratio f of the observed to the expected histograms (left for iSONG and right for AUSTRAL), indicating the fraction of "good" data points. An excess of outliers is evident for $|r_i/\sigma_i| \gtrsim 2$ in the case of iSONG and for $|r_i/\sigma_i| \gtrsim 1.5$ for AUSTRAL.*

range at higher values, namely 800-2000 μHz, as the star is less evolved than β Aquilae.

At first, we searched for a comb-like pattern, following the method used by [48] and the approach described in Section 2.3.4, and found a first value for the mean large separation, $\Delta\nu = 57.76 \pm 0.96\,\mu$Hz, very close to the value found by [48]. The subsequent mode identification for 13 different modes is reported in Table 2.5, as derived from Eq. (2.39) and by means of the échelle diagram shown in Figure 2.17. The notation adopted in Figure 2.17 is similar to that used for Figure 2.10, having the amplitude with the same color-coded background. A total of 6 frequencies has been shifted by the corresponding daily gap of $\pm 11.57\,\mu$Hz in order to be identified according to the asymptotic relation. Only modes whose amplitude was > 0.35 m s^{-1} (or $S/N > 3$) were considered, with a final list of 13 frequencies. As for β Aquilae, the uncertainties on the frequencies were derived by means of the analytical relation adopted by [48] and derived by [237], which now reads

$$\sigma(\nu) = \sqrt{\frac{6}{N}}\frac{1}{\pi T}\frac{\langle\sigma_{\rm v}\rangle}{A} \simeq \frac{0.1}{A_{\rm m/s}}\ \mu{\rm Hz} \tag{2.46}$$

where $\langle\sigma_v\rangle = 2.22$ m s^{-1} and T = 6.28, N = 1106. However, as also stressed in Section 2.3.5, these uncertainties can be several time larger than the true ones, up to the limit of the formal resolution of 1.8 μHz given by the reciprocal of the total observing time of our dataset.

The asymptotic parameters of Eq. (2.39) were derived by means of a linear weighted least-squares fit to the asymptotic relation, as done in Section 2.3.5 and are listed in Table 2.6, together with their 1σ uncertainties. The results for $\Delta\nu$ and ϵ are very close to those obtained by [48], but with a slight different small spacing $\delta\nu_{02}$, although compatible within the error bars.

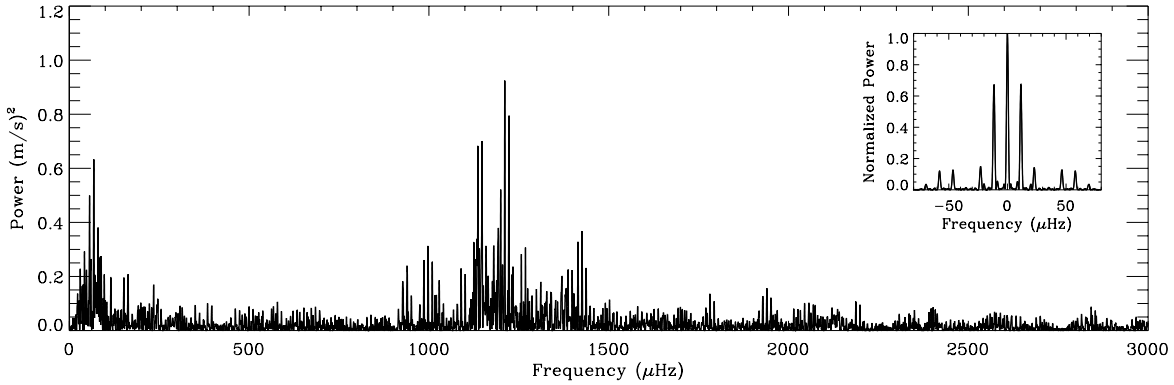

Figure 2.16: *The Power Spectrum of the weighted radial velocity measurements of μ Herculis extracted with the iSONG code from the SARG spectra. An excess of power is clearly visible, with a maximum centered around 1200 μHz. The inset shows the normalized power spectrum of the window function for a sine-wave signal of amplitude 1 m s^{-1}, sampled in the same way as the observations.*

2.5.4 Mode Amplitudes and Lifetimes

The amplitude distribution per radial mode was derived according to the method introduced by [190, 191] and adopted in Section 2.3.6. For this case we computed the amplitude distribution in the range 700-2000 μHz. The maximum amplitude is given by $A_{\max} = 41 \pm 9$ cm s^{-1} and occurs at $\nu_{\max} = 1189\,\mu$Hz. The resulting distribution is plotted in Figure 2.18, where the identification of the radial modes has been marked. Amplitude distributions for dipole and quadrupole modes are derived using the same coefficients adopted in Section 2.3.6. The resulting values for each mode were then used in Eq. (2.46) for estimating the frequency uncertainties, like for β Aquilae.

The resulting maximum amplitude is lower than the value found by [48], who used a different method. Anyhow, according to the results derived in [191] our estimate of the amplitude appears to be reasonable for a subgiant star with $\nu_{\max} \approx 1000\,\mu$Hz.

Lastly, the analysis of the mode lifetimes, performed in an analogous manner to that described in Section 2.3.7, lead us to inconclusive results similarly to the case of β Aquilae. This is likely to rely mainly on the complicated pattern of the mode peaks due to the window function. Hence, longer and more continuous datasets would be required for investigating this aspect of the acoustic modes observed.

2.6 Conclusions

Our observations of β Aquilae show an evident excess of power in the PS region centered at 415 μHz, clearly very well separated from the low-frequency power, and with a position and amplitude that agree with expectations. Although consistently hampered by the single-site window, the comb analysis and the échelle diagram show clear evidence for regularity in the peaks at the spacing expected from the asymptotic theory. The complete identification of six high S/N modes for $\ell = 0, 2$ led to a well-constrained mean large separation of $\Delta\nu = 29.56 \pm 0.10\,\mu$Hz, compatible with a scaled value from the Sun and the value obtained by the power law relation [286], and to a most likely value for the small separation of $\delta\nu_{02} = 2.55 \pm 0.71\,\mu$Hz, whose reliability has yet to be confirmed. The $\ell = 1$ modes found are presumably mixed modes but a theoretical confirmation is needed before adopting these values as real modes of oscillations for this star. Moreover, our results provide a valuable proof that solar-like oscillations in an

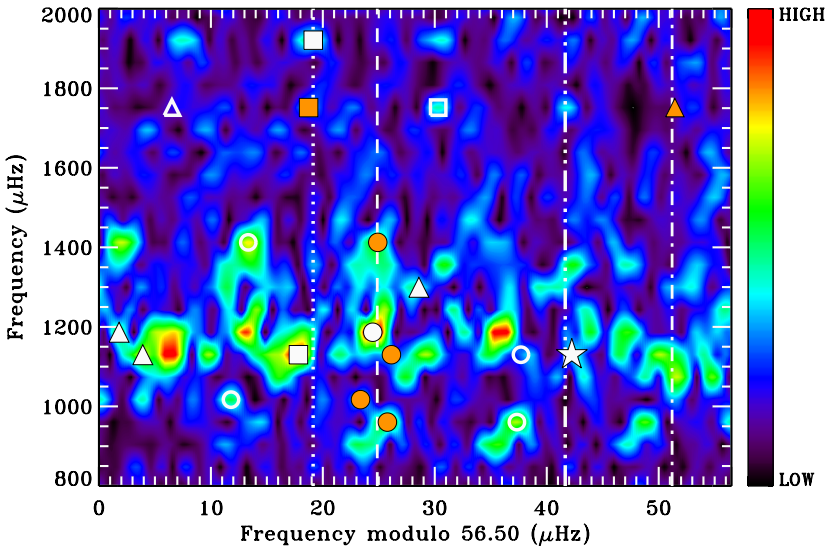

Figure 2.17: *The échelle diagram for μ Herculis overlaid on the amplitude spectrum with a colored background scale. The filled symbols (white and orange) represent the identified modes for $\ell = 0$ (circles), $\ell = 1$ (triangles), $\ell = 2$ (squares) and $\ell = 3$ (stars). The orange symbols are the frequencies shifted for the daily gap of $\pm 11.57\,\mu$Hz while the open symbols correspond to the original unshifted values. The ridges derived from the fit to the asymptotic relation given by Eq. (2.39) are also marked.*

evolved subgiant star show amplitudes that scale as L/M by extrapolating from the Sun.

The revised analysis of μ Herculis by means of the new iSONG time-series led to very similar results to those obtained by [48]. The values of $\Delta\nu = 56.50 \pm 0.06\,\mu$Hz, $\delta\nu_{02} = 5.73 \pm 0.76\,\mu$Hz and $\epsilon = 1.44 \pm 0.02$ appear to be very close to asymptotic parameters derived using the AUSTRAL time-series by [48]. Interestingly, the comparison between iSONG and AUSTRAL shows that iSONG provided radial velocity measurements that suffer less from systematics than those derived by AUSTRAL. In addition, according to Eq. (2.36), we found that the uncertainties in radial velocity were underestimated by AUSTRAL and overestimated by iSONG. This aspect would require further improvements of iSONG for the derivation of the uncertainties.

Therefore, this campaign of observations attained with SARG led to high-precision RV measurements by means of the iSONG code, which was used for the first time in this work. The time-series analysis of the given dataset was able to provide for the first time global asteroseismic parameters and individual p modes together with the evidence for mixed modes. Moreover, this result will be extremely important to develop theoretical models for this star. Multi-site observation campaign with the *SONG* project is highly desirable in a near future. That would then allow us to explore the solar-like oscillations for this target in a detailed way by providing a large number of identified modes. The asteroseismic and astrophysical parameters of this star will then be constrained properly, yielding a deeper understanding of solar-like oscillations in the very difficult region of the HR diagram to which β Aquilae belongs.

Table 2.5: *Mode identification for μ Herculis, in the frequency range 800-2000 μHz. The modes listed show an amplitude > 0.35 m s^{-1} (or S/N > 3). The corrected frequencies reported in the fourth column include the frequencies shifted for the daily gap of ±11.57 μHz. The fifth column represents the uncertainties as derived by means of the analytical relation (2.41). The ℓ = 1 frequencies reported without any radial order are possible avoided crossings.*

ℓ	n	S/N	Corrected frequency[a] (μHz)	Uncertainty[b] (μHz)
0	16	4.5	986.26 (+11.57)	0.38
0	17	3.6	1040.36 (-11.57)	0.34
0	19	3.4	1156.13 (+11.57)	0.25
0	20	8.1	1210.96	0.24
0	24	5,1	1437.41 (-11.57)	0.37
1	-	5.4	1133.92	0.19
1	19	3.9	1181.45 (+11.57)	0.18
1	-	3.2	1188.30	0.18
1	-	3.1	1328.10	0.22
2	18	7.4	1147.82	0.24
2	29	3.2	1770.23 (+11.57)	0.76
2	32	3.2	1940.14	0.71
3	22	3.6	1398.24	0.73

[a] The raw frequencies can be evaluated by adding the daily frequency reported in parentheses.

[b] The actual uncertainties can be several times larger up to the limit of the formal resolution of 1.8 μHz, since the modes are not coherent.

Table 2.6: *Asymptotic parameters for μ Herculis as derived by a linear weighted least-squares fit to the asymptotic relation given by Eq. (2.39).*

Δν (μHz)	$\delta\nu_{02}$ (μHz)	ϵ
56.50 ± 0.06	5.73 ± 0.76[a]	1.44 ± 0.02

[a] The small separation was derived by using the definition from the asymptotic relation and the modes identified for ℓ = 1, 3.

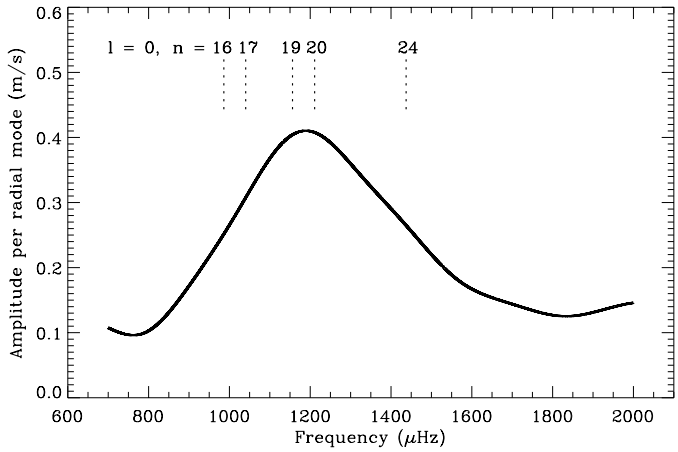

Figure 2.18: *Smoothed amplitude spectrum of μ Herculis showing the amplitude per radial mode computed in the range 700-2000 μHz. The maximum amplitude $A_{max} = 41 \pm 9$ cm s^{-1} occurs at $\nu_{max} =$ 1189 μHz. The positions of the identified ℓ = 0 frequencies as derived from the asymptotic relationship given by Eq. (2.39) are also marked.*

Part II

Ensemble Asteroseismology with *Kepler*

Gilliland, R. L., et al., *Kepler Asteroseismology Program: Introduction and First Results*, 2010, PASP, 122, 131 [128]
Chaplin, W. J., et al., *Ensemble Asteroseismology of Solar-Type Stars with the NASA Kepler Mission*, 2011, Science, 332, 213 [71]

3 The NASA's *Kepler* Mission

The centuries-old quest for other worlds like our Earth has been rejuvenated by the intense excitement and popular interest surrounding the discovery of hundreds of planets orbiting other stars. There is now clear evidence for substantial numbers of three types of exoplanets; gas giants, hot-super-Earths in short period orbits, and ice giants. The challenge now is to find terrestrial planets (i.e., those one half to twice the size of the Earth), especially those in the habitable zone of their stars where liquid water and possibly life might exist.

The NASA's *Kepler* Mission[1], NASA Discovery mission #10, is specifically designed to survey a portion of our region of the Milky Way galaxy to discover dozens of Earth-size planets in or near the habitable zone and determine how many of the billions of stars in our galaxy have such planets. Results from this mission will allow us to place our solar system within the continuum of planetary systems in the Galaxy.

The great interest shown for extrasolar planets detections has incredibly favored the asteroseismology of stars across the entire HR diagram (see Chapter 1), yielding enormous amounts of high quality data, especially subsequently to the launch of the space missions *CoRoT* and *Kepler*. *Kepler* light curves in particular represent one of the main data source used in this dissertation. Thus, in the first part of this chapter we introduce the main features of the NASA's *Kepler* Mission, the spacecraft and instrumentation, and details about its datasets. In the second part of the present chapter we will focus on the importance of the *Kepler* mission for the asteroseismology field, providing a description of the ORK pipeline used to analyze *Kepler* light curves for the research discussed in this thesis, and ending up with a presentation of some of the most important and early results that allowed the investigation of ensemble properties of stars observed in the *Kepler* FOV.

3.1 General Outlook

The NASA's *Kepler* Mission [50] was designed to determine the frequency of Earth-sized planets in and near the habitable zone of Sun-like stars. The habitable zone is the region where planetary temperatures are suitable for water to exist on a planet's surface. During the first 6 weeks of observations, Kepler monitored 156,000 stars, and five new exoplanets with sizes between 0.37 and 1.6 Jupiter radii and orbital periods from 3.2 to 4.9 days were discovered. The density of the Neptune-sized Kepler-4b is similar to that of Neptune and GJ 436b, even though the irradiation level is 800,000 times higher. Kepler-7b is one of the lowest-density planets ($\sim 0.17\,\mathrm{g\,cm^{-3}}$) yet detected. Kepler-5b, -6b, and -8b confirm the existence of planets with densities lower than those predicted for gas giant planets.

[1]The NASA Kepler Mission official website is available at http://kepler.nasa.gov/ and http://www.nasa.gov/kepler.

3.1.1 Scientific Objective

The scientific objective of the *Kepler* Mission is to explore the structure and diversity of planetary systems. This is achieved by surveying a large sample of stars to:

1. determine the abundance of terrestrial and larger planets in or near the habitable zone of a wide variety of stars;
2. determine the distribution of sizes and shapes of the orbits of these planets;
3. estimate how many planets there are in multiple-star systems;
4. determine the variety of orbit sizes and planet reflectivities, sizes, masses and densities of short-period giant planets;
5. identify additional members of each discovered planetary system using other techniques; and
6. determine the properties of those stars that harbor planetary systems.

The *Kepler* Mission also supports the objectives of future NASA Origins theme missions Space Interferometry Mission (SIM) and Terrestrial Planet Finder (TPF),

- by identifying the common stellar characteristics of host stars for future planet searches,
- by defining the volume of space needed for the search and
- by allowing SIM to target systems already known to have terrestrial planets.

3.1.2 Transit Method for Planets detection

When a planet crosses in front of its star as viewed by an observer, the event is called a transit. Transits by terrestrial planets produce a small change in a star's brightness of about 1/10,000 (100 parts per million, ppm), lasting for 1 to 16 hours. This change must be periodic if it is caused by a planet. In addition, all transits produced by the same planet must be of the same change in brightness and last the same amount of time, thus providing a highly repeatable signal and robust detection method.

Once detected, the planet's orbital size can be calculated from the period (how long it takes the planet to orbit once around the star) and the mass of the star using Kepler's Third Law of planetary motion. The size of the planet is found from the depth of the transit (how much the brightness of the star drops) and the size of the star. From the orbital size and the temperature of the star, the planet's characteristic temperature can be calculated. Knowing the temperature of a planet is key to whether or not the planet is habitable (not necessarily inhabited). Only planets with moderate temperatures are habitable for life similar to that found on Earth.

3.1.3 Design

For a planet to transit, as seen from our solar system, the orbit must be lined up edgewise to us. The probability for an orbit to be properly aligned is equal to the diameter of the star divided by the diameter of the orbit. This is 0.5% for a planet in an Earth-like orbit about a Sun-like star (for the giant planets discovered in four-day orbits, the alignment probability is more like 10%). In order to detect many planets, one can not just look at a few stars for transits or even a few hundred. One must look at thousands of stars, even if Earth-like planets are common.

If they are rare, then one needs to look at many thousands to find even a few. *Kepler* looks at more than 100,000 stars so that if Earths are rare, a null or near null result would still be significant. If Earth-size planets are common then *Kepler* should detect hundreds of them.

Considering that the aim is to find planets in the habitable zone of stars like the Sun, the time between transits is about one year. To reliably detect a sequence one needs four transits. Hence, the mission duration needs to be at least three and one half years. Since the *Kepler* Mission has been approved for Mission extension by NASA through fiscal year 2016 [295], it will be able to detect smaller, and more distant planets as well as a larger number of true Earth analogs. It is noticeable already how the total number of planets candidates has increased up to over 2,300 in the first 16 months of observations [18], and it is destined to increase in the extended mission phase.

The *Kepler* instrument is a specially designed 0.95-m diameter telescope called a photometer or light meter (Figure 3.1). It has a very large field of view for an astronomical telescope, 105 square degrees, which is comparable to the area of your hand held at arm's length (Figure 3.2). The fields of view of most telescopes are less than one square degree. *Kepler* needs the large field of view in order to observe the large number of stars. It stares at the same star field for the entire mission and continuously and simultaneously monitors the brightnesses of more than 100,000 stars for at least 3.5 years, the initial length of the mission, which is now extended to more than 6 years.

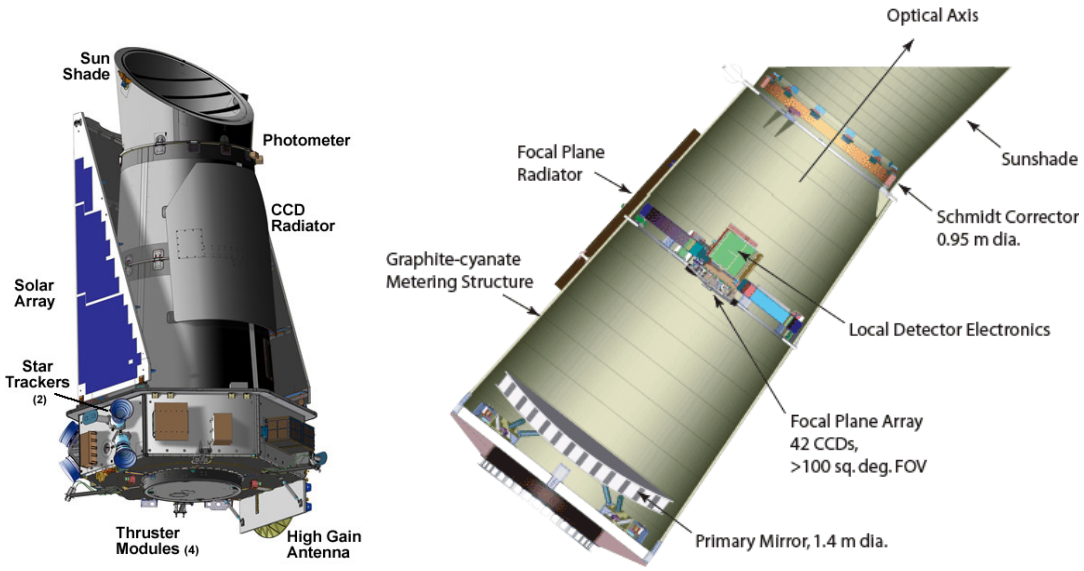

Figure 3.1: *Scheme of the Kepler spacecraft.*

The diameter of the telescope needs to be large enough to reduce the noise from photon counting statistics, so that it can measure the small change in brightness of an Earth-like transit. The design of the entire system is such that the combined differential photometric precision over a 6.5 hour integration is less than 20 ppm (1σ) for a 12^{th} magnitude solar-like star including an assumed stellar variability of 10 ppm. This is a conservative, worse-case assumption of a grazing transit. A central transit of the Earth crossing the Sun lasts 13 hours. And about 75% of the stars older than 1 Gyr are less variable than the Sun on the time scale of a transit.

The photometer must be space-based to obtain the photometric precision needed to reliably see an Earth-like transit and to avoid interruptions caused by day-night cycles, seasonal cycles and atmospheric perturbations, such as, extinction associated with ground-based observing.

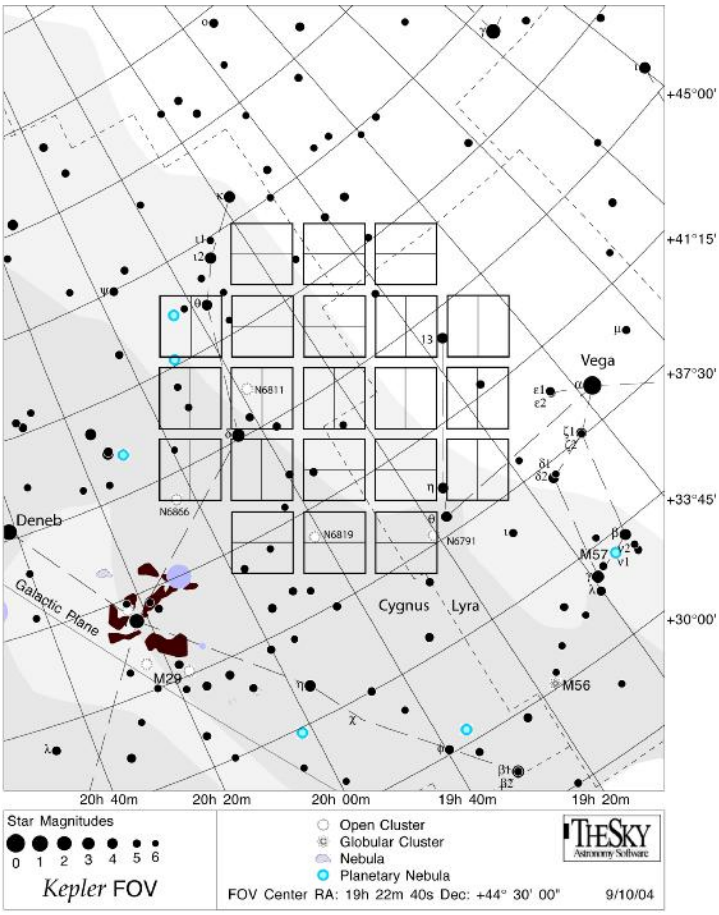

Figure 3.2: *Representation of the Kepler FOV.*

3.1.4 Photometer and Spacecraft

The *Kepler* photometer is a simple single purpose instrument. It is basically a Schmidt telescope design with a 0.95-meter aperture (Figure 3.3) and a 105 deg^2 (about 12 degree diameter) field-of-view (FOV). It is pointed at and records data from just a single group of stars for the four year duration of the mission. The photometer is composed of just one "instrument", which is an array of 42 CCDs (Charge Coupled Devices) (Figure 3.4). Each 50×25 mm CCD has 2200×1024 pixels. The CCDs are read out every 6 seconds to prevent saturation. Only the information from the CCD pixels where there are stars brighter than $m_V = 14$ is recorded (the CCDs are not used to take pictures; the images are intentionally defocused to 10 arc seconds to improve the photometric precision). The data are integrated for 30 minutes.

The instrument has the sensitivity to detect an Earth-size transit of a $m_V = 12$ G2V (solar-like) star at 4σ in 6.5 hours of integration. The instrument has a spectral bandpass from 400 nm to 850 nm, with a central wavelength of 650 nm, as shown in Figure 3.5. Data from the individual pixels that make up each star of the 100,000 main-sequence stars brighter than $m_V = 14$ are recorded continuously and simultaneously. The data are stored on the spacecraft and transmitted to the ground about once a month.

The spacecraft provides the power, pointing and telemetry for the photometer. Pointing at a single group of stars for the entire mission greatly increases the photometric stability and simplifies the spacecraft design. Other than the small reaction wheels used to maintain the pointing and an ejectable cover, there are no other moving or deployable parts The only liquid

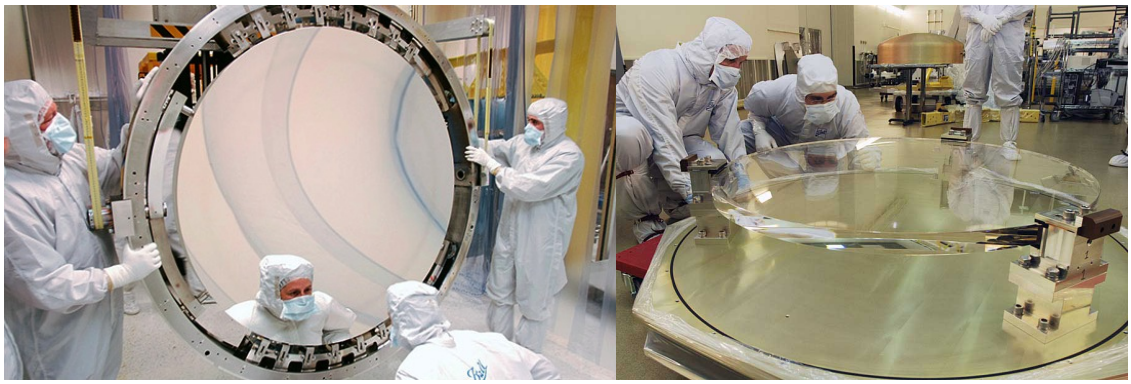

Figure 3.3: *Pictures of the 1.4 m mirror (left) and of the 0.95 m Schmidt corrector (right).*

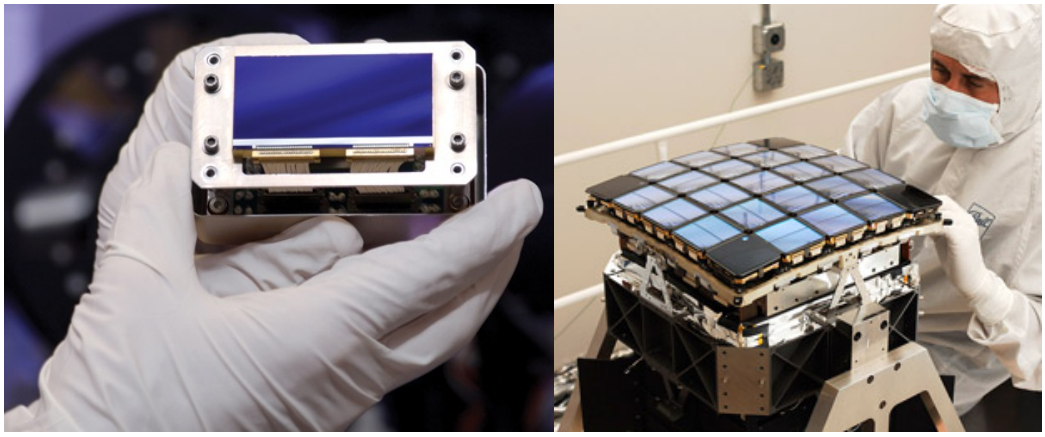

Figure 3.4: *Pictures of the CCD photometer.*

is a small amount for the thrusters which is kept from slosh by a pressurized membrane. This design enhances the pointing stability and the overall reliability of the spacecraft. In particular, the main system characteristics are:

- Spacebased Photometer: 0.95-m aperture.
- Primary mirror: 1.4 meter diameter, 85% light weighted.
- Detectors: 95 mega pixels (21 modules each with two 2200×1024 pixel CCDs).
- Bandpass: 430-890 nm FWHM.
- Dynamic range: 9^{th} to 16^{th} magnitude stars.
- Fine guidance sensors: 4 CCDs located on science focal plane.
- Attitude stability: < 9 milliarcsec, 3σ over 15 minutes.

while the most important mission characteristics can be summarized by

- Continuously point at a single star field in Cygnus-Lyra region except during Ka-band downlink.
- Roll the spacecraft 90 degrees about the line-of-sight every 3 months to maintain the Sun on the solar arrays and the radiator pointed to deep space.

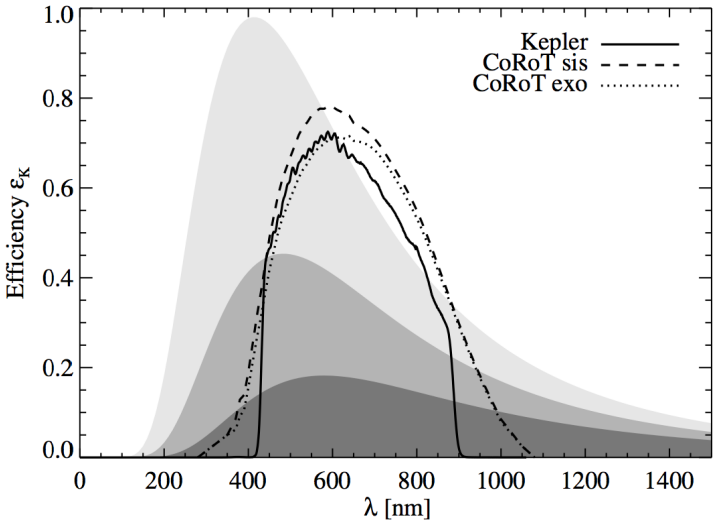

Figure 3.5: *The spectral response ϵ_K of Kepler as a function of the wavelength λ (solid lines), in comparison to that of CoRoT seismo- and exofield (dashed and dotted lines, respectively). Grey profiles indicate the spectra (in arbitrary units) of black bodies with temperatures of 7000, 6000, and 5000 K. Image copied from [12].*

- Monitor 100,000 main-sequence stars for planets.
- Mission lifetime of 3.5 years extended to least 7 years.
- X-band contact twice a week for commanding, health and status.
- Ka-band contact once a month for science data downlink.

3.2 *Kepler* AND ASTEROSEISMOLOGY

Asteroseismology is sometimes considered as the stellar analog of helioseismology [136], being the study of very low amplitude sound waves that are excited by near-surface, turbulent convection, leading to normal-mode oscillations in a natural acoustical cavity. The Sun, when observed as a star without benefit of spatial resolution on its surface, shows ~ 30 independent modes with white light amplitude of a few parts per million (ppm) and periods of 4–8 minutes (see also Chapter 1. In many cases stars with stochastically driven oscillations may deviate strongly from solar size, as in the case of red giants. For our purposes, however, we will broaden the definition of asteroseismology to also include the many types of classical variable stars, e.g. Cepheids, famous for helping to establish the scale of the universe [163].

There are two primary motivations for performing asteroseismology with *Kepler*, which is primarily a planet-detection and characterization mission [50]. First, knowledge of planet properties is usually limited to first order by knowledge of the host star, e.g. *Kepler* easily measures the ratio of planet to star size through transit depths. Turning this into an absolute size for the planet requires knowledge of the host star size which, in favorable cases, asteroseismology is able to provide better than any other approach – in many cases radii can be determined to accuracies near 1%. Second, the instrumental characteristics already required for the exquisitely demanding prime mission [50, 193] can readily support the needs of seismology at essentially no additional cost or modification, which is actually proven by the large number of studies based on the use of *Kepler* light curves.

Oscillations in stars similar to the Sun, which comprise the primary set for the planet detections, have periods of only a few minutes and require use of the Short Cadence (SC) mode [127] with 58.8-second effective integrations. Many of the classical variables may be studied well with the more standard Long Cadence (LC) mode [174], which has 29.4-minute effective integrations.

Kepler comes at a propitious time for asteroseismology. The forerunner missions MOST [221], and especially CoRoT [11, 229], have started the journey taking space-based asteroseismology from decades of promise to the transformative reality that was realistically expected with *Kepler*. The most salient features of the *Kepler* Mission for asteroseismology are: (a) a stable platform from which nearly continuous observations can be made for months to years, (b) cadences of 1 and 30 minutes which support the vast majority of asteroseismology cases, (c) a large 100 square degree field of view providing many stars of great intrinsic interest, (d) a huge dynamic range of over a factor of 10,000 in apparent stellar brightness over which useful asteroseismology (not always of the same type of variable) can be conducted, and (e) exquisite precision that in many cases is well under one ppm for asteroseismology purposes. Initial data characteristics for SC [127] and LC [174] have been shown to support results that nearly reach the limit of Poisson statistics.

The observations conducted by *Kepler* are sectioned into quarters having length of three months each and usually termed as "Q" and the number of the quarter. A quarter itself is also divided into three segments having length of one month each (e.g. Q1.1, Q1.2, Q1.3 for the first quarter). Exceptionally, quarter Q0 had a length of 10 days because it was the commissioning run of the mission, while quarter Q1 was only one month long being that of the survey phase. So far, quarters up to Q13, namely about 36 months of observation ($\sim$ 3 years), are already public for download and can be used for asteroseismic purposes.

3.2.1 The Kepler Input Catalog

The Kepler Input Catalog (KIC)[2] [193, 60] provides knowledge (30 – 50% errors on stellar radii) of likely stellar properties for $\sim$ 4.5 million stars in the Kepler field of view at a level of accuracy necessary to specify the targets to be observed. One application of asteroseismology follows from quantification of stellar radii to more than an order of magnitude better than this for a few thousand giant stars, and several hundred dwarfs, which can then be used to test the KIC entries, and quite possibly provide the foundation for deriving generally applicable improvements to the calibrations enabling redefined entries for the full catalog.

Some 15% of the KIC entries were not classified, thus no radius estimates were available to support selection of stars most optimal for small-planet transit searches. In Q0 (May 2009) and Q1 (May - June 2009) a total of about 10,000 such unclassified stars brighter than $m_{\rm Kep} = 13.8$ were observed for either 10 and/or 33 days respectively. An early application of asteroseismology was to identify stars in this unclassified set that are obviously red giants, a well-posed exercise given the quality of Kepler data [193], and thus allow these to be dropped from further observation in favor of bringing in smaller, and photometrically quieter stars.

In addition, a detailed study conducted by [256], who used *griz* filters from Sloan Digital Sky Survey (SDSS) photometry, has provided new temperature estimates for more than 150,000 stars observed by *Kepler*. A systematic difference of about 200 K between KIC and SDSS temperatures was found. Thus, the catalog of temperature estimates [257] derived by [256] can be a very useful supporting tool for attaining more detailed and reliable investigations of asteroseismic properties of large sample of stars, as we will discuss in Chapter 6.

[2] http://archive.stsci.edu/kepler/kepler_fov/search.php.

3.3 Testing Stellar Astrophysics

The mission as a whole is benefitting from enhanced knowledge of stellar structure and evolution theory. This aspect is advancing by challenging theory with detailed observations of stellar oscillations across a wide range of stellar types. Supporting this goal has led to devoting 0.8% from the available LC target allocations (still a very healthy number of 1,320 stars) to a broad array of classical variables. These LC targets are usually large stars with characteristic periods of variation of hours, to in extreme cases months or even years, and hence are either massive and/or evolved stars. In addition, a set of 1,000 red giants selected to serve as distant reference stars for astrometry [235] provide enticing targets for LC-based asteroseismology. A number of stellar variables, including, of course, close analogs of the Sun can only be probed with use of the SC (1-minute) observations. After the first year a subset of these surveyed targets has been selected for more extended observations. As the mission progressed, a growing fraction of the SC targets has been used to follow up planet detections and candidates, for which better sampling supports transit timing searches for additional planets in the system [155] and oscillation studies of the host stars. In the actual extended mission phase, target selection is not available anymore.

Asteroseismology with *Kepler* is being conducted through the Kepler Asteroseismic Science Consortium (KASC)[3], whose $\sim$ 500 members are organized into working groups by type of variable star. The WG1, related to Solar-like oscillations in main-sequence and subgiant stars is the one of interest for the results presented in this chapter. Chapter 4, instead, deals with targets of interest for the WG2, which studies oscillating stars in Open Clusters, while Chapter 6 is referred to both WG1 and WG8 targets, the latter represented by red giants oscillations.

3.3.1 Solar-like Oscillations

Stars like the Sun, which have sub-surface convection zones, display a rich spectrum of oscillations that are predominantly acoustic in nature (see Chapter 1). As already mentioned in the introductory part of this thesis, the fact that the numerous excited modes sample different interior volumes within the stars means that the internal structures can be probed, and the fundamental stellar parameters constrained, to levels of detail and precision that would not otherwise be possible (e.g. see [134]). Asteroseismic observations of many stars are allowing for multiple-point tests of stellar evolution theory and dynamo theory. They are also important constraints to be placed on the ages and chemical compositions of stars, key information for constraining the evolution of the galaxy (e.g. see [233, 234]). Furthermore, the observations permit tests of physics under the exotic conditions found in stellar interiors, such as those underpinning radiative opacities, equations of state, and theories of convection.

Main Sequence and Subgiant Stars

Kepler observed more than 1500 solar-like stars during the initial survey phase of the asteroseismology program. This allowed the first extensive "seismic survey" performed on this region of the color-magnitude diagram. On completion of the survey, a subset of 50 to 75 solar-like

[3]The Kepler Asteroseismic Investigation (KAI) is managed at a top level by Gilliland, R. L., Brown, T. R., Christensen-Dalsgaard, J., & Kjeldsen, H.. The next level of authorship comprises the KASC working group chairs, and members of the KASC Steering Committee. Data for KASC use first passes through the STScI archive for Kepler, then if SC is filtered to remove evidence of any transits, and then is made available to the KASC community from the Kepler Asteroseismic Science Operations Centre (KASOC) at the Department of Physics and Astronomy, Aarhus University, Denmark. Astronomers wishing to join KASC are welcome to do so by following the instructions at: http://astro.phys.au.dk/KASC/.

targets have been selected for longer-term, multi-year observations. These longer datasets allow tight constraints to be placed on the internal angular momenta of the stars, and also enable "sounding" of stellar cycles via measurement of changes to the mode parameters over time (e.g. see [182, 220], and Chapter 7 for an application to *Kepler* light curves for studying differential rotation in active stars).

Figure 3.6 showcases the potential of the *Kepler* data for performing high-quality asteroseismology of solar-like stars from SC data [70]. The left-hand panels show frequency-power spectra of three 9^{th}-magnitude, solar temperature targets observed during Q1. All three stars have a prominent excess of power showing a rich spectrum of acoustic modes. The insets show near-regular spacings characteristic of the solar-like mode spectra, and highlight the excellent S/N observed in the individual mode peaks . The sharpness of the mode peaks indicates that the intrinsic damping from the near-surface convection is comparable to that seen in solar p modes. The p modes sit on top of a smoothly varying background that rises in power towards lower frequencies. This background carries signatures of convection and magnetic activity in the stars. We see a component that is most likely due to faculae — bright spots on the surface of the stars formed from small-scale, rapidly evolving magnetic field. This component is manifest in the spectra of the top two stars as a change in the slope of the observed background just to the low-frequency side of the p-mode envelope (see arrows). All three stars also show higher-amplitude components due to granulation, which is the characteristic surface pattern of convection. The near-regularity of the oscillation frequencies allows us to display them in so-called échelle diagrams (see Chapter 2 for a definition), in Figure 3.6 (right-hand panels, see also Chapter 1). Here, the individual oscillation frequencies have been plotted against their values modulo $\Delta\nu$ — the average large frequency spacing, see Eq. (1.11). The frequencies align in three vertical ridges that correspond to radial, dipole, and quadrupole modes. As already discussed in Chapter 1, by making use of the individual frequencies and the mean spacings we are able to constrain the masses and radii of the stars to within a few percent. The top two stars are both slightly more massive than the Sun (by about 5 %), and also have larger radii (larger by about 20 % and 30 % respectively). The bottom star is again slightly more massive than the Sun (10 %), and about twice the radius. It has evolved off the main sequence, having exhausted the hydrogen in its core. The ragged appearance of its dipole-mode ridge labeled avoided crossing in Figure 3.6, is a tell-tale indicator of the advanced evolutionary state, placing the star in the subgiant phase of the evolution; the frequencies are displaced from a near-vertical alignment because of evolutionary changes to the deep interior structure of the star.

Red Giant Stars

Red giants have outer convective regions and are expected to exhibit stochastic oscillations that are solar-like in their general properties but occur at much lower frequencies (requiring longer time series). The first firm discovery of solar-like oscillations in a giant was made using radial velocities by [113]. However, it was only recently that the first unambiguous proof of non-radial oscillations in G and K giants was obtained, using spaced-based photometry from the CoRoT satellite [96]. This opened up the field of red giant seismology, which is particularly interesting because important uncertainties in internal stellar physics, such as convective overshooting and rotational mixing, are more pronounced in evolved stars because they accumulate with age.

The extremely high S/N photometry of the *Kepler* observations brings red giant seismology to the next level. With the first 43 days of LC data it was possible to detect oscillations with $\nu_{\rm max}$ ranging from 10 μHz up to the Nyquist frequency around 280 μHz. The results include the first detection of oscillations in low-luminosity giants with $\nu_{\rm max} > 100\,\mu$Hz [30]. These giants

are important for constraining the star-formation rate in the local disk [231]. In addition, *Kepler* power spectra have such a low noise level that it is possible to detect $\ell = 3$ modes (see e.g. [30, 165, 87]), hence increasing the available asteroseismic information. The large number of giants that *Kepler* continuously monitors for astrometric purposes during the entire mission is allowing pioneering research on the long-term interaction between oscillations and granulation [219]. Moreover, as it was already expected from the beginning, the frequency resolution provided by *Kepler* was sufficient to detect rotational splitting in the fastest rotating giants [22, 23], and possibly allow the measurement of frequency variations due to stellar evolution on the red giant branch.

3.4 Ensemble Asteroseismology and First Results

Although not representing the primary goal of the NASA's *Kepler* Mission, one of the main contributes comes from Asteroseismology. In fact, with the great advantage of high quality photometry and continuous observations of large numbers of stars, *Kepler* is allowing for a thorough study of fundamental stellar properties of thousands of pulsating stars, thus contributing to the previous work done by the *CoRoT* satellite (see e.g. [96, 178, 243]).

In this second part of the current chapter we describe the pipeline developed at Catania Astrophysical Observatory (OACT) for the measurement of the mean global asteroseismic parameters $\langle\nu_{\rm max}\rangle$ and $\langle\Delta\nu\rangle$ (see Chapter 1 for a definition) for a number of $\sim$ 800 main-sequence to subgiant stars observed by *Kepler* and showing solar-like oscillations. We then briefly present the main works that adopted the global parameters derived with the ORK pipeline. The ensemble asteroseismology, namely the study of global asteroseismic properties of large sample of stars, is the main character of the works that used our estimates of $\nu_{\rm max}$ and $\Delta\nu$. The ensemble asteroseismology relies on the adoption of scaling relations, namely simple empirical laws used to derive the fundamental stellar properties from our knowledge about the Sun. In particular, the most useful and widely adopted scaling relations are those given by Eqs. (2.43) and (2.45) for $\Delta\nu$ and $\nu_{\rm max}$ respectively (see also [72]). As already mentioned in the introduction of this thesis, these two asteroseismic quantities are directly related to mass, radius and effective temperature of the star. By combining the two equations mentioned beforehand and assuming to have the estimates of both $\nu_{\rm max}$ and $\Delta\nu$, and a measure of temperature of the star either from photometry or spectroscopy, we can estimate its mass and radius up to 2-3 % of precision (see [71]).

3.4.1 The ORK Pipeline

The ORK pipeline (@OACT), is an IDL based pipeline that searches regular patterns of p-mode oscillations in stars selected by the KASC WG1 at the beginning of the *Kepler* mission. An original version of the pipeline is described in [48] and was adopted for the single target μ Herculis in the study of its asteroseismic properties from a set of ground-based observations (see also Chapter 2, for further discussion). The *Kepler* light curves used by the ORK pipeline span from Q1 up to Q3 and are those corrected for KASC members through the method described by [122]. Some additional corrections were also applied: removing "bad points" from the light curves, i.e. those points that deviated by 0.5σ from the mean flux value, σ being the scatter in the observed flux around its mean.

After this preliminary phase, we convert the flux from e^-/cadence, which is a default unit for the pre-processed *Kepler* light curves format [127] of SC, to ppm, i.e. parts per million of

variation in magnitude. From the simple Pogson relation for apparent magnitudes

$$m = -2.5 \log_{10} F \tag{3.1}$$

where F is the flux in e^-/cadence, we evaluate error bars in magnitude through the propagation of the errors

$$\sigma_{\mathrm{m}} = \frac{2.5}{\ln 10} \frac{\sigma_{\mathrm{F}}}{F} \tag{3.2}$$

where σ_{F} is the error on flux from *Kepler* light curves and it is again expressed in e^-/ cadence. Afterwards, we subtract the mean magnitude value from the light curve and multiply the result by 10^6, in order to obtain units of ppm.

Thus, from the light curves cleaned from bad points and ppm units, we compute the amplitude spectrum by means of the LSS method described in Chapter 2. In this case, either adopting a weighted or an unweighted LSS, lead to the same result. Thus, we decided to compute the unweighted amplitude spectrum for simplicity.

The general procedure used in the ORK pipeline guarantees a half-automated extraction of the asteroseismic parameters $\nu_{\max}$ and $\Delta\nu$, and it can be summarized in the following steps:

1. load iteratively the *Kepler* light curve by selecting the star from a list given by WG1
2. apply the corrections mentioned above to the selected light curve
3. compute the LSS in order to derive the corresponding PS (i.e. power spectrum) in a range of frequencies from 50 to 4000 μHz
4. check by eye whether the selected star shows a significant excess of power in that frequency range
5. if positive, continue the analysis to point (6), otherwise discard the star and go to the next one starting from point (2)
6. multiply the value in power of the PS at each bin by the value of the frequency at that bin. This enhances the power excess of the solar-like oscillations for a sufficiently large $\nu_{\max}$ ($\gtrsim 400\,\mu$Hz) and therefore allows us to identify its position in an easier manner
7. apply a Gaussian smoothing, having a FWHM $\approx 50\,\mu$Hz, to the entire PS in order to get rid of the frequency peaks and search for the maximum of the power excess
8. find the maximum of the smoothed PS and save the frequency corresponding to the maximum as our estimate of the mean frequency of maximum power, $\langle\nu_{\max}\rangle$
9. compute a new PS in the region of the power excess by considering the range $\langle\nu_{\max}\rangle \pm \langle\nu_{\max}\rangle/2\pi$ (supposed to be wide enough to contain the entire power excess caused by oscillation signal)
10. CLEAN a number of 20 peaks (can be more if desired) from the last PS computed and save their corresponding frequencies
11. search for a comb-like pattern using the 20 CLEANed peaks and the CR function defined by Eq. (2.40) (see also Chapter 2 or [48] for details)
12. find the mean large separation $\langle\Delta\nu\rangle$ resulting directly from the maximum of the CR function.

The choice of the FWHM of the Gaussian smoothing as well as the number of CLEANed peaks followed from tests done on a small number of stars ($\sim$ 20) having values of $\nu_{\rm max}$ spanning from 400 to 3500 μHz. The results derived with this pipeline, though without uncertainties, contributed to several projects of the KASC, related to both WG1 [71, 72, 73, 216, 310, 63, 275, 218, 317, 45], and joint WG1–WG8 [167]. We will briefly show some of the main results and list all the , since a detailed description of their analyses goes beyond the scope of this thesis.

3.4.2 Results

The first work that adopted our estimates of $\langle\nu_{\rm max}\rangle$ and $\langle\Delta\nu\rangle$ is described in [71]. This allowed, together with the results arising from other pipelines (e.g. see [310] for a general presentation of the different methods), to derive asteroseismic masses and radii for a sample of $\sim$ 2000 stars and to compare the distribution in mass and radius with theoretical computations made by means of the TRILEGAL code [130]. Thanks to the scaling relations given by Eqs. (2.43) and (2.45) we can derive the fundamental stellar properties of these stars in a very straightforward manner. This method is usually known as the "direct" method of estimation of stellar properties (see also [71, 72, 310] for more details) and has the great advantage of being independent of any stellar evolutionary model, hence of great benefit for instructive comparisons with the population synthesis models. Another method that is often adopted is termed the "grid-based" method of estimation, which relies on the adoption of several theoretical models for predicting $\Delta\nu$, having different input values of mass and radius (see e.g. [71, 17]).

Panel (A) of Figure 3.7 shows the HR diagram of the stars studied by [71] with luminosities computed from the scaling relations of $\nu_{\rm max}$ and $\Delta\nu$. Combining together Eqs. (2.43) and (2.45) we obtain the asteroseismic mass and radius

$$\frac{M}{M_\odot} \simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{3} \left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-4} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{1.5}, \tag{3.3}$$

$$\frac{R}{R_\odot} \simeq \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right) \left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{-2} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{0.5}, \tag{3.4}$$

where the values $\nu_{\rm max,\odot} = 3150\,\mu$Hz, $\Delta\nu_\odot = 134.9\,\mu$Hz and $T_{\rm eff,\odot} = 5777\,$K were used. From the error propagation law we can derive the corresponding uncertainties

$$\sigma_M \simeq M\sqrt{9\left(\frac{\sigma_{\nu_{\rm max}}}{\nu_{\rm max}}\right)^2 + 16\left(\frac{\sigma_{\Delta\nu}}{\Delta\nu}\right)^2 + 2.25\left(\frac{\sigma_{T_{\rm eff}}}{T_{\rm eff}}\right)^2}, \tag{3.5}$$

$$\sigma_R \simeq R\sqrt{\left(\frac{\sigma_{\nu_{\rm max}}}{\nu_{\rm max}}\right)^2 + 4\left(\frac{\sigma_{\Delta\nu}}{\Delta\nu}\right)^2 + 0.25\left(\frac{\sigma_{T_{\rm eff}}}{T_{\rm eff}}\right)^2}. \tag{3.6}$$

According to these relations, the uncertainty on the scaled mass is usually about twice larger than that in radius, as also noted by [71]. These scaling relations have been proven to provide reliable estimates of masses and radii for stars homologous to the Sun (up to the subgiant phase), but start to deviate from the real stellar parameters for stars that are already ascending the RGB (see e.g. [232]). However, they have been used extensively for deriving fundamental stellar properties of thousands of main sequence stars, providing successful results.

Panel (B) of Figure 3.7 shows a similar plot to that of panel (A) by using the asteroseismic quantity $\Delta\nu$ instead of luminosity, where the former, as clearly visible, proves to decrease as the stars evolve. The red points marked in both panels represent some cases discussed by [71] and

were used to show that the frequency of maximum power $\nu_{\rm max}$ decreases with the increasing age of the star.

An interesting result of this work arises by comparing the distribution of the asteroseismic masses derived from Eq. (3.3) to that obtained by the population synthesis code TRILEGAL [129]. The result is shown in Figure 3.8, where the observed distributions are represented by black lines while the synthetic ones are indicated in red. As it appears evident, the matching is rather good in the case of radii distributions (bottom panel), while there is a significant lack of dwarfs in the observed masses (top panel) with respect to the models. Indeed, the low number of solar twins in the *Kepler* FOV constitutes one of the main sources of bias for the asteroseismology of the stars observed.

Other works that used the global asteroseismic parameters derived by means of the ORK pipeline are listed below:

- predicting the detectability of oscillations in Solar-Type stars [72]
- evidence for the impact of stellar activity on the detectability of solar-like oscillations [73]
- solar-like oscillations in KIC 11395018 and KIC 11234888 from 8 months of *Kepler* data [216]
- global asteroseismic properties of solar-like oscillations: a comparison of complementary analysis methods [310]
- asteroseismology from multi-month *Kepler* photometry: the evolved Sun-like stars KIC 10273246 and KIC 10920273 [63]
- constructing a one-solar-mass evolutionary sequence using asteroseismic data [275]
- asteroseismic diagrams from a survey of solar-like oscillations with *Kepler* [317]
- testing scaling relations for solar-like oscillations from the Main Sequence to Red Giants [167]
- calibrating convective properties of solar-like stars in the *Kepler* FOV [45]
- Seismic Analysis of Four Solar-like Stars Observed during More Than Eight Months by Kepler [218].

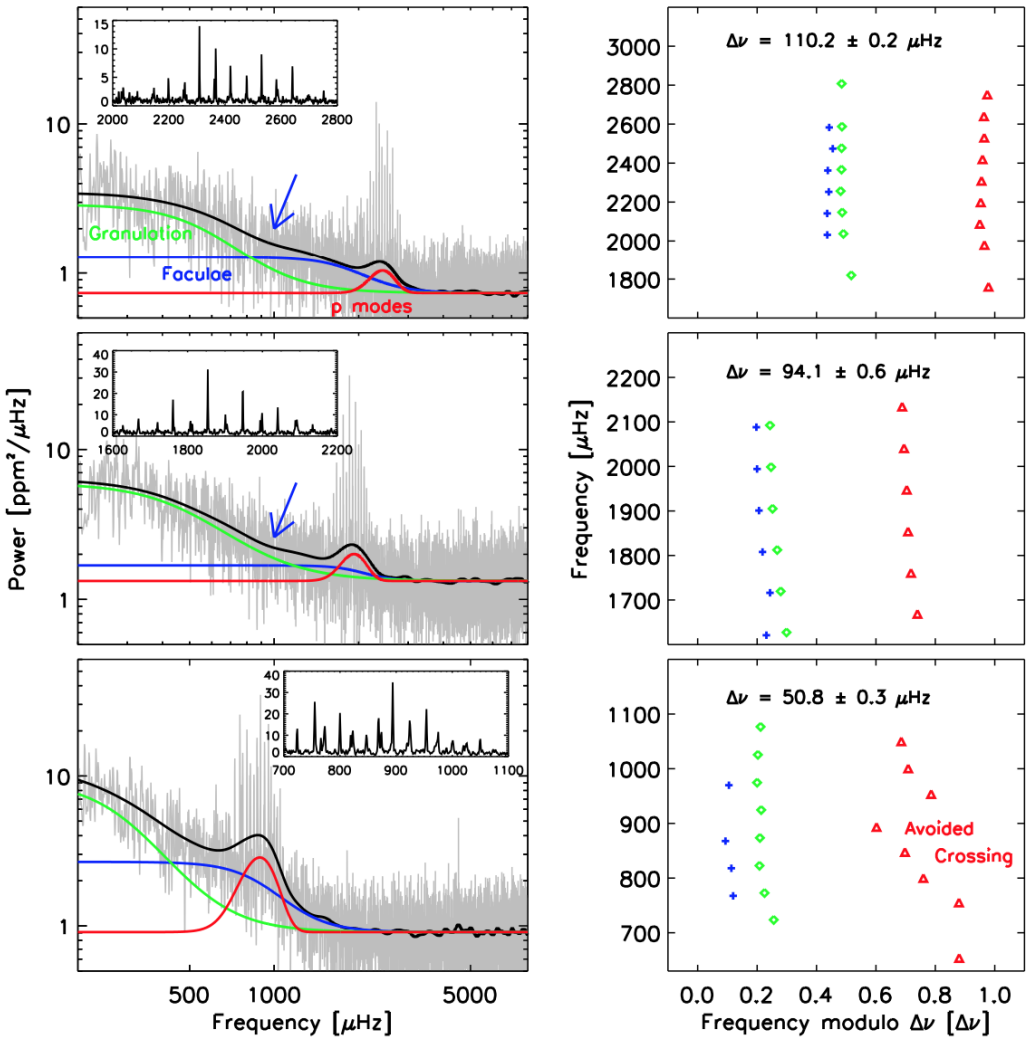

Figure 3.6: *Left-hand panels: Frequency-power spectra of Kepler photometry of three solar-like stars (grey) over 200–8000 μHz. The thick black lines show the result of heavily smoothing the spectra. Fitted estimates of the underlying power spectral density contribution of p modes, bright faculae and granulation as labelled in the top left panel are also shown; these are color coded red, blue and green respectively in the on-line version. These components sit on top of a flat contribution from photon shot noise. The arrows mark a kink in the background power that is caused by the flattening toward lower frequencies of the facular component. The insets show the frequency ranges of the most prominent modes. Right-hand panels: So-called échelle plots of individual mode frequencies. Individual oscillation frequencies have been plotted against the frequencies modulo the average large frequency spacings (with the abscissa scaled to units of the large spacing of each star). The frequencies align in three vertical ridges that correspond to radial modes ($\ell = 0$, diamonds), dipole modes ($\ell = 1$, triangles) and quadrupole modes ($\ell = 2$, crosses). Image copied from [128].*

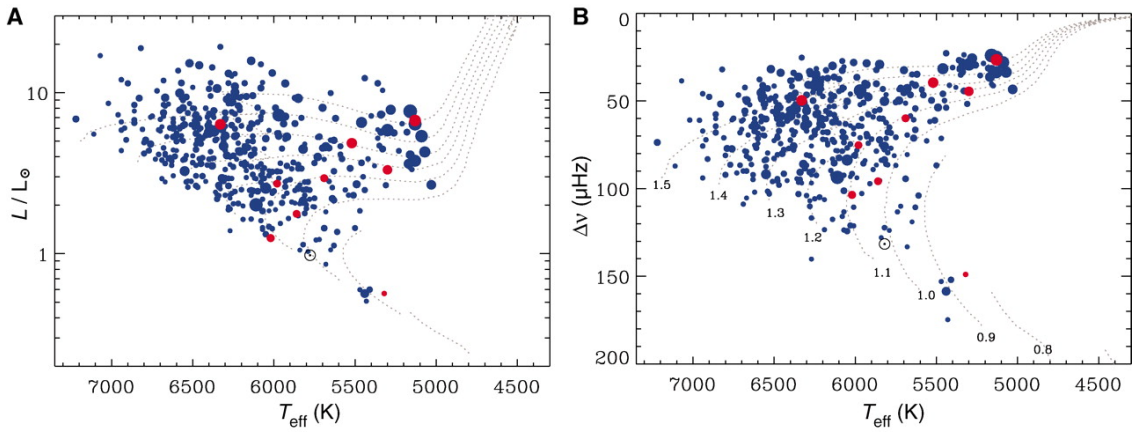

Figure 3.7: *Panel (A): Estimates of the luminosities of the stars (in units of the solar luminosity) of the ensemble of Kepler stars showing detected solar-like oscillations, plotted as a function of effective temperature, $T_{\rm eff}$. Stars with red symbols are those described in [71]. Panel (B): Average large frequency separations, $\Delta\nu$, against $T_{\rm eff}$. The symbol sizes are directly proportional to the prominence of the detected oscillations (i.e., the signal-to-noise ratios). These ratios depend both on stellar properties (e.g. the photometric amplitudes shown by the oscillations, and the intrinsic stellar backgrounds from convection) and the apparent brightness of the stars. The dotted lines show predicted evolutionary tracks [212] for models of different stellar mass (0.8 to 1.5 $M_{\odot}$, in steps of 0.1 $M_{\odot}$). The Sun is marked with a solar symbol (⊙).*

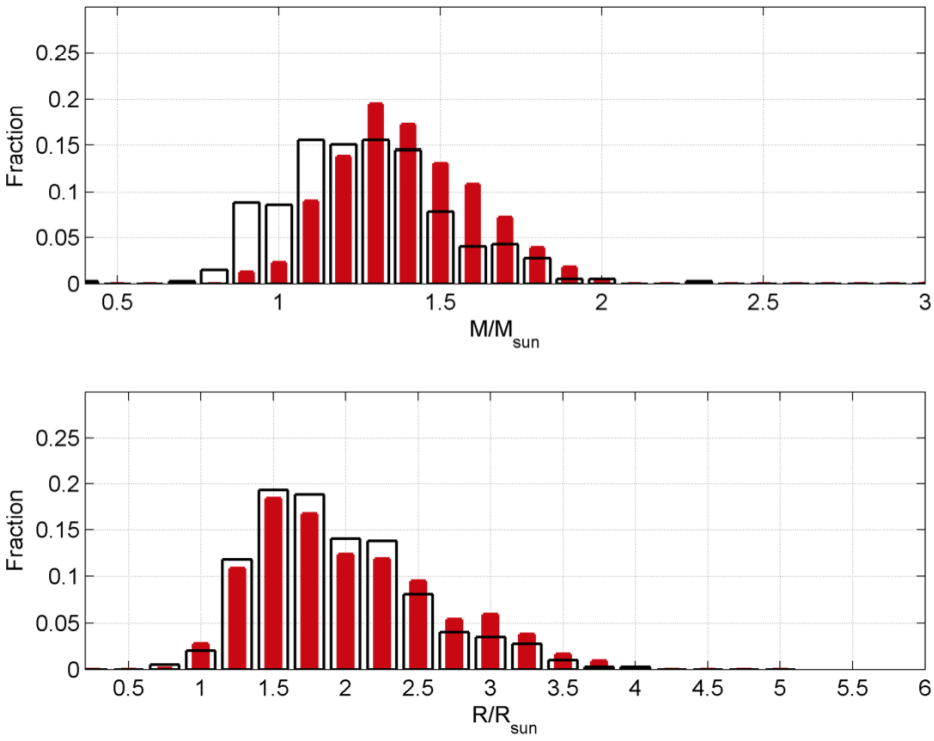

Figure 3.8: *Black lines: Histograms of the observed distribution of masses (top) and radii (bottom) of the Kepler ensemble. In red, the predicted distributions from population synthesis modeling, after correction for the effects of detection bias. The population modeling was performed using the TRILEGAL code [130].*

4 Pulsating Red Giants in Open Clusters

Many recent studies of solar-like oscillations of red giant stars have been focused on large ensembles of stars, made possible by the flood of high quality photometric data provided by the space missions CoRoT [96, 178, 243]. and *Kepler* [50, 193, 30]. Particular attention has been given to the three open clusters NGC 6791, NGC 6811, and NGC 6819 in the *Kepler* field [289, 17, 150, 291, 292, 232], due to the well-known advantage of cluster stars sharing common properties, which allows for more stringent investigations of stellar evolution theory.

Among the highlights in recent results relevant for our study are the measurements of the small frequency separations $\delta\nu_{02}$, $\delta\nu_{01}$ and of the dimensionless term ϵ, their correlation with the large frequency separation (see [30, 165, 179] for previous results on field red giant stars) and their dependence on stellar mass [236, 179]. Also, the results on the ensemble échelle diagrams have allowed for the investigation of ensemble properties of the modes, including the measurement of the mean small spacing $\delta\nu_{03}$ [30, 165] and the linewidths of the dipole modes and their correlation to fundamental stellar properties [69, 19, 7, 35], which are important for the comprehension of the physics responsible for the excitation and damping of solar-like oscillations. Finally, the period spacing analysis for the investigation of the evolutionary stage of red giants [32, 242] now allows us to distinguish between He-core burning red giants and those only burning hydrogen in a shell.

Here, we study 115 red giants belonging to the above mentioned clusters, continuously observed for 19 months by the NASA *Kepler Mission* (see [175, 128, 60] for details on the data pipeline and acquisition, for a general introduction to the asteroseismic program, and for a description of the Kepler Input Catalog, respectively). Our study is made along the same lines as those described by [165], who analyzed the first 4.5 months of *Kepler* observations of field red giants. In contrast to [165], our cluster red giants have the great advantage of providing more homogeneous samples because age, metallicity and mass [17, 232], are about the same. In particular, NGC 6791 is a very old, (~ 8.3 Gyr [56]), and metal rich ([Fe/H] $= 0.29 \pm 0.03$ (random) ± 0.07 (systematic) [55]) open cluster, with average masses $M_{\rm RGB} = 1.20 \pm 0.01\, M_{\odot}$ [17] and $M_{\rm RC} = 1.15 \pm 0.03\, M_{\odot}$ [232] for red giant branch (RGB) and red clump (RC) stars respectively (see also [56] for recent results from eclipsing binaries). NGC 6819 is a middle aged (2-2.4 Gyr [17]) open cluster, with solar metallicity ([Fe/H] $= 0.09 \pm 0.03$ [51]), and average masses $M_{\rm RGB} = 1.68 \pm 0.03 M_{\odot}$ and $M_{\rm RC} = 1.65 \pm 0.04$ for RGB and RC stars respectively. The third open cluster, NGC 6811, is characterized by a young (0.7 ± 0.1 Gyr [131]) and possibly solar metallicity star population (suggested by two independent spectroscopic investigations by Bruntt et al., in prep, and Molenda-Żakowicz et al., in prep.), where a small number of RC stars has been observed, showing an average mass $M_{\rm RGB} = 2.35 \pm 0.04\, M_{\odot}$ [291, 292]. The temperature estimates for both NGC 6791 and NGC 6819 were derived by [150]. In particular, they used color-temperature calibrations by [262] and JHK photometry from the 2MASS catalog [277], which is available for all the stars of the sample. V magnitudes are taken from

[294] for NGC 6791 and from [154], in order to derive temperatures based on the $(V-K)$ color. The adopted reddenings are $E(B-V) = 0.16 \pm 0.02$ for NGC 6791 [55] and $E(B-V) = 0.15$ for NGC 6819 [51]. Lastly, [17] estimated DMs for both NGC 6791 and NGC 6819 by adopting the extinction $A_V = 3.1E(B-V)$, which yielded to $(m-M)_0 = 13.11 \pm 0.06$ and $(m-M)_0 = 11.85 \pm 0.05$, respectively.

After briefly introducing the parameters involved in our study in Section 4.1, we introduce the code developed for this work in Section 6.1 (see Appendix A for all the details), which concerns the analysis of the average p-mode structure in the power spectrum. In Section 4.3 we show the resulting asteroseismic ensemble diagrams and the linewidths of radial modes as a function of fundamental stellar parameters, while Section 4.4 presents the analysis of the period spacing of mixed dipole modes following the approach of [32]. Finally, we conclude in Section 4.5.

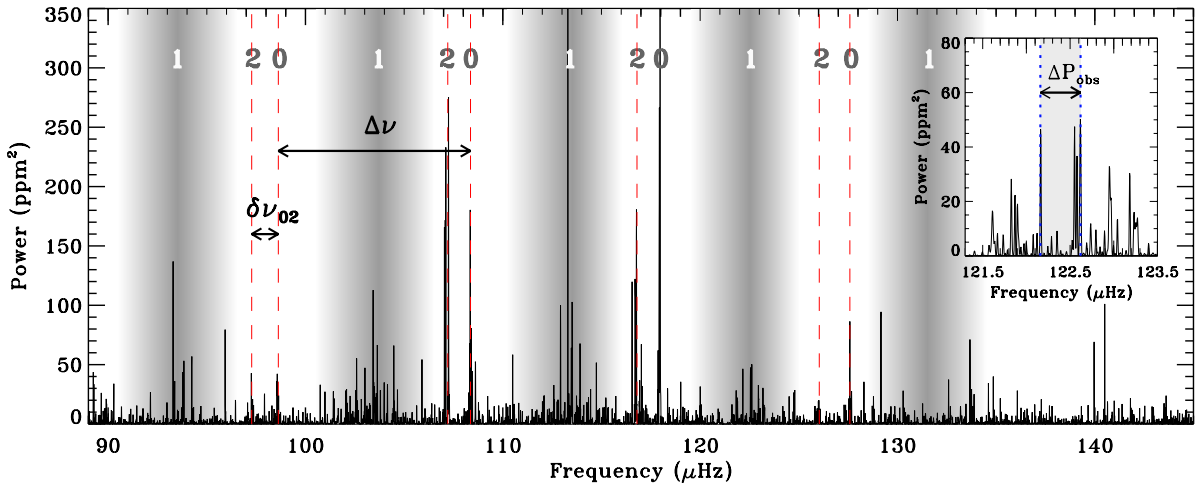

Figure 4.1: *Power spectrum of KIC 2436593, a typical low luminosity RGB star belonging to NGC 6791. Mode identification for some of the peaks is shown. Shaded regions in gray indicate mixed $\ell = 1$ modes. $\Delta\nu$ and $\delta\nu_{02}$ are also marked. The inset shows the detail of one of the $\ell = 1$ shaded regions, where an indication of the observed period spacing, $\Delta P_{\rm obs}$, is shown.*

4.1 Asymptotic Parameters Investigated

Before proceeding with the description of the data analysis, it is helpful to introduce the physical quantities that we study in this work. As we remind from the introduction of this thesis given in Chapter 1, the asymptotic theory of solar-like oscillations show us that the acoustic standing waves (also known as p modes) with low angular degrees, ℓ, and high radial orders, n, are regularly spaced in frequency, according to the relation given by Eq. (1.10) (see also [308, 299, 133]). The term $\delta\nu_{0\ell}$, known as the small frequency separation, for $\ell = 1, 2$ and 3 it is defined as

$$\delta\nu_{02} = \nu_{\mathrm{n},0} - \nu_{\mathrm{n}-1,2}, \tag{4.1}$$

$$\delta\nu_{01} = \frac{1}{2}\left(\nu_{\mathrm{n},0} + \nu_{\mathrm{n}+1,0}\right) - \nu_{\mathrm{n},1}, \tag{4.2}$$

$$\delta\nu_{03} = \frac{1}{2}\left(\nu_{\mathrm{n},0} + \nu_{\mathrm{n}+1,0}\right) - \nu_{\mathrm{n},3}. \tag{4.3}$$

As stressed in Chapter 1, Section 1.2, the small frequency separations are related to the sound speed gradient in the stellar core, hence to the mean molecular weight, which increases as the star evolves. Mixed modes occur as the frequencies of the g modes in the core and the p modes in the envelope become similar during the subgiant and red giant phase. As the star evolves, its mixed modes will undergo avoided crossings causing so called mode bumping, which broadens the ridges in the échelle diagram (e.g. see [293]). As argued by [30] and [165] the small spacing $\delta\nu_{03}$ is therefore preferred in red giants over the more conventional $\delta\nu_{13}$. We note that mode bumping mostly affects the dipole modes as they penetrate deeper into the star, and hence couple more strongly to the g modes in the core [103], see also Figure 4.2. The dipole modes are therefore sensitive to the core properties of the star, which allows us to determine which red giants burn helium or not (see Section 4.4).

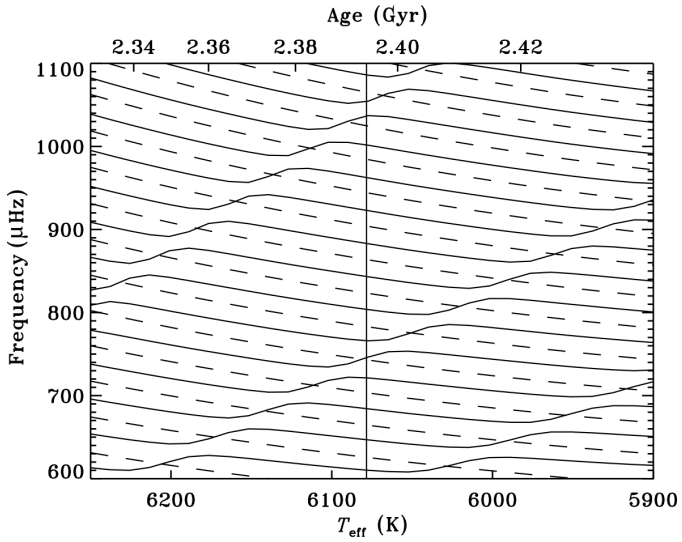

Figure 4.2: *Example of avoided crossing occurring in low degree p modes as the star evolves. Evolution of adiabatic frequencies with age of a model of mass 1.60 $M_\odot$, where age is measured by the effective temperature. The dashed lines correspond to modes of degree $\ell = 0$, and the solid lines to $\ell = 1$. The vertical solid line indicates the $T_{\rm eff}$ of η Bootis. As it is clear from the mode frequency evolution, the dipole modes are bumped at some times, resulting in an increase of their frequency with respect to the value predicted by the asymptotic relation given by Eq. (1.10). Image copied from [25].*

Figure 4.1 shows a typical power spectrum of a low luminosity RGB star, KIC 2436593, observed in NGC 6791. The mode identification for some of the modes is shown, together with the indication of regions containing mixed $\ell = 1$ modes, represented by the gray-shaded strips. The large separation $\Delta\nu$ and the small spacing $\delta\nu_{02}$ are indicated as well. The inset shows a zoom-in of one of the gray-shaded strips, where the observed period spacing of the dipole modes, $\Delta P_{\rm obs}$ is marked.

4.2 Observations and Data Analysis

The photometric time series of the 115 red giants used in this work were obtained in *Kepler*'s long cadence mode ($\Delta t \sim 30$ min, [174]) between 2009 May 13 and 2010 December 22. This corresponds to the observing quarters 1−7, providing a total of almost 18,000 data points per star (see [122] for details on the detrending of the data). We followed the approach described

by [292] for merging the quarters, and we discarded the stars that they classified as seismic non-members in their study. We also note that, according to their classification based on the color-magnitude diagram (CMD), the cluster stars in NGC 6811 are all He-burning stars, with one star appearing to be in a late He-core burning phase towards the asymptotic giant branch (AGB).

For the present study we developed the Asymptotic Analysis of Red Giants (AARG) code (Appendix A), with the purpose of deriving asymptotic parameters for p modes and observed period spacings for mixed modes in red giant stars (see Appendix A for more details). AARG performs a multi-step analysis in a semi-interactive way, allowing the user to follow the results at each step and make any necessary corrections. We calculated background-corrected power spectra and measured $\Delta\nu$ using the SYD pipeline [164]. As a check we compared $\Delta\nu$ values with those derived using other methods [240, 215, 149, 177] and found good agreement. We focus first on the analysis of p modes, which represents the main part of the work, leaving the discussion of period spacings to Section 4.4.

The analysis of p modes, performed for each star, was done in three steps: (i) collapse the échelle diagram using the measured $\Delta\nu$; (ii) identify the centroids ν_0, ν_1, ν_2 of the $\ell = 0, 1, 2$ ridges by fitting three Lorentzian profiles to the collapsed échelle diagram, which gives the small spacings $\delta\nu_{02}$ and $\delta\nu_{01}$, and ϵ (see the next paragraph and [165]); and (iii) simulate 500 power spectra by perturbing the observed power spectrum of the star according to a χ^2 statistics with 2-degrees of freedom [319], perform the first two steps of the analysis for each simulation in order to derive a new set of asymptotic parameters, and evaluate their uncertainties by computing a robust rms of the results. Figure 4.3 shows an example of a collapsed échelle diagram obtained with the AARG code. The centroids of the ridges $\ell = 0, 1$, and 2 are marked by dotted lines, while the Lorentzian profiles used to fit the different ridges are shown with thick solid lines (red, blue, and green, respectively).

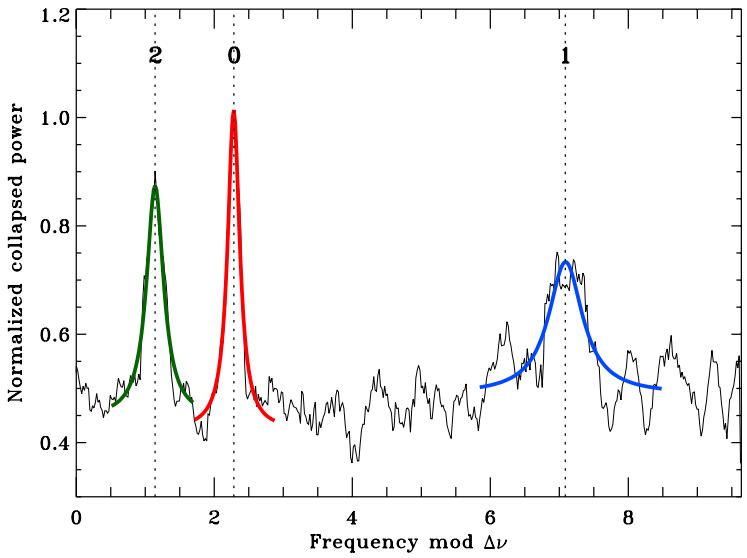

Figure 4.3: *Collapsed échelle diagram of KIC 2436593. The identification of the ridges $\ell = 0, 1$, and 2 is shown, together with their Lorentzian fits (red, blue, and green solid lines, respectively).*

We now describe step (ii) in slightly more detail. AARG requires an initial guess of ϵ for each star, given by manually marking the position of the $\ell = 0$ peak in the collapsed échelle diagram. This is followed by a Lorentzian fit to obtain the centroid, ν_0. According to the values

shown by [165] and [317], we expected to have $0.5 < \epsilon < 1.5$ for red giants with $\Delta\nu < 15$ μHz. We therefore either added or subtracted 1 to the measured ϵ to ensure it would fall within this range (see also [243, 316] for further discussion of the position of the centroids and the ϵ diagrams).

Next, AARG makes a first guess for the centroid ν_2 of the $\ell = 2$ peak by adopting an empirical relation $\nu_0 - \nu_2 \equiv \delta\nu_{02} = c\Delta\nu$, where c is small. Although the relation has a slight mass dependence [236], using a fixed value of $c = 0.123$ offered a reliable first guess for every star in our sample (note that our value is very close to the one measured by [30]). As for ν_0, a Lorentzian fit centered on the first guess for the $\ell = 2$ ridge position provides the final value of the centroid ν_2, and hence also $\delta\nu_{02}$. The search for the $\ell = 1$ peak is performed automatically by finding the maximum in the regions of the collapsed échelle diagram laying outside the $\ell = 0, 2$ peaks. A third Lorentzian fit is then performed, providing the centroid ν_1, which gives $\delta\nu_{01} = \nu_0 + \Delta\nu/2 - \nu_1$ according to the convention by [30]. For a few stars (~ 10) our method did not perform well. This was mainly caused by partly overlapping $\ell = 0, 2$ peaks (especially in NGC 6811 because of the higher mass of its stars) and strongly affected $\ell = 1$ peaks due to mixed modes. These stars were manually analyzed afterwards. We could successfully derive the asymptotic parameters of p modes for a total of 115 stars: 60 for NGC 6791, 5 for NGC 6811, and 50 for NGC 6819.

4.3 Results

In the present section we provide the results arising from our analysis of p modes. A complete table of values can be found in Appendix B.

4.3.1 ϵ Diagram

The ϵ term of Eq. (1.10) was shown to be highly correlated with $\Delta\nu$ for red giant stars [165, 243]. The ϵ diagram is shown in Figure 4.4 for the clusters NGC 6791, NGC 6811, and NGC 6819, where 1σ error bars were derived by means of Eq. (1.10). We note that the RC stars (identified from the CMDs by [292] but adjusted for a few stars based on our analysis of the period spacing, presented in Section 4.4) form distinct groups with slightly lower ϵ than the RGB stars at $\Delta\nu \simeq 3.7\,\mu$Hz for NGC 6791, $\Delta\nu \simeq 8\,\mu$Hz for NGC 6811, and $\Delta\nu \simeq 4.8\,\mu$Hz for NGC 6819. In particular, we measured a weighted average of ϵ for clump stars and RGB stars in the same $\Delta\nu$ range of RC stars, and found them to be $\langle\epsilon^{\mathrm{RC}}\rangle = 0.829 \pm 0.031$ and $\langle\epsilon^{\mathrm{RGB}}\rangle = 0.915 \pm 0.039$ for NGC 6791, and $\langle\epsilon^{\mathrm{RC}}\rangle = 0.970 \pm 0.018$ and $\langle\epsilon^{\mathrm{RGB}}\rangle = 1.015 \pm 0.017$ for NGC 6819. In both cases, $\langle\epsilon^{\mathrm{RC}}\rangle$ appears to be significantly different from $\langle\epsilon^{\mathrm{RGB}}\rangle$. This is in good agreement with [32] and [179]. Although a lower mass of RC stars can result in lower $\Delta\nu$, and hence lower ϵ, one should note that this effect alone cannot explain the observed difference in ϵ between RC and RGB stars of similar $\Delta\nu$ (see also [232] for a detailed study about the mass difference between the RGB and RC stars). The difference in evolutionary state also needs to be taken into account to fully explain the observed difference in phase shift [179].

A least-squares fit to the RGB stars of the clusters was computed, using the log-relation

$$\epsilon = A + B \log \Delta\nu \,, \tag{4.4}$$

adopted by [243]. Since the fits computed to the RGB stars of NGC 6791 and NGC 6819 are not significantly different, we give the result for all the RGB stars in our sample, providing single values for the coefficients A and B. The result is shown as a solid black line in Figure 4.4, where $A = 0.601 \pm 0.025$ and $B = 0.632 \pm 0.032$. The fit from [243], who used a five-month

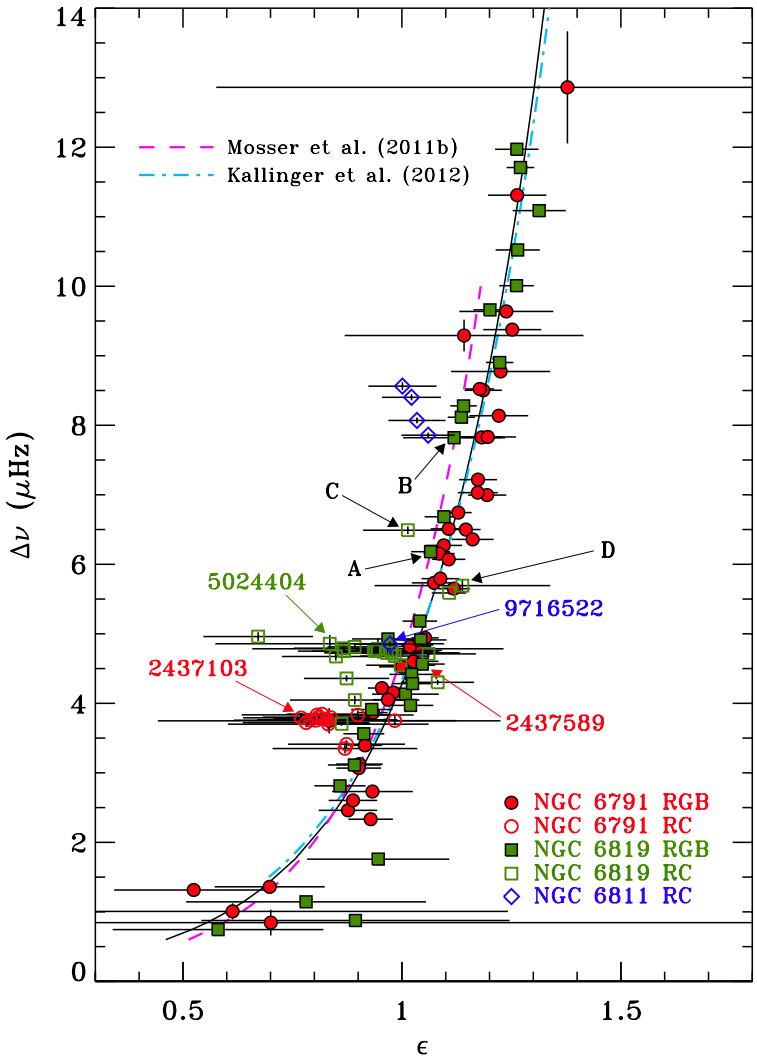

Figure 4.4: *ϵ diagram for the clusters NGC 6791 (red circles), NGC 6811 (blue diamonds), and NGC 6819 (green squares). Open symbols represent RC stars while filled symbols are RGB stars. 1σ uncertainties are displayed for both quantities. A fit to the RGB stars of all the clusters using Eq. (4.4) is added (solid black line), as well as the one from [243] (dashed purple line) and [179] (dot-dashed cyan line). Stars marked with labels and arrows represent special cases that are discussed in Section 4.4.2.*

data set, is added for comparison and plotted in the $\Delta\nu$ range $[0.6, 10]$ μHz covered in that study (dashed purple line). The fit by [179], based on more than 900 field red giants observed by *Kepler* for about 600 days, is almost indistinguishable from ours (dot-dashed cyan line). We also tested a power-law form of the ϵ-$\Delta\nu$ relation and found the χ^2 to be very similar to that derived from Eq. (4.4). The log-relation was finally chosen to allow a direct comparison with the results by [243] and [179].

Lastly, we note that the uncertainties on $\Delta\nu$, and hence on ϵ, become quite large for values of $\Delta\nu$ below $< 2\,\mu$Hz due to the limited frequency resolution and small number of orders observed. For the star with the highest $\Delta\nu$, the large uncertainty is caused by the low S/N level, due to its low oscillation amplitude.

4.3.2 C-D Diagrams

In the C-D diagram one plots the small spacing $\delta\nu_{02}$ versus the large spacing $\Delta\nu$ [75], which for MS stars enables one to discriminate stars of different age and mass. A new version of the C-D diagram proposed by [222] and by [236] for MS and RGB stars respectively, is constructed by considering $\delta\nu_{01}$ instead of $\delta\nu_{02}$ (see also [316]). It has been shown that for red giants the C-D diagrams can not be used to investigate age [317] but that it is still useful to discriminate mass (e.g. see [30, 165, 236, 179]). The results for both $\delta\nu_{02}$ and $\delta\nu_{01}$ are shown in Figures 4.5(a) and (c), for the three clusters. As before, open symbols represent RC stars while filled symbols are RGB stars. As a first approximation, we represent the relation between $\delta\nu_{02}$ and $\Delta\nu$ by the linear relation $\delta\nu_{02} = a_{02} + b_{02}\Delta\nu$, which we fitted with a standard least-squares method to the RGB stars. The results are shown in Fig. 4.5 with a dashed red line for NGC 6791 and a dot-dashed green line for NGC 6819. Their equations are given by

$$\delta\nu_{02}^{(6791)} = (0.121 \pm 0.003)\,\Delta\nu + (0.035 \pm 0.012)\ \mu\text{Hz}\,, \tag{4.5}$$

and

$$\delta\nu_{02}^{(6819)} = (0.114 \pm 0.003)\,\Delta\nu + (0.019 \pm 0.012)\ \mu\text{Hz}\,. \tag{4.6}$$

Only error bars on $\delta\nu_{02}$ were considered for the fits, but the results obtained by including uncertainties on both quantities were indistinguishable from the ones presented here. The coefficients of the $\delta\nu_{02}$-$\Delta\nu$ relation estimated from our fit agree within a few percent with those derived by [179] for field stars.

The typical mass for an RGB star, $M_{\rm RGB}$, is expected to be different for each cluster but about the same within a given cluster. For cluster RGB stars we therefore have a much tighter constraint on the stellar mass than for field stars. [17] found $M_{6791} = 1.20 \pm 0.01\,M_\odot$ and $M_{6819} = 1.68 \pm 0.03\,M_\odot$ as the averages for the RGB stars, which were derived using grids of stellar models that incorporated scaling relations for $\nu_{\rm max}$ and $\Delta\nu$. We refer to [232] for further discussion about the mass estimates for these stars. Following the theoretical work by [236], who showed that $\delta\nu_{02}$ depends on mass, we relate the difference in the slopes, b_{02}, in Eqs. (4.5) and (4.6) to the difference in $M_{\rm RGB}$. Assuming the linear relation

$$b_{02} = \alpha_{02} + \beta_{02}\left(\frac{M_{\rm RGB}}{M_\odot}\right)\,, \tag{4.7}$$

we obtain $\alpha_{02} = 0.138 \pm 0.012$ and $\beta_{02} = -0.014 \pm 0.008$ by solving the system of two equations (one for each cluster).

As done for $\delta\nu_{02}$, we fitted a linear relation $\delta\nu_{01} = a_{01} + b_{01}\Delta\nu$ to the RGB stars of NGC 6791 and NGC 6819. The linear trends are shown in Figure 4.5(c) with the same notation

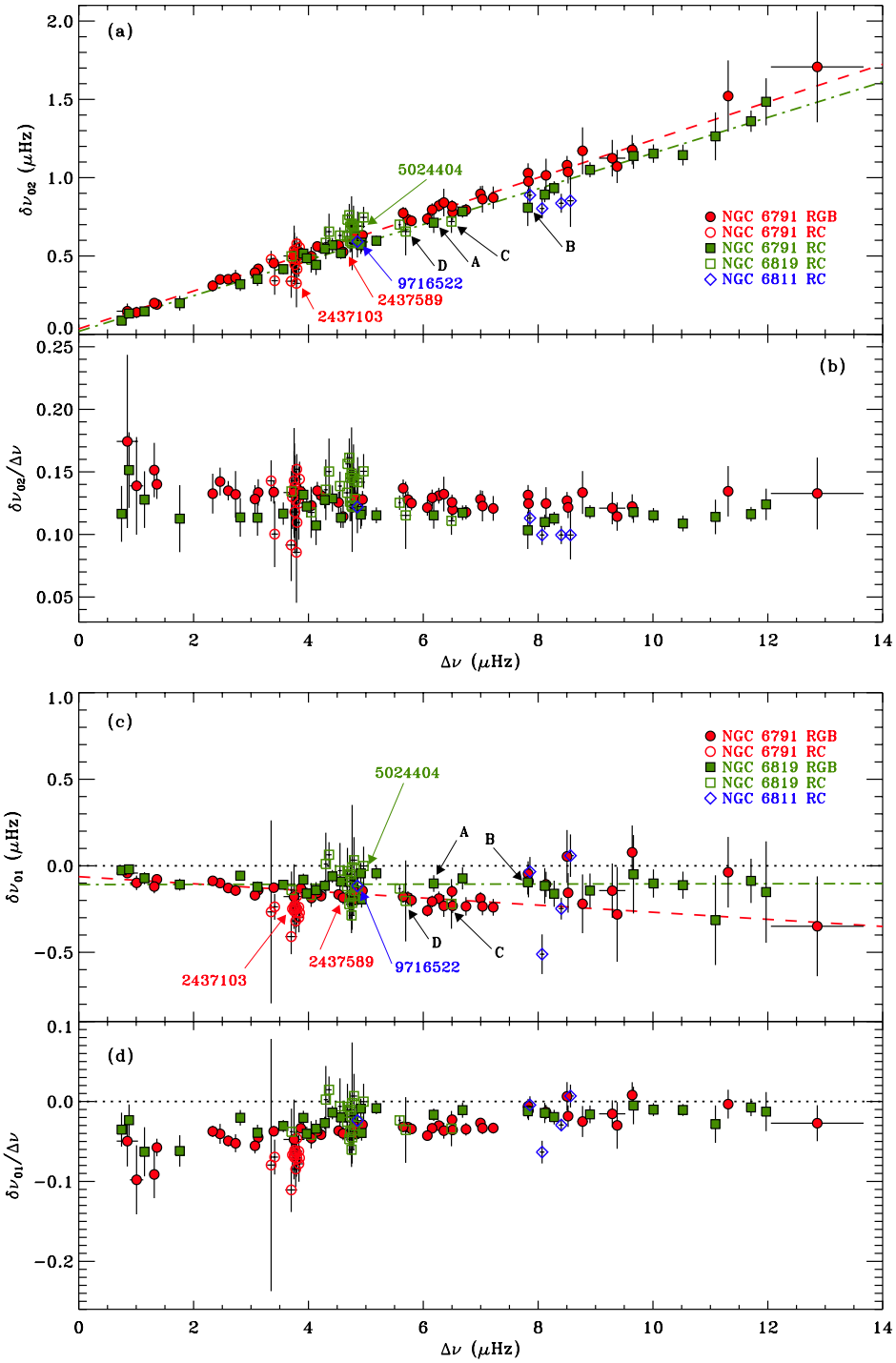

Figure 4.5: ***(a)***, ***(c)***: *C-D diagrams of the small spacings $\delta\nu_{02}$ and $\delta\nu_{01}$ for the clusters NGC 6791 (red circles), NGC 6811 (blue diamonds), and NGC 6819 (green squares). Open symbols represent RC stars while filled symbols are RGB stars. Error bars show 1σ uncertainties. The linear fits to the RGB stars are shown for both NGC 6791 (dashed red line) and NGC 6819 (dot-dashed green line). Stars marked with labels represent special cases that are discussed in Section 4.4.2.* ***(b)***, ***(d)***: *modified C-D diagrams of the ratios $\delta\nu_{02}/\Delta\nu$ and $\delta\nu_{01}/\Delta\nu$ with the same notation adopted for the upper panels.*

as Figure 4.5(a), and the results are

$$\delta\nu_{01}^{(6791)} = -\left(0.021 \pm 0.003\right)\Delta\nu - \left(0.063 \pm 0.011\right)\ \mu\text{Hz} \tag{4.8}$$

and

$$\delta\nu_{01}^{(6819)} = \left(0.000 \pm 0.003\right)\Delta\nu - \left(0.109 \pm 0.012\right)\ \mu\text{Hz}\,. \tag{4.9}$$

Once again, the uncertainties are quite similar for the two clusters. Like $\delta\nu_{02}$, we also see a mass dependence on $\delta\nu_{01}$ for the RGB stars. But unlike $\delta\nu_{02}$, the trend appears to go in the opposite direction, with higher $\delta\nu_{01}$ for higher masses. This is in qualitative agreement with the theoretical results by [236], whose Figure 5(b) shows a slight increase in $\delta\nu_{01}$ for increasing mass along the RGB. As for $\delta\nu_{02}$, we relate the slopes b_{01} in Eqs. (4.8) and (4.9) to M_{RGB}, assuming the linear relation

$$b_{01} = \alpha_{01} + \beta_{01}\left(\frac{M_{\text{RGB}}}{M_\odot}\right)\,, \tag{4.10}$$

and find $\alpha_{01} = -0.073 \pm 0.012$ and $\beta_{01} = 0.044 \pm 0.008$. We find that $|\beta_{01}| \simeq 3|\beta_{02}|$, hence it appears that $\delta\nu_{01}$ is more sensitive to mass than $\delta\nu_{02}$ by about a factor of three. But at this stage we would caution overinterpretation of this result as further theoretical investigations are required to fully understand how $\delta\nu_{01}$ depends on the fundamental parameters and internal structure of red giants (Section 4.3.3).

Figures 4.5(b) and (d) show the so-called modified C-D diagrams which plot the relative ratios $\delta\nu_{02}/\Delta\nu$ and $\delta\nu_{01}/\Delta\nu$. The reason for considering the ratio $\delta\nu_{0\ell}/\Delta\nu$ is that models show it to be less sensitive to surface layer effects (e.g. see [316]) and that the small spacings $\delta\nu_{0\ell}$ approximatively scale with $\Delta\nu$. Our results appear to be in agreement with previous results on red giants [30, 165, 243, 179] and with the theoretical studies by [236].

4.3.3 The small spacings of Red Clump Stars

It is interesting to compare the average small spacings for the RC stars relative to the RGB stars in each cluster. In the following we denote this quantity $\Delta\langle\delta\nu_{0\ell}\rangle \equiv \langle\delta\nu_{0\ell}^{\text{RC}}\rangle - \langle\delta\nu_{0\ell}^{\text{RGB}}\rangle$. It is evident in all four panels of Figure 4.5, but slightly more so in Figures 4.5(b) and (d), that the RC stars on average show different small spacings than RGB stars of similar $\Delta\nu$. We will first discuss $\Delta\langle\delta\nu_{02}\rangle$.

For NGC 6819, $\Delta\langle\delta\nu_{02}\rangle = 0.112 \pm 0.016\,\mu$Hz, while for NGC 6791 we have $\Delta\langle\delta\nu_{02}\rangle = 0.012 \pm 0.021\,\mu$Hz. Given the relation between $\delta\nu_{02}$ and mass for RGB stars (Section 4.3.2), one might speculate that a similar relation would exist for RC stars. However, we note that $\Delta\langle\delta\nu_{02}\rangle$ for NGC 6819 is about twice as large as the difference in $\delta\nu_{02}$ between the two RGB populations in NGC 6791 and NGC 6819 at a similar $\Delta\nu$. Hence, if we applied the relation in Eq. (4.7) to the RC stars, we would find that the RC stars in NGC 6819 have a mass of about $0.7\,M_\odot$ (corresponding to a mass loss of about $1\,M_\odot$), in stark disagreement with the results by [232], who found $\Delta\langle M\rangle = -0.03 \pm 0.04$. Hence, there is certainly something else dominating the different values of $\Delta\langle\delta\nu_{02}\rangle$ we see for the two clusters.

Turning our attention to the other small spacing, we have $\Delta\langle\delta\nu_{01}\rangle = 0.004 \pm 0.025\,\mu$Hz for NGC 6819 and $\Delta\langle\delta\nu_{01}\rangle = -0.113 \pm 0.020\,\mu$Hz for NGC 6791. Hence we see that $\Delta\langle\delta\nu_{01}\rangle_{6791} < \Delta\langle\delta\nu_{01}\rangle_{6819}$, which was also the case for $\Delta\langle\delta\nu_{02}\rangle$. We recall that the mass dependencies of $\delta\nu_{02}$ and $\delta\nu_{01}$ were opposite for the RGB, both in observations (Section 4.3.2) and models [236]. Hence, we would also expect an opposite trend for $\Delta\langle\delta\nu_{01}\rangle$ ($\Delta\langle\delta\nu_{01}\rangle_{6791} > \Delta\langle\delta\nu_{01}\rangle_{6819}$). The fact that we do not observe this is further evidence that a simple relation with mass alone cannot explain the observed differences in small spacings between RC and RGB stars.

A possible explanation is the internal structural changes of the stars that occur during the He-flash phase [42] between the tip of the RGB and the RC. These changes could be significantly different for stars of different masses ($M_{6791} = 1.20 \pm 0.01\,M_\odot$ and $M_{6819} = 1.68 \pm 0.03\,M_\odot$, [17]), composition ($[\mathrm{Fe/H}]_{6791} = 0.29 \pm 0.03$ (random) ± 0.07 (systematic), [55]; $[\mathrm{Fe/H}]_{6819} = 0.09 \pm 0.03$, [51]), and rotation rates [223]. Further investigation requires modeling of both the RGB and RC stars in these clusters.

The dependence of $\delta\nu_{01}$ on stellar properties was investigated by [236] using stellar models covering $0.7 - 2.3\,M_\odot$ on the RGB and $2.5 - 5.0\,M_\odot$ in the He-core burning phase. They found that small values of $\delta\nu_{01}$ were predominantly seen among RGB models, and we would therefore expect the RC stars to show larger $\delta\nu_{01}$ on average, which is in contrary to what we observe for NGC 6791. However, we note that all the He-core burning models in the [236] sample were more massive than the stars in the two clusters considered here, and a direct comparison is therefore not possible. The physical cause of a lower value of $\delta\nu_{01}$ was not firmly established by [236], but they argued that there was a tendency for low $\delta\nu_{01}$ values in models where the inner turning point of the $\ell = 1$ modes was well inside the convective envelope, corresponding to stars with deep convective envelopes. Clearly, these issues deserve further study.

4.3.4 Ensemble échelle Diagrams

Following [165], we computed the so-called ensemble échelle diagrams for both NGC 6791 and NGC 6819. When dealing with a large number of stars, ensemble échelle diagrams are very helpful for studying the evolution of features such as ridge width and position. In particular, the measurement of the average position of the $\ell = 3$ ridge becomes possible also when low S/N in the power spectra does not allow one to make a clear detection of the corresponding peak in a single star. The results are shown in Figures 4.6(a) and (c), where the stars are numbered by increasing $\Delta\nu$. Each row in the plot represents the collapsed échelle diagram of a single star using the large separation adopted in the analysis, and shifted in order to have the $\ell = 0$ ridge fall on $(\nu/\Delta\nu \bmod 1) = 0.3$ (see also the discussion by [293], Sec. 2). The RC stars (red star symbol) clearly show strong broadening of the $\ell = 1$ and 2 ridges. We see that even the $\ell = 0$ ridge appears broader for RC stars in both clusters.

The result of collapsing the ensemble échelle over the entire sample of stars is shown in Figures 4.6(b) and (d) (thick black line). Results for RC stars (red line) and RGB stars (blue line) are also plotted for both clusters. The presence of an $\ell = 3$ peak becomes evident for NGC 6791, while for NGC 6819 a hint of $\ell = 3$ is visible only for the RGB stars. For NGC 6791, the $\ell = 3$ hump seems to arise from several stars, particularly those with $\Delta\nu < 7\,\mu$Hz, as visible from Figure 4.6(a). The position of the marked $\ell = 3$ peaks of the two clusters, and hence their average small spacings $\delta\nu_{03}$, are in agreement with the results of [30, 165, 243] and [179].

It is noticeable that the $\ell = 1, 2$ ridges move away from the $\ell = 0$ ridge as the stars evolve from H-shell to He-core burning red giants, a result that was already discussed by [165]. We also note that the hump visible in Figure 4.6(d), on the left slope of the $\ell = 1$ peak (red line), is caused by only two stars having strong peaks that occur at $(\nu/\Delta\nu \bmod 1) \simeq 0.7$ and this is therefore not an indication of a general feature. Referring to the effect on $\delta\nu_{01}$ discussed in Section 4.3.3, we notice that the $\ell = 1$ ridge of the RC stars of NGC 6791 (Figure 4.6(a)) is shifted towards the right-hand side of the diagram, i.e towards lower values of the small spacing, while this shift is not apparent in NGC 6819 (Figure 4.6(c)).

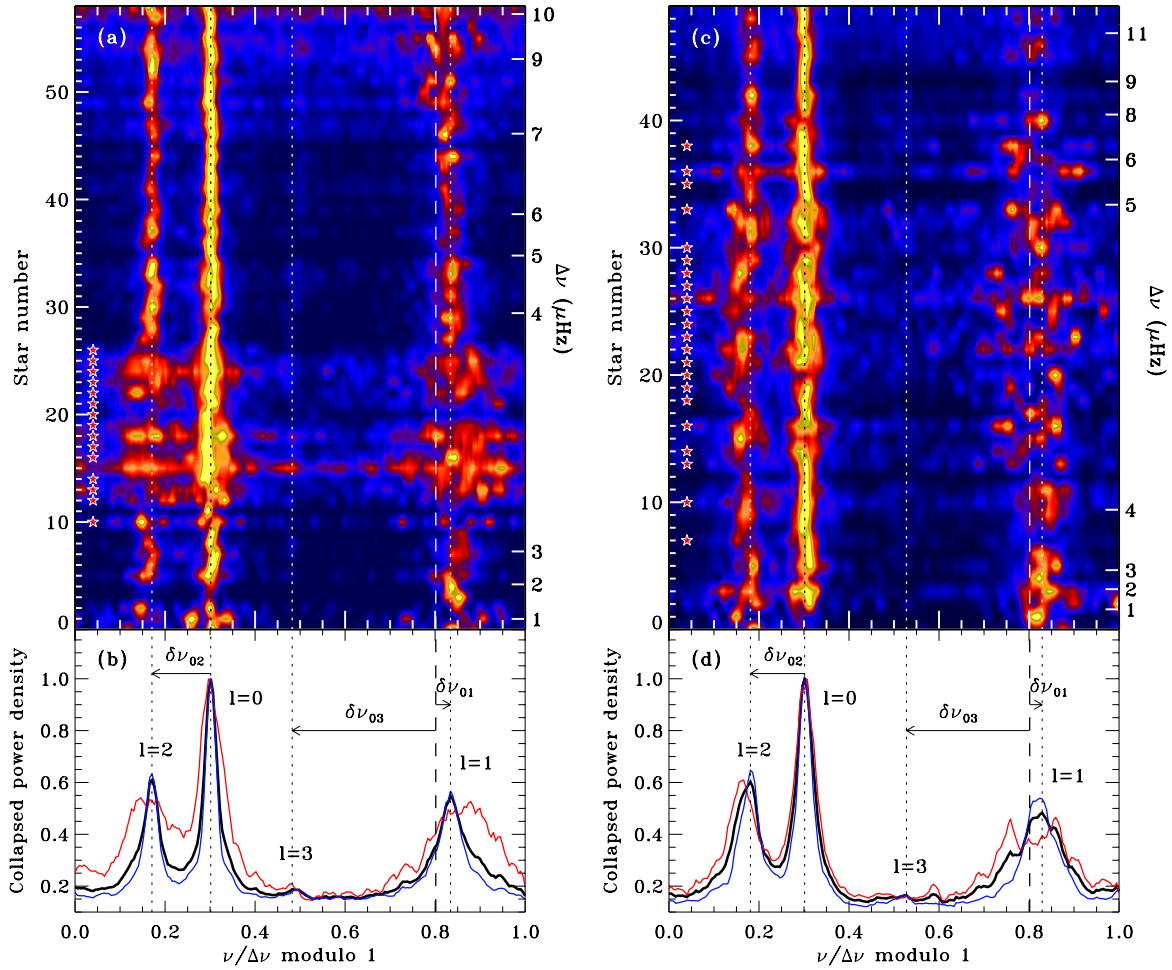

Figure 4.6: ***(a)***, ***(c)***: *ensemble échelle diagrams of the clusters NGC 6791 and NGC 6819, respectively, where* $\ell = 0$ *centroids were aligned by shifting the* $\ell = 0$ *ridge of each star to align with* $(\nu/\Delta\nu \mod 1) = 0.3$. *The number of the stars, ordered by increasing* $\Delta\nu$, *is shown on the left coordinate, and the corresponding* $\Delta\nu$ *is shown on the right axis. Red star symbols mark the clump stars identified in the clusters. Note that each row corresponds to the collapsed échelle of one star, normalized to unity.* ***(b)***, ***(d)***: *diagrams showing panels (a), (c) collapsed over the entire range of* $\Delta\nu$ *(thick black line) normalized to unity. Results for RC stars in red and RGB stars in blue are also shown. Ridge identifications and definitions of small separations are indicated. In both panels, the dotted lines represent the centroids of the* $\ell = 0, 1, 2$ *and* 3 *ridges, while the dashed line is the position of the midpoint of two adjacent* $\ell = 0$ *modes.*

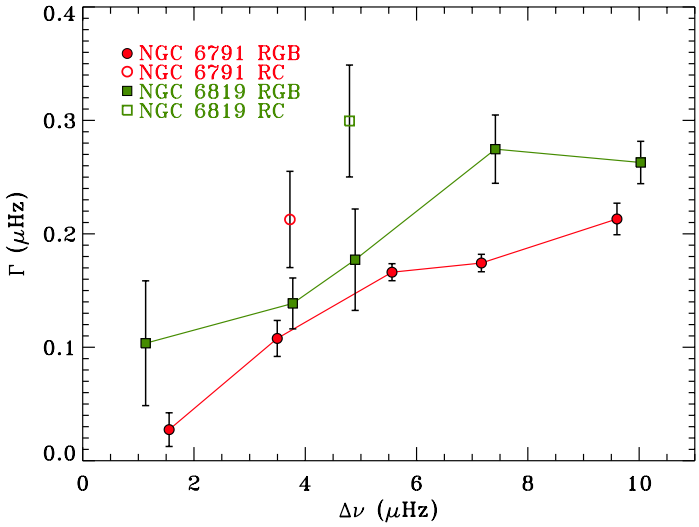

Figure 4.7: *FWHM of the $\ell = 0$ ridge as a function of $\Delta\nu$ for RGB stars in NGC 6791 (filled red circles) and NGC 6819 (filled green squares). Each point represents the average of values within a subset of stars with similar $\Delta\nu$. The error bars are the 1σ uncertainties on the mean for each subset. Open symbols at $\Delta\nu \simeq 3.7\,\mu Hz$ and $\Delta\nu \simeq 4.8\,\mu Hz$ represent the measurements for the subsets of RC stars.*

4.3.5 Mode Linewidths

Measuring the linewidths of p modes and studying how they correlate to the fundamental stellar properties has important consequences for the understanding of the damped nature of solar-like oscillations. In fact, the physics responsible for the damping mechanism that acts in the convective envelope of low mass stars is not yet fully understood (e.g. see [158, 103, 35]).

In the present work we provide estimates of the linewidths of radial modes derived through the AARG code. In particular, the widths of the ridges in the collapsed échelle diagrams (Figures 4.6(b) and (d)) give a rough estimate of the mode linewidths. Figure 4.7 shows the FWHM for the $\ell = 0$ ridge from the Lorentzian fit to the corresponding peak in the collapsed échelle diagram, for the RGB stars of both NGC 6791 and NGC 6819. Each point is the average from a subset of stars sorted in bins of $\Delta\nu$, while the overlaid error bars are 1σ uncertainties on the mean for each bin. The open symbols at $\Delta\nu \simeq 3.7\,\mu$Hz and $\Delta\nu \simeq 4.8\,\mu$Hz show our measurements for the RC stars. We see a clear increasing trend when moving to higher $\Delta\nu$, a result that was already apparent from the analysis by [165] of field red giants, despite of the shorter data set that was available (see also [179], who obtained a similar result by using a different method). This increasing trend is also visible in Figures 4.6(a) and (c), in that the scaled width $\nu/\Delta\nu$ of the $\ell = 0$ ridge at low $\Delta\nu$ is about the same of that at high $\Delta\nu$. In Figure 4.7 we also notice a systematic difference between the ridge widths of the two clusters, a feature that is already visible from the collapsed échelle diagrams of Figures 4.6(b) and (d).

To see whether the difference in ridge width between the two clusters and between stars with different $\Delta\nu$ arises from the difference in temperature of the stars as contemplated by [69, 19, 7, 35], we plot our measurements of FWHM as a function of $T_{\rm eff}$ in a log-log scale in Figure 4.8. This shows indeed that all the cluster stars follow an almost common trend, which supports that the observed difference in ridge width largely follows the difference in temperature. We also show the linewidth measurements of a sample of main sequence (MS) and subgiant stars (blue diamonds) from [7], where we have taken temperatures from [65, 66]. It is remarkable how well all the stars are aligned in Figure 4.8. Note that our measure of the

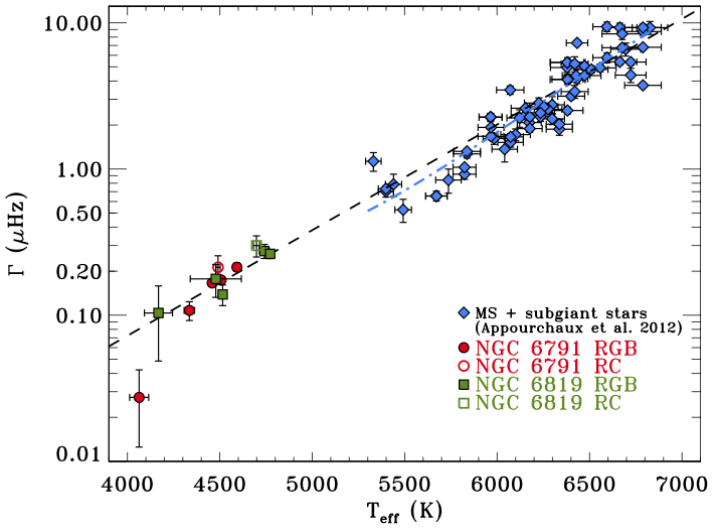

Figure 4.8: *FWHM of the $\ell = 0$ ridge plotted against $T_{\rm eff}$ for the stars of NGC 6791 (red circles) and NGC 6819 (green squares). Also shown are measured linewidths for MS and subgiant field stars (blue diamonds) from [7]. Each cluster point represents the same subset of stars plotted in Figure 4.7. The error bars are the 1σ uncertainties on the mean for each subset. The fit to the MS and subgiant stars taken from [7] is also shown (dot-dashed blue line). The dashed black line shows an exponential fit (Eq. (4.11)) to all stars.*

ridge width only provides an upper limit to the 'true' mode linewidths because of the slight curvature of the ridges in the échelle diagram. The fit to the linewidths across all stars is represented by an exponential function

$$\Gamma = \Gamma_0 \, exp\left(\frac{T_{\rm eff} - 5777\,{\rm K}}{T_0}\right), \tag{4.11}$$

where $\Gamma_0 = 1.39 \pm 0.10\,\mu$Hz and $T_0 = 601 \pm 3\,$K (dashed black line). A detailed study using linewidths found by direct mode fitting (peak bagging) of MS and red giant stars (e.g. [69, 19, 7]) goes beyond the scope of this work. The power law fit with a background component proposed by [7] is here added for comparison in its range of validity (5300 K−6800 K, dot-dashed blue line). However, we can conclude that our measurements, combined with *Kepler* results on MS and subgiant stars, follow a single exponential trend with temperature.

4.4 Mixed Modes

Mixed modes have the great advantage of being sensitive to the core structure, while at the same time being observable at the surface. They were recently used as a way to successfully distinguish between RC and RGB stars [32, 242]. Although their amplitude is lower than of pure p modes [244], long datasets enable us to identify many of them due to their long lifetimes [103]. Even in cluster red giants, which are generally fainter than the *Kepler* field stars, we can detect many mixed $\ell = 1$ modes in the best cases. The main features of mixed modes relevant for the analysis presented in this work are discussed in Section 4.4.1, while our results on their period spacings are described in Section 4.4.2.

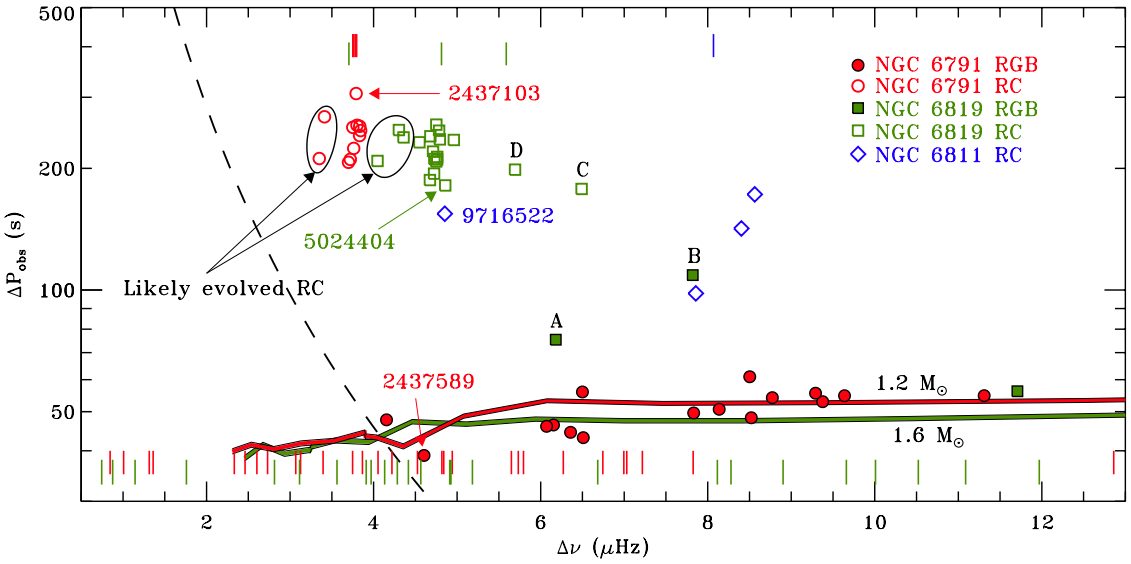

Figure 4.9: *Period spacings of the three clusters NGC 6791 (red circles), NGC 6811 (blue diamonds), and NGC 6819 (green squares). Open symbols represent RC stars, while filled symbols are RGB stars. Tracks for* 1.2 $M_\odot$ *(thick red line) and* 1.6 $M_\odot$ *(thick green line), and* $Z = 0.017$*, are shown [316]. Tick marks drawn at the top (RC) and bottom (RGB) refer to stars that could not be identified with our period spacing analysis, and are colored according to the notation adopted in the rest of the paper. The dashed black line represents the minimum period spacing one can measure with a 19 months-long time-series. Special cases discussed in Section 4.4.2 and listed in Table 4.1 are marked.*

4.4.1 Period Spacings of Mixed Dipole Modes

While p modes are equally spaced in frequency, pure g modes are approximately equally spaced in period, following the asymptotic relation given by Eq. (1.12) [299, 78]. From Eq. (1.12), the period spacing of dipole g modes is given by

$$\Delta P_{\rm g} = \Pi_0/\sqrt{2}\,, \tag{4.12}$$

which appears the most interesting quantity to investigate because of the strong coupling between p and g modes for $\ell = 1$ [103, 236, 293, 24].

However, in contrast to the large separation for p modes, the period spacing of pure g modes, $\Delta P_{\rm g}$, cannot always be directly measured in red giants because all the non-radial modes are mixed in the red giant phase [78]. Nevertheless, from recent studies it seems to be possible to infer Π_0 in some cases (see [32, 245]). Fortunately we can readily measure the period spacing of the mixed modes, $\Delta P_{\rm obs}$, which can serve as a proxy for $\Delta P_{\rm g}$. $\Delta P_{\rm obs}$ is lower than $\Delta P_{\rm g}$ by about a factor of 0.6-0.8 (e.g. see [32, 245]).

4.4.2 $\Delta P_{\rm obs}$-$\Delta\nu$ Diagram

As mentioned in the introduction of the present chapter and in 4.2, AARG measures period spacings using the approach used by [32] (see Appendix A for further details). As a first step, it modifies the power spectrum for each star by erasing the regions containing all the $\ell = 0, 2$ modes, whose positions come directly from our analysis of p modes (see Section 6.1). This new power spectrum shows only $\ell = 1$ mixed modes, and possibly some low amplitude $\ell = 3$ modes, and is then expressed in period rather than frequency. The power spectrum of this

power spectrum is then calculated, which is converted back into period. To obtain a first guess for the observed period spacing, a manual marking of the position of the excess of power is required. Finally, a Gaussian fit to the selected hump of power provides our measure of $\Delta P_{\rm obs}$.

We were able to measure $\Delta P_{\rm obs}$ for a total of 53 stars in our sample: 27 from NGC 6791, 4 from NGC 6811, and 22 from NGC 6819. The results are shown in Figure 4.9, where RC and RGB stars are marked with open and filled symbols, respectively. The dashed black line is the limit set by the frequency resolution. Overlaid are theoretical tracks for $1.2\,M_{\odot}$ and $1.6\,M_{\odot}$ at near-solar metallicity (Z = 0.017), as calculated by [316], which are representative of the RGB stars of NGC 6791 and NGC 6819, respectively. Using stellar models, we verified that changing the metallicity over the range spanned by the two clusters has no significant effect on $\Delta P_{\rm g}$ for RGB stars.

The tick marks at the top (RC) and bottom (RGB) represent stars for which the period spacing could not be clearly measured by our analysis, classified by [292] using the CMD. We see that the fraction of stars with measured period spacings is much higher for RC stars than for RGB stars even after taking into account the limit set by the frequency resolution. In particular, for NGC 6791 these fractions are ∼36% (RGB) and ∼73% (RC), while for NGC 6819 they are ∼10% (RGB) and ∼86% (RC). This could be explained by a weaker coupling between the p-mode and g-mode cavities for the RGB stars [103], which makes the resonances narrower in frequency, resulting in a smaller number of observable mixed modes.

4.4.3 Discussion of Special Cases

The stars labeled from A to D (KIC 5112361, KIC 4937770, KIC 5024414 and KIC 5024476) are outliers in the $\Delta P_{\rm obs}$-$\Delta\nu$ diagram (Figure 4.9), while KIC 2437103, KIC 2437589, and KIC 5024404 have period spacings that imply a different stage of evolution to the one based on the CMDs [292]. KIC 9716522 represents a star on its way up towards the AGB, as already noted by [292] and now supported by our measurement of its high period spacing in agreement with that of other He-burning stars (Figure 4.9).

All these highlighted stars are also marked in Figures 4.4 and 4.5, and in the CMDs of Figure 4.10 (as derived by [292]). We also list these stars in Table 4.1, together with all their asteroseismic parameters derived in this work. To further support the discussion presented below, we derived the masses of all stars near the RC in the CMD, including the outliers A−D and the stars that we have marked as 'likely evolved RC' (Figure 4.9), which are also shown in Figure 4.10 and listed in Table 4.1. To estimate the masses we use the scaling relation for the asteroseismic mass of a star, introduced in Chapter 3 and given by Eq. (3.3), where we adopted $\nu_{\rm max,\odot} = 3100\,\mu$Hz, $\Delta\nu_{\odot} = 135\,\mu$Hz and $T_{\rm eff,\odot} = 5777\,$K (e.g. [232]). The result is shown in Figure 4.11, with masses plotted against V magnitude and 1σ error bars overlaid (see also Chapter 3). Blue lines represent the mean masses of RC stars (solid) and their 1σ uncertainties (dashed), as derived by [232] by adopting Eq. (3.3). To provide corrected estimates of mass for clump stars, the $\Delta\nu$ scaling relation was corrected by 2.7 % and 1.9 % for NGC 6791 and NGC 6819, respectively, according to the study by [232].

We first discuss possible causes for the outliers, A−D. All four are potentially binary stars. Three of them (A, C, and D) are listed as binary stars in the radial velocity study by [154], and the fourth star (B) shows a low oscillation amplitude, which could be indicative of a binary star, as argued by [291]. All four stars also appear relatively blue in the CMD (Figure 4.10). Stars A and B fall below the RC in the CMD and are in line with the rest of the RGB stars in Figures 4.4 and 4.5, suggesting that they are RGB stars with no clear sign of an abnormal mass (Figure 4.5(a) and (b)). This is confirmed by our estimate of their masses according to Eq. (3.3) (Figure 4.11), whose values are similar to the average mass of the RGB stars of

NGC 6819 found by [17] (see Section 4.3.2). Binarity seems like the most plausible explanation for their $B-V$ colors being lower than the other RGB stars. However, their power spectra do not show oscillations from two components, and their higher-than-expected $\Delta P_{\rm obs}$ is therefore difficult to explain. Perhaps it could come from a different core structure of these stars caused by binary interaction. We note that the stars do not seem to be the result of a merger event, given their apparently 'normal' masses. In conclusion, stars A and B are most likely both binaries, with one component on the RGB, whose seismic signal we detect, and a fainter less-evolved component. Stars C and D have luminosities typical to that of the RC. Our measurement of $\Delta P_{\rm obs}$ suggests that the stars indeed belong to the RC. The position of the stars in the sequence of He-core burning stars going from low mass (low $\Delta\nu$) to high mass (high $\Delta\nu$) spanned by the three clusters indicates that stars C and D have higher masses than the other RC stars in NGC 6819 (Figure 4.9). This is confirmed by our estimate of their masses (Figure 4.11). Our conclusion that they are high-mass RC stars is in good agreement with [268], who mention these stars along with others with this position in the CMD to be potential descendants of blue stragglers, meaning that they experienced mass transfer and therefore have a component with a mass significantly higher than the cluster's turn-off mass.

Concerning the next three stars, our period spacing analysis shows that KIC 2437103 ($\Delta P_{\rm obs} = 306\,$s) is an RC star, and KIC 2437589 ($\Delta P_{\rm obs} = 39\,$s) is an RGB star, as argued by [232], and KIC 5024404 ($\Delta P_{\rm obs} = 182\,$s) is an RC star. It seems that KIC 2437589 is an evolved blue straggler in the RGB phase, as suggested by [56]. This would explain its unusual position in the CMD (top panel of Figure 4.10), and is also supported by a mass of about $1.7\,M_{\odot}$, as derived from Eq. (3.3), greater than the mass of the other RGB stars of the cluster.

Lastly, six stars (two in NGC 6791 and four in NGC 6819), were found to be possible candidates for RC stars that are starting to evolve towards the AGB. We list them as 'likely evolved RC' in Table 4.1. Our suggestion arises from our measurement of their $\Delta P_{\rm obs}$, which corresponds to that of RC stars, and from their $\Delta\nu$, which is lower than that of the other RC stars. Their masses (Figure 4.11) are similar to that of the average RC star which, in combination with their lower $\Delta\nu$, confirm that they have a radius significantly larger than the other RC stars.

4.5 Summary of the Results

To summarize and conclude on the main results of the analysis presented in this chapter:

1. The fit of the ϵ-$\Delta\nu$ relation to the RGB stars of our sample, computed using Eq. (4.4), is compatible with the result derived by [243], although it deviates slightly towards higher values of $\Delta\nu$, where our sample has more stars and benefits from longer observations. Our fit is almost indistinguishable from that by [179], which was based on more than 900 field red giants observed for a similar length of time. Moreover, we tested a power-law form of the ϵ-$\Delta\nu$ relation and found it to provide a very similar fit to that derived from Eq. (4.4). Lastly, the average ϵ of clump stars appears to be significantly different from that of their RGB counterparts for both NGC 6791 and NGC 6819, a result in agreement with previous findings on field RGs.

2. The linear fits to the $\delta\nu_{02}$-$\Delta\nu$ relation for the RGB stars of our sample, given by Eqs. (4.5) and (4.6), appear to be compatible within a few percent with the results by [165] and [179] on field red giants. A direct measure of the mass-dependence for the small spacings $\delta\nu_{02}$ and $\delta\nu_{01}$ is derived for the first time for cluster stars. The result indicates that $\delta\nu_{01}$

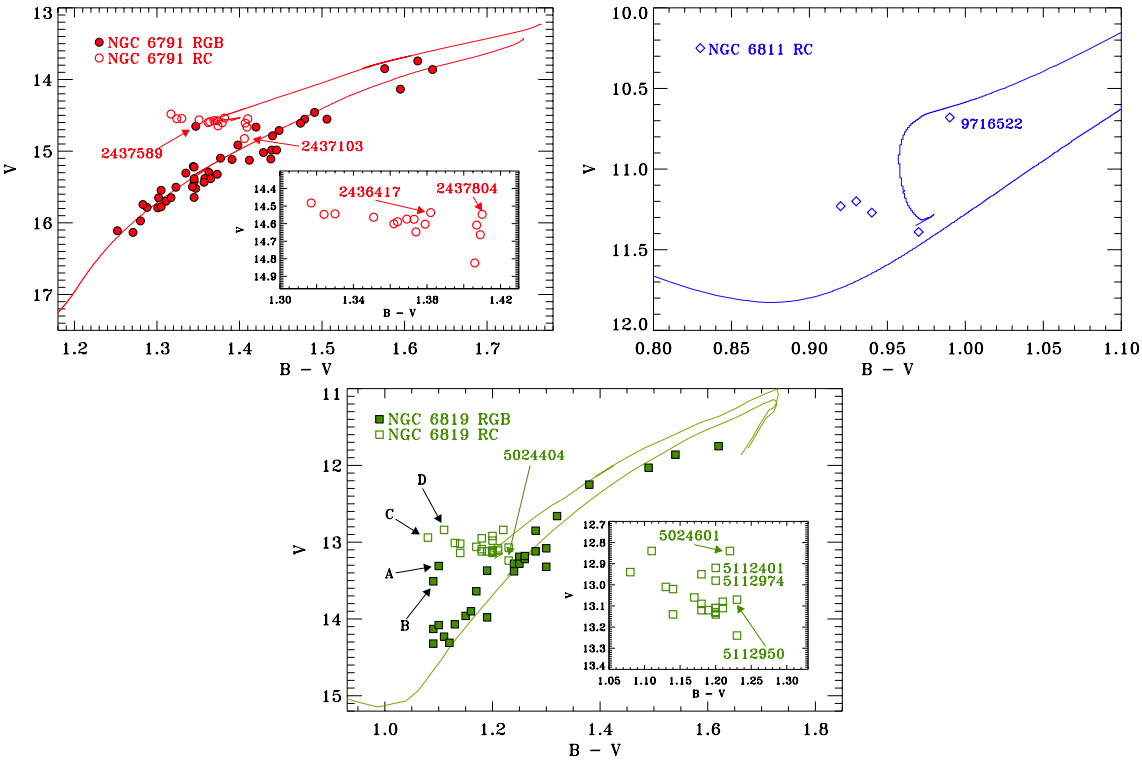

Figure 4.10: *CMDs of the clusters NGC 6791 (top panel), NGC 6811 (middle panel), and NGC 6819 (bottom panel) as derived by [292]. Both RC and RGB stars are shown, with open and filled symbols respectively, according to the classification obtained by the membership study of [292] and our analysis of period spacings. Stars marked with labels represent special stars discussed in Section 4.4.2 and listed in Table 4.1. Isochrones are shown for all the clusters (solid lines, see [292], for details).*

is about three times more sensitive to a mass difference than $\delta\nu_{02}$. The mass-dependence for $\delta\nu_{02}$ is compatible with the results by [179] on field red giants. Furthermore, both $\delta\nu_{02}$ and $\delta\nu_{01}$ show dependence on mass that is qualitatively in agreement with theoretical studies of red giant stars by [236].

3. It is notable that the RC stars of NGC 6791 behave differently from those of NGC 6819 for both $\delta\nu_{02}$ and $\delta\nu_{01}$, as visible **in** Figures 4.5 and 4.6. We quantified this unexpected feature through the difference in $\langle\delta\nu_{0\ell}\rangle$ between RC and RGB stars, which is significantly different from one cluster to the other for both the small spacings. As discussed in Section 4.3.3, further theoretical investigations concerning differences on mass, metallicity and rotation between the two cluster populations, are required for the full interpretation of our results.

4. The positions of the ridges in the collapsed ensemble échelle diagrams (Figure 4.6) confirm the results from [165], with the $\ell = 1$ and 2 ridges moving away from the $\ell = 0$ ridge as the stars evolve from the H-shell to the He-core burning phase. The position of the $\ell = 3$ ridges, hence of the average small spacings $\delta\nu_{03}$ (Figures 4.6(b) and (d)), is also in agreement with results by [30, 165, 243, 179] on field red giants. The FWHM of $\ell = 0$ ridge, which represents an upper limit of the mode linewidths, increases with $\Delta\nu$ for both NGC 6791 and NGC 6819, a result that agrees with the studies by [165] and by [179] on field red giants. A systematic difference of the FWHM between the two clusters is

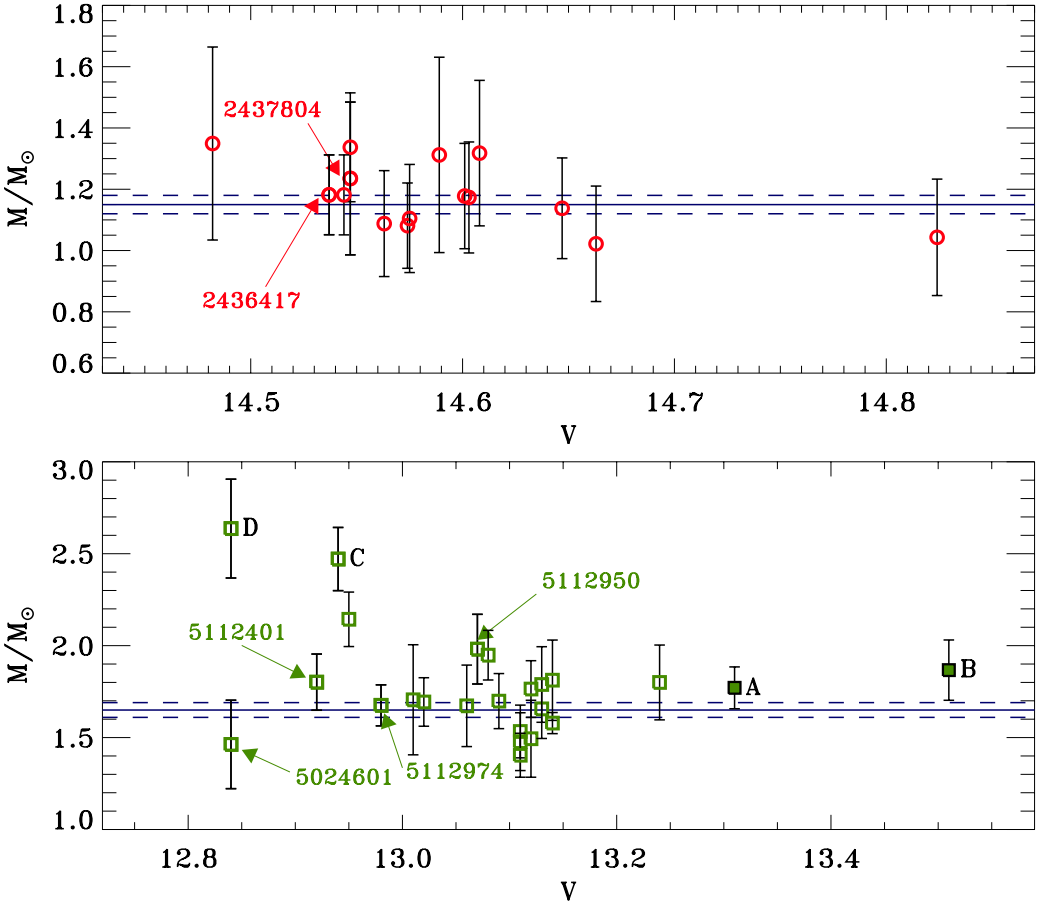

Figure 4.11: *Mass of stars near the RC of NGC 6791 (red circles) and NGC 6819 (green squares) with applied correction in the $\Delta\nu$ scaling of 2.7 % and 1.9 % respectively (see [232]). Error bars show $1-\sigma$ uncertainties derived according to Eq. (3.3). Outlier stars $A-D$ and 'likely evolved RC' stars discussed in Section 4.4.2 and listed in Table 4.1 are marked. Blue lines represent the corrected mean masses of RC stars (solid) and their 1σ uncertainties (dashed) derived by [232].*

shown, which is largely explained by the temperature dependence of mode linewidths (Figure 4.8), as discussed in Section 4.3.5. This result shows the first evidence for an exponential correlation between mode linewidth and temperature in red giants, which is consistent with extrapolating *Kepler* results for main-sequence and subgiant stars derived by [7].

5. The analysis of period spacings using the method described by [32] allowed for the successful identification of almost half of the stars in our sample as either H-shell or He-core burning red giants. The fraction of RGB stars with clearly detectable period spacings is much lower than for RC stars, as seen in Figure 4.9, and could be due to a weaker coupling of the p- and g-mode cavities in RGB stars. We see a number of outliers in Figure 4.9 which require further investigations. It is likely that all of them are binaries and two of them appear to be evolved blue stragglers, as suggested by [268] and supported by their higher masses (Figure 4.11). In addition, our analysis confirms the suggestion by [232] that stars KIC 2437103 ($\Delta P_{\rm obs} = 306\,$s) and KIC 2437589 ($\Delta P_{\rm obs} = 39\,$s), are an RC and an RGB star, respectively, and that KIC 5024404 ($\Delta P_{\rm obs} = 182\,$s) is an RC star.

Lastly, we find a number of possible candidates for evolved RC stars in both NGC 6791 and NGC 6819, as suggested by our measurement of their masses (Figure 4.11), which indicate they have a radius larger than the other RC stars. The special cases discussed in Section 4.4.3 represent potentially interesting targets for detailed theoretical modeling.

Table 4.1: *Overall asteroseismic parameters for some interesting targets.*

KIC ID	NGC	Notes[a]	$\nu_{\max}$ (μHz)	$\Delta\nu$ (μHz)	ϵ	$\delta\nu_{02}$ (μHz)	$\delta\nu_{01}$ (μHz)	$\Delta P_{\rm obs}$ (s)
5112361	6819	(A) Outlier	67.4 ± 1.4	6.181 ± 0.025	1.066 ± 0.044	0.712 ± 0.066	-0.102 ± 0.048	75
4937770	6819	(B) Outlier	93.8 ± 2.4	7.821 ± 0.076	1.119 ± 0.117	0.808 ± 0.117	-0.096 ± 0.070	109
5024414	6819	(C) Outlier	77.1 ± 1.5	6.490 ± 0.056	1.013 ± 0.103	0.720 ± 0.072	-0.220 ± 0.143	178
5024476	6819	(D) Outlier	67.0 ± 1.7	5.693 ± 0.097	1.138 ± 0.201	0.656 ± 0.152	-0.203 ± 0.234	199
2437103	6791	Misclassified CMD	29.7 ± 1.7	3.791 ± 0.064	0.770 ± 0.132	0.325 ± 0.153	-0.242 ± 0.114	306
2437589	6791	Misclassified CMD	46.5 ± 1.5	4.603 ± 0.026	1.026 ± 0.057	0.526 ± 0.042	-0.184 ± 0.038	39
5024404	6819	Misclassified CMD	48.8 ± 0.7	4.857 ± 0.126	0.835 ± 0.261	0.689 ± 0.095	-0.122 ± 0.072	182
9716522	6811	AGB	54.9 ± 1.0	4.852 ± 0.036	0.973 ± 0.084	0.592 ± 0.102	-0.116 ± 0.099	154
2436417	6791	Likely evolved RC	26.7 ± 0.8	3.412 ± 0.058	0.874 ± 0.133	0.342 ± 0.090	-0.237 ± 0.074	268
2437804	6791	Likely evolved RC	26.5 ± 1.6	3.350 ± 0.070	0.870 ± 0.165	0.478 ± 0.054	-0.266 ± 0.529	212
5024601	6819	Likely evolved RC	31.8 ± 1.7	3.704 ± 0.028	0.862 ± 0.065	0.498 ± 0.061	-0.140 ± 0.107	-
5112401	6819	Likely evolved RC	38.2 ± 0.7	4.047 ± 0.068	0.892 ± 0.158	0.476 ± 0.082	-0.169 ± 0.069	209
5112950	6819	Likely evolved RC	42.8 ± 1.3	4.302 ± 0.036	1.082 ± 0.083	0.584 ± 0.104	0.010 ± 0.181	249
5112974	6819	Likely evolved RC	41.7 ± 0.7	4.358 ± 0.045	0.874 ± 0.099	0.655 ± 0.115	0.064 ± 0.073	239

[a] Target description as presented in Section 4.4.2

Part III

Bayesian Methods applied to *Kepler* Targets

Bolstad, W. M., *Introduction to Bayesian Statistics, 2nd Edition*, Wiley & Sons, 2007 [44]
Trotta, R., *Bayes in the sky: Bayesian inference and model selection in cosmology*, 2008, Contemporary Physics, 49, 71 [303]

5 Bayesian Statistics and Markov Chain Monte Carlo

The Bayesian approach to data analysis has become more frequent in the last decades [303], showing great potential among many different fields of interest for the human kind, e.g. economy, biology, finance, astrophysics, and in particular for the asteroseismology [53, 57, 315, 144, 63, 6, 179], the main topic discussed in this dissertation, and the study of stellar differential rotation in active stars [93, 117, 115, 118], to which the Chapter 7 is dedicated.

The substantial growth of scientific works that employ Bayesian methods is mostly caused by a continuously-increasing amount of information contained in the data, which is therefore more difficult to be extracted and properly interpreted. Of course, the advances in techniques adopted in building the instrumentations represent one of the main goals for scientists to improve the quality of the data to be analyzed, a consequence of the increasing request for obtaining more accurate and reliable results that could broad our knowledge about astrophysical phenomena. One of the most plausible examples of this aspect is represented by the NASA's *Kepler* Mission, presented in Chapter 3, whose primary aim is the discovery of earth-sized exoplanets. The advanced quality of the instrumentation installed on board[1] allows scientists to acquire light curves with precisions up to few 10^{-6} mag, a value that sets a new limit in the precision of photometric observations. In order to gain a deeper understanding and a better interpretation of the information contained in such high quality data, a renovated type of tools for statistical analysis is certainly required for exploiting all of their potential. Nowadays Bayesian methods are providing results of great scientific value in the field of asteroseismology commonly by involving the extraction of the stars' asteroseismic parameters that characterize their Fourier power spectra. The reliability and robustness of such results appear to be better than the ones obtained with classic statistical methods (e.g. the minimum χ^2 criterion or Maximumx Likelihood Estimators). The reason relies mainly on the fact that the Bayesian approach offers a solution to the controversial problems of model selection.

On the one hand, the Bayesian statistics applied to solar-like oscillations is usually addressed to the extraction of the main asteroseismic properties of stars by exploiting the Fourier analysis of time-series (see Chapters 1 and 2 for a description of the classical approach to the analysis of time-series), where a low number of free parameters is involved. On the other hand, Bayesian analyses for the study of differential rotation in active stars usually deal with high-dimensional problems, where the number of free parameters can be even above one hundred. A detailed Bayesian analysis applied to a low-dimensional problem for the study of amplitude scaling relations of solar-like oscillations is presented in Chapter 6, while a description of a spot modeling for *Kepler* targets aimed at measuring the differential rotation, is given in Chapter 7,

[1] We remind to the reader that the photometer's sensitivity enables the detection of a earth-sized planet transiting on a G2V star with magnitude $m_V = 12^m$ and an integration time of 6.5 hours. For more details see Chapter 3.

together with a presentation of the results derived.

In the first half of this chapter we shall introduce a new formalism, which enables us to describe the *Bayesian probability theory*, whose fundamental meaning is enclosed in its most important statement, represented by the Bayes' Theorem[2]. The differences between Bayesian probability theory and the classic probability theory, known as the *frequentist* school of thought, will be described. Subsequently, we will show how to solve the so-called Bayesian Inference problem on the light of numerical simulations techniques in the form of *Markov Chain Monte Carlo* (MCMC) simulations (a brief introduction to the advantages carried out by the *Principal Component Analysis* (PCA) for reducing the dimensionality of the problem is given in Appendix D). Details about uncertainty estimation in the Bayesian approach are discussed in Appendix C.

In the second part of this chapter we will focus on the *Bayesian Evidence* by presenting the general problem related to model selection and its solution according to Bayesian's principle of Occam's razorm. Therefore, we provide a detailed description on how this method is able to weigh the quality of a model on the light of the data and the number of free parameters, together with some practical examples.

5.1 Introduction to Probability

The concept of probability as a degree of belief on a given *proposition* (event) was introduced for the first time by the mathematician and Presbyterian minister Thomas Bayes[3] [20] (1702 - 1761). Nevertheless this new definition of probability has been adopted consistently only within the latest years. This aspect should not be surprising if we consider the remarkable progress obtained in the field of computation, which obviously allowed for the use of many and onerous numeric techniques serving as auxiliary tools for statistics. An outline on the differences between the two most important schools of thought for probability theory is provided in the following, together with the reasons that lead us to consider the Bayes' statistics the preferred tool for data analysis and model selection.

5.1.1 The Frequentist Probability

A first, although approximate, concept of probability was given by Aristotle already in the IV century B.C. in his "Rhetoric"

> *"The likely is what occurs for most of the times."*

The classical idea of probability is called **frequentist probability**, asserted for the first time by Robert Leslie Ellis in 1842 and subsequently exposed in a more systematic and elaborated form by John Venn. Anyhow the rise of the frequentist school of thought is typically assigned to the mathematicians Jerzy Neyman and Egon Pearson, but for sure Pierre-Simon de Laplace, Adrien-Marie Legendre, Ronald Fisher and Harold Jeffreys and the contemporaries Edwin T. Jaynes and Larry Brethorst should be considered the fathers of the statistics and of the frequentist inference as it is actually known. The current definition of probability furnished by the classic view states that

[2] See [44] for a comprehensive introduction to Bayesian Statistics

[3] It was also reproduced in Biometrika, 45, 293-315, 1958.

"The probability of an event is defined as the number of times the event occurs over the total number of trials, in the limit of an infinite series of equiprobable repetitions."

Although apparently clear, this definition is unsatisfactory in many aspects, which we summarize in the following.

- The definition itself appears to be *circular*, i.e. it assumes that repeated trails have the same probability of outcomes, but it was the fundamental notion of probability itself that we were trying to define as first step.
- It cannot handle with unrepeatable events, a limit that excludes a class of physical phenomena which are generally related to a single sample of a particular physic system.
- The definition only holds exactly for an infinite sequence of repetitions. This never occurs in the reality because in practice we always handle with a finite number of measurements, which are sometimes only in a very small number (e.g. one of the biggest limits in asteroseismology is represented by the finite observing time). The frequentist approach cannot answer to the question "how can we manage events if we have only a handful of repetitions?", although some complicated *ad hoc* adjustments have been introduced for small sample size. Unfortunately physicists tend to forget about the infinite series requirement and apply the frequentist definition for any number of measurements they are dealing with, consequently affecting the results that go with them.
- Another more subtle aspect concerns the notion of "randomness". In order to understand this concept we refer to the paradigmatic example of coin tosses. By carrying out a long enough sequence of tosses, we will come up to the conclusion that the coin is deemed to be "fair" if the probability of getting heads is $p_H = 0.5$. If this probability is not satisfied we could assert that the coin has some physical imperfections. At this point it might appear plausible that the primary task becomes to determine such physical properties (e.g. a tensor of inertia symmetric about the plane of the coin) in order to understand how they affect the outcome of the coin itself. Anyhow, as forcefully argued by the american physicist Jaynes, who devoted the large part of his studies on mechanical statistics, the probability of the outcome of a sequence of tosses has nothing to do with the physical properties of the coin being tested [170]. In fact, a skilled coin-tosser (or equivalently a purpose-built machine [98]) can heavily influence the outcome even if the coin is not symmetric. The key of the problem is represented by the definition of random toss. Every time the coin-tosser influences the outcome, e.g. by modifying the speed or the spin of the coin, the randomness of the experiment will be clearly spoiled (this is what we call "cheating"). It appears thus that the outcome depends on our state of knowledge about the initial conditions of the system (namely angular momentum and velocity of the toss). Is therefore correct to assert that a lack of precise information about the initial conditions results in a state of indifference about the possible outcome.

5.1.2 The Bayesian Probability

Many of the limitations above exposed can be avoided and paradoxes solved by considering another definition of probability, know as the **Bayesian probability**. The Bayesian viewpoint is based on the simple and intuitive tenet that

"Probability is a measure of the degree of belief about a proposition."

It is immediately clear that this statement can be applied to any kind of proposition because the notion of probability becomes a state of knowledge in presence of a limited number of information. Moreover another advantage is that it deals with uncertainties independently of their origin, i.e. there is no distinction between "statistical uncertainty", arising from the finite precision of the measurement apparatus and the relative random noise, and "systematic uncertainty", deriving from deterministic effects whose nature is only partially known (e.g. calibration uncertainty of a spectrograph, computation errors introduced within the codes used for the data analysis, instrumental drifts in the light-curves caused by temperature variations of the spacecraft). From the coin tossing example we learn that the randomness is directly related to our lack of information about the initial conditions of the system (if we know precisely the way the coin is flipped we could then predict the outcome of any toss with certainty).

The growing complexity of theoretical models and of the information contained in the data are causing an increasing request for more sophisticated statistical tools, which will soon represent the limit of our interpretations. Moreover due to the non-unlimited resources available for scientific projects, the optimization of their success is becoming a tough problem. In this aspect the Bayesian statistics is considerably helpful, especially for what concerns the estimation of outcomes arising from a rigorous ponderation of our degree of uncertainty about that particular field.

It seems that the above mentioned arguments strongly favor the Bayesian view of probability, and for physicists the approach that yields demonstrably superior results ought to be preferred. We now list some good reasons to prefer such type of approach:

- It can be shown that the *application of Bayes' Theorem recovers frequentist results (in the long run) for cases simple enough where such results exist*, while remaining applicable to propositions that cannot even be formulated in a frequentist context.

- Bayesian inference deals effortlessly with the so called *nuisance parameters*. These parameters have to be considered within the data interpretation but no any physical information is carried with them (dealing analytically with this aspect will be discussed below).

- In many situations of physical relevance the *prior information* is highly outstanding. In fact, a lack of information on our state of knowledge before any experiment can seriously affect the outcomes of our inference. A typical example concerns the use of constrained likelihood methods for the determination of parameters whose values are restricted by their physical meaning, e.g. mass, counts, luminosity and distances. In the frequentist approach the result over these parameters can provide also negative values, while we know that all these quantities must be positive. Instead by considering prior information accounted for in the final inference we ensure that physically meaningless results are weeded out from the beginning.

- Bayesian statistics only deals with *data that were actually observed*, while frequentist methods usually focus on the distribution of possible data that have not been obtained! Thus *frequentist results depend on what the experimenter thinks about the probability of data that have not been observed* (this is called the "stopping rule" problem). The argument above is clearly absurd and consequently the classic viewpoint is providing an inference lacking in rigorousness. In fact, a serious and robust inference should not absolutely depend on what could happen but should be conditional only to whatever has already occured. Bayesian methods directly include this aspect because their inferences are by construction conditional on the observed data.

5.2 Bayes' Theorem

The analytic expression used to describe the degree of belief on a proposition is called Bayes' Theorem. It can be derived from a set of basic consistency requirements for plausible reasoning, known as Cox axioms [89]. Therefore Bayesian probability theory can be shown to be the unique generalization of a logical deduction when the available information is incomplete.

5.2.1 Cox Axioms

In order to introduce a mathematical treatment of the Bayesian statistics we adopt a fairly relaxed notation which can be easily understood. Let us consider a generic proposition A, where $p(A)$ could be for instance the probability to obtain head with a coin toss and $p\left(\bar{A}\right)$ its complementary. The *sum rule* reads

$$p(A \mid I) + p\left(\bar{A} \mid I\right) = 1\,, \tag{5.1}$$

where I represents any relevant information that is assumed to be true (it can be important to prove the veracity of A) and $p(A \mid I)$ is the conditional probability of A, i.e. the probability that A occurs given as verified the information I. By introducing a second proposition B the *product rule* gives

$$p(A, B \mid I) = p(A \mid B, I)\, p(B \mid I)\ , \tag{5.2}$$

which says that the *joint probability* of A and B equals the probability of A given that B occurs times the probability of B occurring on its own (both conditional on information I). If we are interested in the probability that B occurs singularly, assuming the existence of A, then we have to sum over all the possible outcomes of the proposition A. By adopting the sum and the product rule together we obtain

$$\begin{aligned} p(B \mid I) &= \sum_A p(A, B \mid I) \\ &= \sum_A p(A \mid B, I)\, p(B \mid I)\ . \end{aligned} \tag{5.3}$$

The quantity shown in the left-hand side of the equation is called *marginal probability* of B.

Since obviously $p(A, B \mid I) = p(B, A \mid I)$ we can derive the *Bayes' Theorem* directly by rewriting the product rule, giving

$$p(B \mid A, I) = \frac{p(A \mid B, I)\, p(B \mid I)}{p(A \mid I)}\,. \tag{5.4}$$

The interpretation can be simple to be understood if one replaces for A the observed data d (e.g. the fluxes in a light-curve attained with a photometer) and for B the hypothesis H we want to asses. We then obtain

$$p(H \mid d, I) = \frac{p(d \mid H, I)\, p(H \mid I)}{p(d \mid I)}\,. \tag{5.5}$$

On the left-hand-side $p(H \mid d, I)$ represents the *posterior probability*, namely the probability that hypothesis H is verified taking the data into account. It is proportional to the *sampling distribution* of the data $p(d \mid H, I)$, which is not a probability distribution but encodes how the degree of plausibility of the hypothesis changes when we acquire the data. Considered as a function of the hypothesis H (for a given dataset d) the sampling distribution is called

likelihood function and we can employ the shortcut notation $\mathcal{L}(H) \equiv p(d \mid H, I)$. The second term $p(H \mid I)$ is usually known as the *prior probability.* which is heavily conditioned to our degree of knowledge about the measurements before we acquire data, as already announced above. The last term on denominator of the right-hand-side of Eq. (5.5) is called the *marginal likelihood* (also known as the Bayesian Evidence), which can be expressed by means of Eq. (5.3) as

$$p(d \mid I) \equiv \sum_{H} p(d \mid H, I)\, p(H \mid I)\,. \tag{5.6}$$

In the forthcoming sections we will show how the posterior probability distribution allows for the derivation of parameter estimates. Nonetheless, posteriors do not play a fundamental role for the model comparison issue, which is determined by the marginal likelihood and prior choice instead, as we shall describe in the second part of the current chapter. Thus, the marginal likelihood represents just a normalization constant within the Bayesian Inference problem and as a consequence it is not changing the posterior probability pattern but the "amplitude", say, of the likelihood function employed.

5.2.2 Prior Probability

One of the most important and at the same time potentially dangerous and criticized aspects of the Bayesian approach is related to the subjectivity conditioning the posterior probability. In fact, it is historically believed that results obtained from considering subjective elements may appear as non objective. As a matter of fact the state of knowledge the experimenter has got about the observed event is exactly what makes the data interpretation objective, as we shall argue below.

The inference problem has no solution without any prior assumptions, as the guideline of the Bayesian logic states. In fact, any choice of such assumptions requires high accuracy level in order to afford the best representation of our degree of knowledge before acquiring the data. Nevertheless priors should be regarded as a feature of the Bayesian statistics, rather than a limitation. Thus, priors allow for a correct weighting of results obtained from different experimenters, who are possessing different states of knowledge about the event (e.g. observational and theoretician points of view), accordingly for an advantageous comparison of the data interpretation.

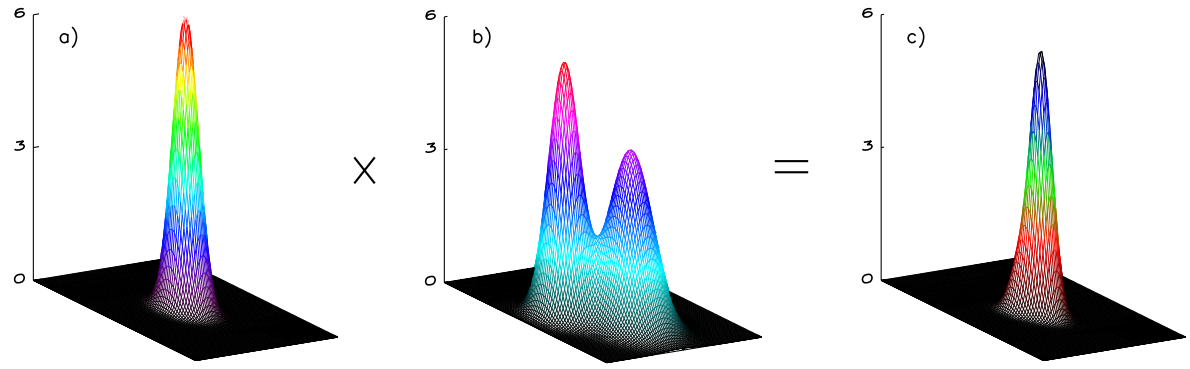

Figure 5.1: *Example of a 2-dimensional parameter inference dominated by data information, where distributions are represented as functions of two variables along the vertical z-axis (arbitrary units). The posterior probability distribution (c) appears to be clearly well constrained as the likelihood function does (a) despite of two different assumptions representing the priors of two different experimenters (b). Hence, priors distribution is not influencing the outcome since the dataset employed is strong enough to constrain the result. As a consequence both the experimenters are converging to the same result through the Bayes' Theorem.*

Example

We can now mention an example about the role of priors that might help the reader to understand how priors can influence the results. When the information carried on by the data sampling distribution is relevant with respect to the one contained in the priors then the posterior probability derived by two different scientists will converge to the same result, i.e. objective inference on the hypothesis, as shown in Figure 5.1. This means that the information included in the dataset is exhaustive enough for solving the inference problem. On the other hand, when the priors play a fundamental role for the posterior probability then a great care must be given on their choice. This situation is often reproduced in real cases since the data are usually not strong enough to override the prior (e.g. for small sample sizes or for dimensionality problems of the hypotheses space, see Figure 5.2). Hence we have learned something useful about the constraining power (or lack thereof) of the data.

After all we can conclude that the objectivity of the Bayesian method relies on the following statement: *two different experimenters with the same state of knowledge about the event must assign the same priors, hence obtaining identical posterior probability distributions by using the same dataset.*

Anyhow, any existing typologies of priors can be used for the data analysis (see also [124] and references therein for further details on prior choice). In order to be probability distributions, priors must be *proper* i.e. normalizable to unity probability content. Essentially the most interesting ones are:

1. *Ignorance Priors*: reflect the state of indifference with respect to the symmetries of the problem taken into account.

2. *Flat Priors*: priors of standard choice in Bayesian parameter estimation and especially for the analysis described in the following chapters. A flat prior distribution is given when we assume a constant state of knowledge for all the possible outcomes of our parameters (Figure 5.3). Flat priors are in general *improper*, in the sense they are not normalizable to unity as a probability density function (PDF) has to, which is the case when an infinite

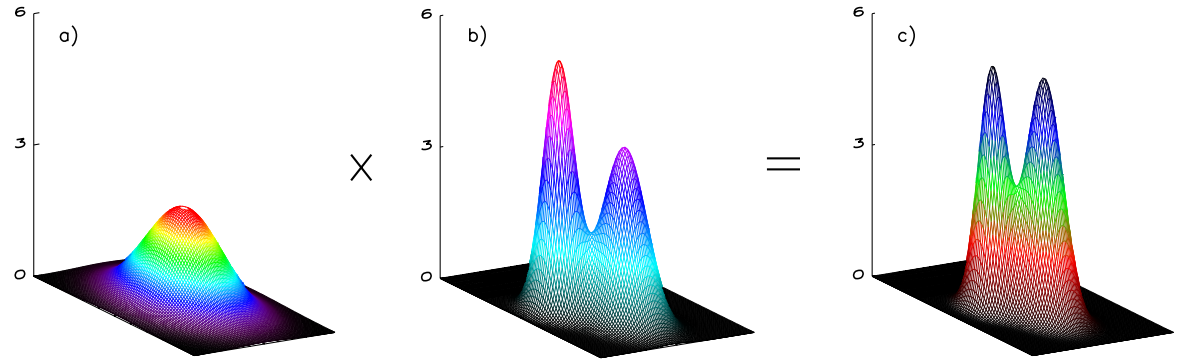

Figure 5.2: *Example of a 2-dimensional parameter inference dominated by priors information (arbitrary units along the vertical z-axis). The posterior probability distribution (c) is clearly hampered by the prior distribution patterns of the two different experimenters (b). Hence, since the dataset employed is not large enough to provide all the information, the likelihood function is not ensuring the same outcome for both experimenters through the Bayes' Theorem.*

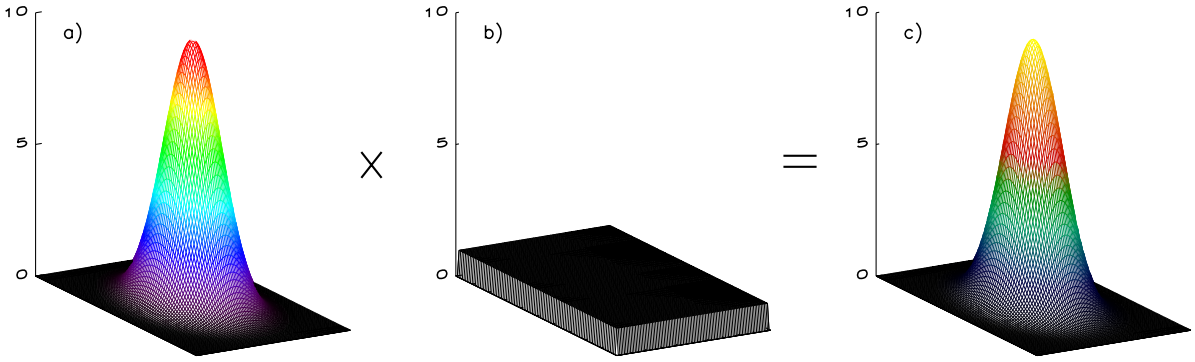

Figure 5.3: *Example of a 2-dimensional flat prior distribution (b). The posterior probability distribution (c) has the same 2-dimensional Gaussian shape of the likelihood function (a), meaning that our state of knowledge about the parameters is non-committal, i.e. we have no information about the expected outcomes and hence priors are not conditioning the result given by the Bayes' Theorem.*

range of values of the parameter is taken into account (e.g. the real axis). Anyhow, the improperness is generally not an issue because one usually constraints the number of expected values by limiting the parameter to a known interval of variation (see also [297]). For instance, the *proper* flat prior of a physical quantity β limited to the range $[\beta_{\min}, \beta_{\max}]$ is simply given as $p(\beta) = (\beta_{\max} - \beta_{\min})^{-1}$, i. e. it gives a constant value equal to the reciprocal of the length of the interval of values where β is supposed to be measured. Such choice allows the prior to satisfy the normalization condition

$$\int_{\beta_{\min}}^{\beta_{\max}} p(\beta) d\beta = 1\,, \tag{5.7}$$

Thus, flat priors reflect an absolute lack of information about the parameter for which they are adopted. For instance, one can consider again a one-dimensional case dealing with the parameter β (e. g. the inclination angle of the rotational axis of a star). When a particular value of β is not supposed to be predominant among the others, a flat and space limited prior distribution is required (e.g. for the inclination angle we expect to observe values within the range $[-\pi/2, \pi/2]$). Nevertheless we should not forget that a flat prior

distribution for β does not turn into a flat prior distribution for any non-linear function $\psi(\beta)$, because a uniform distribution is not invariant under re-parametrization (see the discussion in [297]). In fact, being $p(\beta)$ a probability *density* function (PDF), $p(\beta)\,d\beta$ is the probability to find β within the range β and $\beta + d\beta$. Since we are considering the same physical information (e. g. the inclination of the rotational axis of a star), although through two different functions of the parameter, the two probability distributions must coincide. Thus we have

$$p(\beta)\,d\beta = p(\psi)\,d\psi\,, \tag{5.8}$$

and hence the PDF for the new function $\psi(\beta)$ will be

$$p(\psi) = p(\beta)\left|\frac{d\beta}{d\psi}\right|. \tag{5.9}$$

This relation describes how a flat prior for the parameter β may be extremely constraining for a function $\psi(\beta)$ of the same parameter instead. This relies on the derivative term $|d\beta/d\psi|$ appearing in the right-hand-side of the equation (in a multi-dimensional case the derivative is represented by the determinant of the Jacobian matrix for the transformation). Parameters inference can be sometimes facilitated by adopting an adequate function of the parameter which transforms a non-flat prior distribution into a flat one. A very simple and clarifying example is represented by the inclination angle i of the rotational axis of a star. As well known, the observed component of the rotational velocity of a star is proportional to $\sin i$, where i is taken as the angle between our line of sight and the rotational axis of a star. Hence the expected prior distribution $p(i)$ for the parameter i will follow a sine law (Figure 5.4a), clearly peaked at the best angle for measuring the rotational velocity (angles toward 90° are more likely to be observed). On one hand, a non-flat prior distribution must be used if we consider the parameter i itself, which will clearly yield to more complex structures in the posterior PDF (see Section 5.3). On the other hand, the problem can be solved by simply using the function $\psi(i) = \cos i$, which is able to flatten our prior distribution (Figure 5.4b) through Eq. (5.9). In fact, the advantage of flat priors is that they can be included in the inference problem as a simple interval of values for the parameter. Therefore the choice of the new parameter $\cos i$ limited to the range $[-1, 1]$, will overcome further complexities in the inference problem.

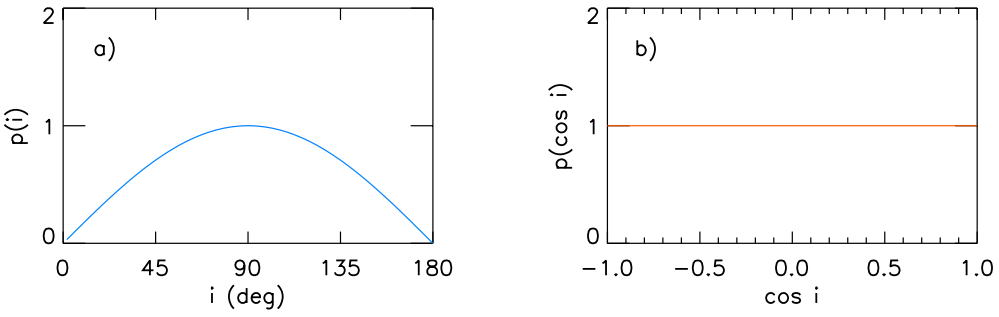

Figure 5.4: *The inclination angle i of the rotational axis of a star. The prior distribution of the expected values for the angle i is shown in panel (a), deriving from the observed component of the rotational velocity of the star, which follows a sine law. An adequate choice of a function of the parameter will lead to a flat prior distribution, panel (b), through Eq. (5.9), where $\psi(i) = \cos i$. Hence the enclosing of such prior over the parameter within the inference, will be simply expressed as the range $[-1, 1]$ of allowed values imposed on the new parameter.*

3. *Reference Priors*: exploit the experience acquired by the scientist to construct the least informative state of knowledge. They were first described by [40] and subsequently further developed by [39]. The method adopted for deriving reference priors is usually known as the Berger-Bernardo method.

4. *Jeffreys' Priors*: a particular class of reference priors [184] used when we are ignorant about the *scale* of a given quantity β. In this case, it can be shown that the appropriate prior is flat on $\ln \beta$ because it gives equal weight to all orders of magnitude [170]. From Eq. (5.9) we can derive that the corresponding Jeffreys' prior for β is given by $p(\beta) \propto \beta^{-1}$.

   Another important feature of the Jeffreys' priors concerns the treatment of posteriors probability distributions of different but correlated physical parameters, that is of parameters that carry on the same physical information. In order to understand the problem we can provide the example of two couples of physical parameters, i.e. the period and frequency (P, ν) related to a p-mode oscillation and the radius and area (r, A) of a flat circular spot laying on a stellar photosphere. Since clearly $P = \nu^{-1}$ and $A = \pi r^2$, the Bayes' Theorem will provide two different posterior distributions for both P, ν and r, A, without allowing for any statistical comparison of the same physical information contained. Conversely, if we adopt the new parameters $p = \ln P$ and $a = \ln A$ we obtain $p \propto \ln \nu$ and $a \propto \ln r$. Hence the PDFs of the new parameters p and a will give identical statistical moments whether we consider periods and radii or frequencies and areas. Assuming those parameters to have flat priors, then P and A will have Jeffreys' priors given by $p(P) \propto P^{-1}$ and $p(A) \propto A^{-1}$. The adoption of such priors is of importance in the physics field because different functions of the same physical observable can be treated equivalently, in a manner independent of our choice of the function. Applications of this explanation are treated in Chapters 6 and 7.

5.3 Parameters Inference

Once we have introduced the main elements and features of the Bayesian mindset we can proceed with the description and solution of the parameters inference problem. We remind to the reader that in statistics the *inference problem* represents the estimation of population parameters by means of quantities derived from a single (or more than one) sample, which is usually only a small subset of the entire population[4]. It has been shown [281] that if the sample size $n \geq 30$, many times a quantity derived from the sample with the purpose of estimating a population parameter, hereafter denoted as *statistic*[5], appears to be normally distributed. The probability distribution of a sample statistic, is known as the *sampling distribution* of the statistic. In particular, a statistic is called an *unbiased estimator* of the corresponding population parameter if the mean (or expectation) of the statistic, which called *unbiased estimate*, is equal to the population mean.

In the context of Bayesian statistics, we first choose a model, which we denote as $\mathcal{M}$, containing a set of hypotheses H formalized into a vector of free parameters, $\boldsymbol{\xi}$, known as the *parameters vector*. The parameters might describe any aspect of the model, but usually they identify some physically meaningful quantities (e.g. the frequency, amplitude and damping time of a p-mode oscillation). Together with the model assumption, the priors for the parameters

[4] In general, a population can be either finite or infinite. In most of the cases, however, only small samples of the entire population can be analyzed.

[5] Any quantity arising from a sample of measurements of a given population that is derived with the purpose of estimating population parameters is called *sample statistic* or briefly *statistic*. In general, depending on the population parameter we want to estimate, there will be a precise statistic to be derived from the sample.

must be specified. Besides, the prior for a new observation might be taken to be the posterior from a previous measurement.

The central step of the Bayesian inference is the construction of the likelihood function, which usually reflects the way the data are obtained. For instance, a measurement with a Gaussian noise will be represented by a Normal distribution, while cosmic rays counts on a detector will have a Poisson distribution for a likelihood. Thus the Bayes' Theorem given by Eq. (5.5) reads

$$p\left(\boldsymbol{\xi} \mid d, \mathcal{M}\right) = \mathcal{L}(\boldsymbol{\xi}) \frac{\pi(\boldsymbol{\xi} \mid \mathcal{M})}{p(d \mid \mathcal{M})}\,, \tag{5.10}$$

where $p\left(\boldsymbol{\xi} \mid d, \mathcal{M}\right)$ is the posterior PDF, $\mathcal{L}(\boldsymbol{\xi}) = p(d \mid \boldsymbol{\xi}, \mathcal{M})$ is the likelihood function, $\pi(\boldsymbol{\xi} \mid \mathcal{M})$ the chosen priors, and $p(d \mid \mathcal{M})$ is the Bayesian evidence, the latter described in Section 5.6.2.

5.3.1 Marginalization Problem

We therefore consider the parameters vector $\boldsymbol{\xi}$ divided into physically interesting parameters, $\boldsymbol{\xi}_I$ and no physically meaningful parameters (*nuisance* parameters), $\boldsymbol{\xi}_N$. Hence, since we are considering the vector $\boldsymbol{\xi} = (\boldsymbol{\xi}_I, \boldsymbol{\xi}_N)$, Eq. (5.10) can be rewritten as

$$p\left(\boldsymbol{\xi}_I, \boldsymbol{\xi}_N \mid d, \mathcal{M}\right) = \mathcal{L}(\boldsymbol{\xi}_I, \boldsymbol{\xi}_N) \frac{\pi(\boldsymbol{\xi}_I, \boldsymbol{\xi}_N \mid \mathcal{M})}{p(d \mid \mathcal{M})}\,. \tag{5.11}$$

We assume to have a continuous *parameter space*[6], hereafter denoted as $\Omega_{\mathcal{M}}$. If we suppose to have m interesting parameters, $\boldsymbol{\xi_I} = \{\xi_{I_1}, \xi_{I_2}, \ldots, \xi_{I_m}\}$ and n nuisance parameters $\boldsymbol{\xi_N} = \{\xi_{N_1}, \xi_{N_2}, \ldots \xi_{N_n}\}$, and assuming each of them to have a delimited range of variation, the parameter space can be expressed as

$$\begin{aligned} \Omega_{\mathcal{M}} = &\left[\xi_{I_1}^{\max} - \xi_{I_1}^{\min}\right] \times \left[\xi_{I_2}^{\max} - \xi_{I_2}^{\min}\right] \times \cdots \times \left[\xi_{I_m}^{\max} - \xi_{I_m}^{\min}\right] \\ &\times \left[\xi_{N_1}^{\max} - \xi_{N_1}^{\min}\right] \times \left[\xi_{N_2}^{\max} - \xi_{N_2}^{\min}\right] \times \cdots \times \left[\xi_{N_n}^{\max} - \xi_{N_n}^{\min}\right]\,, \end{aligned} \tag{5.12}$$

namely the cartesian product of the intervals of the free parameters that formalize the hypotheses of the model, and whose ranges of variation are fixed by our choice of the priors.

At this stage, we can finally solve the inference problem by marginalizing the posterior PDF over the nuisance parameters. This is done by integrating Eq. (5.11) over all the possible values of the nuisance parameters, obtaining

$$p(\boldsymbol{\xi}_I \mid d, \mathcal{M}) \propto \int_{\Omega_{\mathcal{M}}} \mathcal{L}(\boldsymbol{\xi}_I, \boldsymbol{\xi}_N) \pi(\boldsymbol{\xi}_I, \boldsymbol{\xi}_N \mid \mathcal{M}) d\boldsymbol{\xi}_N\,, \tag{5.13}$$

where we excluded the marginal likelihood $p(d \mid \mathcal{M})$ since it is irrelevant for the parameter inference (but central for model comparison, see Section 5.6). We can further marginalize our solution for each ξ_{I_j}, where $j \in [1, m]$, by integrating the posterior $p(\boldsymbol{\xi}_I \mid d, \mathcal{M})$ with respect to

[6] In Bayesian statistics a parameter space is the multi-dimensional space identified by our choice of the priors, thus its total volume is given by the product of the intervals of each free parameter considered in the inference problem. In general, we refer to prior parameter space and posterior parameter space as the effective volume of the prior and posterior PDFs, respectively. The dimensionality of the parameter space equals the number of free parameters involved within the model chosen to represent the data. A continuous parameter space is a space where parameters' values are expressed as continuous variables. Although continuous parameter spaces are useful for a mathematical treatment of the argument, only discrete parameter spaces are considered for practical cases.

the remaining m-1 parameters. We will then obtain a one-dimensional posterior PDF for each parameter ξ_{I_j}. Without losing in generality we can assume $j = 1$ and write

$$p(\xi_{I_1} \mid d, \mathcal{M}) = \int_{\Omega_{\mathcal{M}}} p(\boldsymbol{\xi}_I \mid d, \mathcal{M}) d\xi_{I_2} d\xi_{I_3} \dots d\xi_{I_m} , \tag{5.14}$$

where we have assumed that $p(\xi_{I_1} \mid d, \mathcal{M})$ satisfies the normalization condition

$$\int_{\xi_{I_1}^{min}}^{\xi_{I_1}^{max}} p(\xi_{I_1} \mid d, \mathcal{M}) d\xi_{I_1} = 1 . \tag{5.15}$$

Eq. (5.14) is called the *marginal* PDF of the parameter ξ_{I_1}. In Bayesian statistics, all the information about the parameter are drawn from its marginal PDF, namely represented by the statistical moments of the distribution (usually 1^{st} and 2^{nd} momentum are the most used quantities). Eq. (5.14) constitutes the so called *marginalization problem* and its solution will be discussed in Section 5.5, while a technical description about the calculation of the uncertainties is provided in Appendix C.

5.4 Markov Chain Monte Carlo

Although apparently simple in its formulation and meaning, the marginalization problem has usually no analytical solution. For example, for models representing the star spot activity as the one described in this dissertation, the number of parameters employed for reproducing the modulation of an observed light-curve might be up to the order of a hundred, increasing the complexity of the shape of the posterior PDF. Hence Eq. (5.13) becomes completely unsolvable with analytical treatment.

Nowadays the computational power available is encouraging the introduction and development of sophisticated numerical techniques. One of the most interesting and powerful inference techniques, which is actually the base of the wide applicability of the Bayesian method, is obtained by means of the Markov Chain Monte Carlo (MCMC). An introduction about the notion of Monte Carlo numerical simulations and about the properties of a Markov Chain will be therefore provided [41, 228, 261, 247]. Later in this chapter we will show how MCMC can solve the parameters inference when the marginalization problem presented in Section 5.3 has no analytical solution.

5.4.1 Monte Carlo Methods

The term *Monte Carlo* is referred to a wide class of computational algorithms that rely on repeated random sampling to compute their results. This class of algorithms is particularly useful for simulating physical and mathematical systems (where a large number of coupled degrees of freedom is involved). The idea of Monte Carlo simulation was coined in 1940s by Enrico Fermi and Stanislaw Ulam while they were studying neutron collisions at the Los Alamos Scientific Laboratory. Since the problem of the collision had not an analytical solution the charge to create a secret code for simulating the experiment was committed to John von Neumann, which named it in homage to the famous Monte Carlo casino. The fundamental feature distinguishing a Monte Carlo simulation from a classic one is related to the manner of reconstructing a deterministic problem. In fact, Monte Carlo simulations inverts systematically the way to approach to a deterministic problem by starting from the research of a probabilistic analog, conversely to what normal algorithms do. Many different approaches are used within a Monte Carlo simulation. However some common features characterizing such simulations can be identified in the following pattern:

1. Define a domain of possible inputs.
2. Generate inputs randomly from a probability distribution over the domain.
3. Perform a deterministic computation on the inputs.
4. Aggregate the results.

For instance consider a circle inscribed in a square. Given that the circle and the square have a ratio of areas that is $\pi/4$ (a result this one that can be attained from geometry), the value of π can be approximated using a Monte Carlo method:

1. Draw a square on the ground, then inscribe a circle within it.
2. Uniformly scatter some objects of uniform size (grains of rice or sand) over the square.
3. Count the number of objects inside the circle and the total number of objects.
4. The ratio of the two counts is an estimate of the ratio of the two areas, which is $\pi/4$. Multiply the result by 4 to estimate π.

In this procedure the domain of inputs is the square that circumscribes our circle. We generate random inputs by scattering grains over the square, then perform a computation on each input (test whether it falls within the circle). Finally, we aggregate the results to obtain our final result, the approximation of π.

Other two important properties are arising from the above description. First, the quality of the randomness whereby random numbers are generated should be always high. Second, there should be a large number of inputs, generally by adopting long run simulations. Obviously the more is the randomness of the inputs the more will be the quality of the results, as well as their precision will increase by extending the duration of the simulation.

Lastly, we can complete this overview on the Monte Carlo numerical simulations by listing the most interesting areas in which they are employed. Presently these methods are extremely important for physical sciences, especially in computational physics, chemical and statistical physics, molecular modeling, particle physics, QCD, galaxy and cosmological modeling and of course star spot modeling and p-mode oscillations modeling. Monte Carlo methods are also applied to other fields such as engineering, applied statistics, design and 3D graphic, finance and business, telecommunication and artificial intelligence for video games. Moreover they are very useful for computational mathematics related to integration methods, numeric optimization and inverse problem solution.

5.4.2 Stochastic Processes

In order to introduce and understand the notion of Markov Chain we firstly have to describe the meaning of a random process. A random process, also called *stochastic*, is the counterpart to a deterministic process. Instead of dealing with only one possible reality of how the process might evolve under time (e.g. for solutions of an ordinary differential equation), in a stochastic or random process there is some indeterminacy in its future evolution described by probability distributions. This means that a given physical system laying on a state s_i at the time t_i, will not pass to a subsequent state s_{i+1} at the time t_{i+1} in a deterministic manner, but accordingly to a certain probability distribution. Hence even if the initial condition (or starting point) is known, there are many possibilities the process might go to, but some paths may be more likely and others less so.

A mathematical formalization for the notion of stochasticity of a system requires the introduction of some terminologies. First of all we assume to deal with discrete states[7], which are collected in a finite or countable space known as the *state space*. We will denote this space with $\mathcal{S} = \{s_i\}_{i\in I}$, where $I \subset \mathbb{N}$ is a finite or countable set and s_i is the ith state occupied by the system. Sometimes this state will be known while in other cases one can just talk about the probability for the state to be occupied. Therefore it is reasonable to introduce a probability distribution λ on $\mathcal{S}$. It can be represented by a set of values λ_i, where $i \in I$, such that

$$\lambda_i \geq 0 \quad \forall i \in I, \quad \sum_{i\in I} \lambda_i = 1\,, \tag{5.16}$$

namely greater than (or equal to) zero and whose sum is equal to the unity. A distribution having $\lambda_i = 1$ for some i and $\lambda_j = 0$ when $j \neq i$ is called *concentrated* at point i. Then the state of our system becomes *deterministic*. When $\sum_{i\in I} \lambda_i \neq 1$ instead, then λ is called a *measure* on $\mathcal{S}$. If such measure is finite, i.e. $\sum_{i\in I} \lambda_i < \infty$, then λ can be transformed into a probability distribution by the normalization

$$\tilde{\lambda}_i = \frac{\lambda_i}{\sum_{j\in I} \lambda_j}\,, \tag{5.17}$$

since we have

$$\sum_{i\in I} \tilde{\lambda}_i = \frac{\sum_{i\in I} \lambda_i}{\sum_{j\in I} \lambda_j} = 1\,. \tag{5.18}$$

On the other hand, when the measure is infinite, i.e. when $\sum_{i\in I} \lambda_i = \infty$, it is still possible to transform λ into a probability distribution by considering a finite subset $J \subset I$ such that $\lambda(J) = \sum_{i\in J} \lambda_i < \infty$.

As a second step we have to introduce the concept of *discrete-time* stochastic processes. We refer to systems laying in a single state for each single time-step. Time-steps are formally associated to integer numbers, $n \in \mathbb{N}$. For example we can couple the times $(0, 1, \ldots, m)$ to the states $(s_0, s_1, \ldots, s_m)$. Generally a discrete-time stochastic process is represented as a set $\mathbf{\Phi} = \{\Phi_0, \Phi_1, \ldots, \Phi_n\}$ of individually measurable random variables, where $\Phi_i \in \mathcal{S} \quad \forall i \in [0, n]$, hence belonging to the state space.

We can now describe the stochastic mechanism that allows the system to transit between two successive states. The transition is made possible by means of the so called *transition matrix* or *matrix of transition probabilities* $\mathbf{P}$, whose entries p_{ij}, with $i, j \in I$, are the probabilities for the system to transit from a state s_i to a state s_j in a single time-step. The nth power of the transition matrix, $\mathbf{P}^n$, will provide the transition probability for n consecutive time-steps. Lastly, a transition matrix is called *stochastic* if it satisfies the properties

$$0 \leq p_{ij} \leq 1 \quad \forall i, j \in I\,, \quad \sum_{j\in I} p_{ij} = 1 \quad \forall i \in I\,, \tag{5.19}$$

where each entry in $\mathbf{P}$ is non-negative but not greater than 1, and the sum of entries along every row equals the unity. By analogy, a probability distribution λ on $\mathcal{S}$ is often called a *stochastic vector*. Then a stochastic matrix is one where every row is a stochastic vector. In fact, we can think of p_{ij} as the conditional probability for the system to occupy the state s_j starting from the state s_i, after a single time-step.

[7]Discrete stochastic processes are the simplest and most interesting to be treated since when dealing with numerical simulations only discrete quantities are involved. Anyhow continuous stochastic processes are studied analytically [296].

The simplest case of a stochastic matrix is the one in a two-state space. We refer to the states $A = s_0$ and $B = s_1$ and we denote with t_0, t_1 the two corresponding time-steps. Without loss of generality, the stochastic matrix has the form of the 2×2 matrix

$$\mathbf{P} = \begin{array}{c} \\ A \\ B \end{array} \begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix} \end{array}$$

where $0 \leq \alpha, \beta \leq 1$. In particular, $\alpha = \beta = 0$ gives the identity matrix $\mathbf{I}$ and $\alpha = \beta = 1$ the anti-diagonal matrix. A system with the identity transition matrix remains in the initial state forever; in the anti-diagonal case it flips state every time, from A to B and *vice versa*. If we consider $\alpha = \beta = 1/2$ instead, the system may stay either in the same state or flip into the other one with equal probabilities. It is sometimes convenient to represent the transition matrix by an oriented graph where arrows show all the possible transitions, see Figure 5.5.

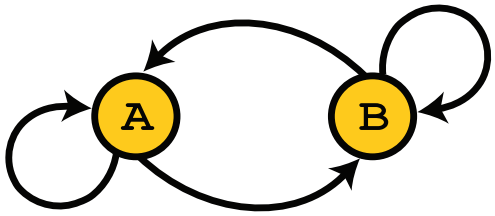

Figure 5.5: *Oriented graph of a 2-states stochastic process.*

5.4.3 Markov Processes

Very common examples of stochastic processes concern random movements, namely *random walk* or *Brownian motion*. Here below we would like to stress their most important and useful features.

The russian mathematician Andrey A. Markov (1856 - 1922) is best known for his work on theory of stochastic processes. In 1907 he focused his research on a new idea of processes related to the probability, where the outcome of some experiment can influence the outcome of a subsequent experiment. Such subclass of stochastic processes was then called *Markov Chain*. Although at first judged an unimportant discover by Markov himself, Markov Chains are nowadays widely used for simulations of random walks in the exploration of posterior probability distributions or likelihood spaces.

Formally, a discrete-time Markov Chain (DTMC)[8] is a discrete-time stochastic process for which the *Markov property* (or memoryless property) holds. For a mathematical introduction of a DTMC we refer to the terminology presented in Section 5.4.2. We denote as Φ_n the state occupied by the system at the time n and we assume to have an initial probability distribution λ on $\mathcal{S}$ and a transition matrix $\mathbf{P}$. The framework is completed by specifying the initial state of the system, that is the state where the system lies at the initial time $n = 0$. Therefore we assume that the initial state of our DTMC is $\Phi_0 = s_i$, where $s_i \in \mathcal{S}$ and it is occupied with a probability λ_i. Subsequently the rules defining a Markov Chain can be formulated as:

1. Φ_0 has a probability distribution λ:

$$\mathbb{P}(\Phi_0 = s_i) = \lambda_i \quad \forall i \in I\,,$$

[8] We refer to this category of Markov Chain since it is the one of interest for practical studies.

meaning that the probability the system occupies one of the states of the state space $\mathcal{S}$ at its initial time $n = 0$ is given by the initial probability distribution λ,

2. more generally, for all time-steps n and the states $s_0, s_1, \ldots, s_n \in \mathcal{S}$, the total probability $\mathbb{P}(\Phi_0 = s_0, \Phi_1 = s_1, \ldots, \Phi_n = s_n)$ that the system occupies states $s_0, s_1, \ldots, s_n$ at times $0, 1, \ldots, n$ is written as

$$\mathbb{P}(\Phi_0 = s_0, \Phi_1 = s_1, \ldots, \Phi_n = s_n) = \lambda_0 \cdot p_{01} \cdot p_{12} \cdots p_{(n-1)n}\,, \tag{5.20}$$

namely the product between the probability to occupy the initial state and the transition probabilities from a state to the successive one. Of course (1) is a particular case of (2) with $n = 0$.

The Markov property is expressed by an important corollary of Eq. (5.20), which defines the conditional probability $\mathbb{P}(\Phi_{n+1} = s_b | \Phi_0 = s_0, \Phi_1 = s_1, \ldots, \Phi_n = s_a)$ that the state occupied by the system at time $n + 1$ is $s_{n+1} = s_b$, given that the states $s_0, s_1, \ldots, s_{n-1}$ and $s_n = s_a$ have been occupied at times $0, 1, \ldots, n-1, n$:

$$\begin{aligned} &\mathbb{P}(\Phi_{n+1} = s_b | \Phi_0 = s_0, \ldots, \Phi_{n-1} = s_{n-1}, \Phi_n = s_a) \\ &= \frac{\mathbb{P}(\Phi_0 = s_0, \ldots, \Phi_{n-1} = s_{n-1}, \Phi_n = s_a, \Phi_{n+1} = s_b)}{\mathbb{P}(\Phi_0 = s_0, \ldots, \Phi_{n-1} = s_{n-1}, \Phi_n = s_a)} \\ &= \frac{\lambda_0 \cdot p_{01} \cdots p_{(n-1)a} \cdot p_{ab}}{\lambda_0 \cdot p_{01} \cdots p_{(n-1)a}} = p_{ab}\,, \end{aligned} \tag{5.21}$$

meaning that the conditional probability on the successive state Φ_{n+1} does not depend on the states $s_0, s_1, \ldots, s_{n-1}$ but only on the state Φ_n at the last preceding time n. The *memorylessness* illustrated by Eq. 5.21 is the main feature of all the Markov processes. Hence *only the last occupied state can influence the probability for successive states of the system to be occupied and no memory of the preceding states remains.*

5.4.4 Equilibrium and Convergence of a Markov Chain

A peculiar feature of DTMC regards the equilibrium condition and the convergence for a long enough run of the process. It is crucial for a successful numerical inference, as will be discussed in the forthcoming Section 5.5.

Let $\mathbf{\Phi} = \{\Phi_0, \Phi_1, \ldots, \Phi_n\}$ be a DTMC with transition probability matrix $\mathbf{P}$. An initial probability distribution λ is called an *equilibrium* distribution (also *stationary*, or an *invariant* distribution) if it is preserved in time. That is

$$\lambda_j = \mathbb{P}(\Phi_0 = s_j) = \mathbb{P}(\Phi_1 = s_j) = \cdots = \mathbb{P}(\Phi_n = s_j)\,, \quad \forall j \in I \tag{5.22}$$

and we will denote it as a probability distribution π on $\mathcal{S}$. This means that the stochastic vector $\pi = (\pi_i)$ is an invariant vector for $\mathbf{P}$ (that is, an eigenvector with the eigenvalue 1) since it satisfies the eigenvalues equation $\pi\mathbf{P} = \pi$. If $\mathbf{P}$ is an irreducible matrix[9] it allows a unique equilibrium distribution at most (one or none). Moreover if $\mathbf{P}$ is also finite, so as for computationally treated systems, then it will always admit a unique equilibrium distribution. For instance, referring to the case illustrated in Section 5.4.2, we have that

[9] A matrix $\mathbf{R}$ is denoted as reducible if exists a matrix $\mathbf{B}$ for a change of base such that $\mathbf{B}^{-1}\mathbf{R}\mathbf{B}$ is a block-triangular matrix. A non reducible matrix is also called *irreducible*.

- if $\alpha + \beta > 0$, the systems has the only equilibrium distribution

$$\pi = \left(\frac{\beta}{\alpha + \beta}, \frac{\alpha}{\alpha + \beta} \right) ;$$

- if $\alpha = \beta = 0$ then $\mathbf{P} = \mathbf{I}$ and all (x, y) vectors will be invariant.

Convergence to equilibrium means that, as the time progresses, the Markov Chain "forgets" about its initial distribution λ. Clearly, this is related to the asymptotic properties of the n-step matrix $\mathbf{P}^n$ as $n \to \infty$. Consider again the case of a finite DTMC and suppose to have a $m \times m$ n-step finite transition matrix $\mathbf{P}^n$, converging to a *limiting matrix* $\mathbf{\Pi} = ||\pi_{ij}||$ such that

$$\lim_{n \to \infty} p_{ij}^{(n)} = \pi_{ij} \quad \forall i, j \in I \,, \tag{5.23}$$

where $p_{ij}^{(n)}$ are the transition probabilities given as entries of $\mathbf{P}^n$. Hence each row $\pi^{(i)}$ of $\mathbf{\Pi}$ is an equilibrium distribution, that is

$$\pi^{(i)} \mathbf{P} = \pi^{(i)} \quad \text{or} \quad \pi_{ij} = \sum_m \pi_{im} p_{mj} \,.$$

Besides if $\mathbf{P}$ is irreducible then a unique equilibrium distribution is ensured, so that all rows $\pi^{(i)}$ coincide, $\pi^{(1)} = \pi^{(2)} = \cdots = \pi^{(m)} = \pi$. Therefore

$$\lim_{n \to \infty} \mathbb{P}(\Phi_n = s_j) = \pi_j \quad \forall j \in I \,.$$

In other words, each row of the limiting matrix is a repetition of the same, unique vector representing the equilibrium distribution for $\mathbf{P}$. Hence the *DTMC forgets about its initial distribution λ and converges to a unique stationary distribution π.*

In conclusion, some definitions about the nature of states have to be introduced before the convergence condition can be provided. A state s_i is called:

- *periodic* with period k if the system transits on the same state after k time-steps. Conversely it is called *aperiodic.*

- *transient* if, assuming it is the initial state, there is a non-zero probability that the state will not be occupied again. Otherwise it is denoted as *recurrent* or *persistent.*

- *positive recurrent* if the expectation time for the state to be occupied again is finite. On the contrary, when it is infinite then the state is called *null recurrent.*

Such definitions can be subsequently extended to the transition matrix when its entries, that is the transition probabilities, are verifying the conditions for each type of state. The convergence to an equilibrium distribution is then ensured by the *Convergence Theorem.*

Theorem 1. *Given an irreducible, aperiodic and positive recurrent transition matrix $\mathbf{P}$, then $\mathbf{P}^n \to \mathbf{\Pi}$ as $n \to \infty$.*

When a DTMC converges to an equilibrium condition, it is typically denoted as a *relaxed* DTMC.

5.5 Numerical Inference

Once the main properties of a Markov Chain have been given, we can come back to the marginalization problem presented in Section 5.3.1. In this section we shall try to understand both how a free parameter of a given model can be inferred from a posterior PDF through numerical techniques and how a marginal PDF can be derived in the case of a multi-dimensional problem.

Markov Chain Monte Carlo algorithms are used to develop a sequence of points (which is called a *chain*) in the parameter space fixed by our model and whose density is proportional to the posterior probability density distribution $p(\xi \mid d, \mathcal{M})$ given by Eq. (5.10). In the Bayesian inference framework, the equilibrium distribution which a DTMC converges to, introduced in Section 5.4.4, is represented by such posterior probability distribution.

Chosen an algorithm capable of generating a Markov Chain with some particular features (among the most known we remember the Metropolis-Hastings algorithm [225, 147], which is briefly illustrated in Section 5.5.1 below, the Gibbs sampling [278] and the Hamiltonian Monte Carlo [145], the derivation of statistical moments for the expectation values obtained by means of Monte Carlo estimations becomes rather trivial. Consider a finite DTMC $\boldsymbol{\phi} = \{\phi_0, \phi_1, \ldots, \phi_{M-1}\}$, that is a Markov Chain with M different values for a given parameter ϕ, which might be, for instance, the longitude of a star spot on the star's photosphere. Then the posterior mean of the parameter ϕ, whose expectation value (or 1st momentum) from the posterior probability can be denoted as $\langle \cdot \rangle$, is given by

$$\langle \phi \rangle \approx \int p(\phi \mid d, \mathcal{M}) \phi d\phi = \frac{1}{M} \sum_{i=0}^{M-1} \phi_i \,, \tag{5.24}$$

where the equality with the mean of the sample from the MCMC follows because the samples ϕ_i are generated from the posterior by construction. In general, one can easily obtain the expectation value for any function $f(\phi)$ of the parameter by writing

$$\langle f(\phi) \rangle \approx \frac{1}{M} \sum_{i=0}^{M-1} f(\phi_i) \,. \tag{5.25}$$

Similarly to Eq. (5.24), the posterior variance (or 2nd momentum) of the parameter ϕ is given by

$$\sigma_\phi^2 \approx \int p(\phi \mid d, \mathcal{M}) \left(\phi - \langle \phi \rangle\right)^2 d\phi = \frac{1}{M-1} \sum_{i=0}^{M-1} \left(\phi_i - \langle \phi \rangle\right)^2 \,, \tag{5.26}$$

where the equality follows with the same arguments adopted for the expectation value. Further details about the statistical estimators from a marginal PDF and are presented in Appendix C.

In order to extend the discussion to a m-dimensional problem, we assume ϕ to be one of the m parameters ξ_j of the chosen model $\mathcal{M}$, represented by the parameters vector $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_m\}$. Hence, without losing in generality, we can put $\phi = \xi_1$ and write Eq. (5.14) as

$$p(\phi \mid d, \mathcal{M}) = \int_{\Omega_\mathcal{M}} p(\boldsymbol{\xi} \mid d, \mathcal{M}) d\xi_2 d\xi_3 \ldots d\xi_m \,, \tag{5.27}$$

with $p(\boldsymbol{\xi} \mid d, \mathcal{M}) = p(\phi, \xi_2, \ldots, \xi_m \mid d, \mathcal{M})$ the posterior PDF in a m-dimensional parameter space and $p(\phi \mid d, \mathcal{M})$ the one-dimensional marginal PDF of the parameter ϕ. $\Omega_\mathcal{M}$ represents again the parameter space of the free parameters involved. The left-hand side of Eq. (5.27) can be trivially computed through a Markov Chain process since, being its points part of the

parameter space itself, their density reflects the total posterior PDF. It suffices discretizing the range for the expected values of ϕ by dividing the interval of variation in a series of *bins*[10] and hence count the number of points of the Markov Chain that fall in each bin simply by ignoring the reference to all the remaining coordinates $\xi_2, \ldots, \xi_m$. Thus, assuming each parameter ξ_i has M_i different values $\xi_i^{(j)}$, with $j = 0, 1, \ldots, M_i - 1$ (i.e. its values are divided into a grid of M_i different bins), numerically the marginal PDF of the parameter ϕ can be computed as

$$p(\phi \mid d, \mathcal{M}) = \sum_{i=1}^{m-1} \sum_{j=0}^{M_i-1} p(\phi, \xi_2, \ldots, \xi_m \mid d, \mathcal{M}) \ , \tag{5.28}$$

where we note that the sum over all the values of the parameter $\xi_1 = \phi$, given for $i = 0$, has been skipped in order to marginalize over the remaining parameters.

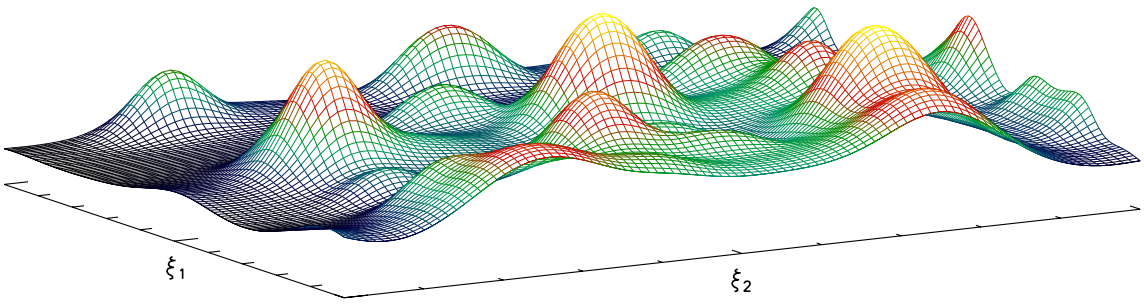

Figure 5.6: *Example of a posterior PDF depending on two free parameters ξ_1, ξ_2. As clearly visible, even in a low dimensional problem the plotted surface of the posterior probability may have many local maxima (multimodal distribution).*

Nonetheless, when dealing with high-dimensionality problems $m \geq 3$), particular attention must be paid in using MCMC methods for exploring and sampling the posterior distributions. In fact, the shape of the posterior PDF may be rather complex and potentially its integration may become a very tough issue. For example, in the case of a photometric star spots modeling [93, 117, 115, 118] when 7-8 (or more) spots are involved, the number of free parameters can be larger than a hundred, possibly leading to posterior PDFs that are sharply peaked with several local extremes. A probability distribution having many peaks (or conversely dimples) is called a *multimodal distribution*, since two or more modes[11] are coexisting in the same distribution (e.g. see Figure 5.6). Therefore, the probability for a Markov Chain to fall inside a local extreme and get stuck within itself becomes very high. In this case the exploration of the entire parameter space is likely to fail, yielding to big mistakes on parameter inference. For this reason, particular criteria for the exploration of posterior PDFs that show highly complex structures have been developed in the last decades. We shall describe one of the most known and used criteria in Section 5.5.1 below.

5.5.1 The Metropolis-Hastings algorithm

Nowadays many algorithms capable of generating a Markov Chain are available. In principle, their choice strongly depends on the features of the problem we want to solve. For the star spots

[10] For example, if ϕ is a longitude then we have to discretize the range $[0, 2\pi]$.

[11] In statistics, the mode is the value that occurs most frequently in either a dataset or a probability distribution. Modal values are usually referred to the most likely values given by a probability distribution.

modeling issue [93, 117, 115, 118], we will refer to and make use of the so called Metropolis-Hastings algorithm [225, 147]. In order to illustrate the mechanism of such algorithm we consider to have again set of m parameters identifying the chosen model and whose values can be defined by the m-dimensional parameters vector denoted as $\boldsymbol{\xi}$. Hence the MCMC approach builds a smart random walk inside the m-dimensional parameter space, saving the actual values of the parameters vector $\boldsymbol{\xi}$ at each time-step. At the successive time-step a new vector of actual parameter values will be then initialized, differing by a fixed displacement vector with respect to the precedent one (i.e. the parameters vector saved at the last point of our random walk). The new point of the random walk will be then accepted or rejected accordingly to a given criterion. The Metropolis-Hastings criterion asserts that a new step in a Markov Chain random walk is accepted if $p_{i+1}q_i/p_iq_{i+1} > u$ and conversely rejected, where u is random number between 0 and 1 ($u \in [0, 1]$) and q_i/q_{i+1} is the ratio of the probabilities for the new solution. In other words this means we always accept a new point in the random walk if it increases the fit quality of the model to the data (the probability value is higher by advancing in the random walk) otherwise it is *sometimes* (randomly) accepted if the fit becomes worse (that is when the probability value is lower). This technique allows the exploration of regions of the parameter space that would be otherwise inaccessible if the random walk falls into a local extreme. In computational analysis the application of such criterion is summarized by the so-called *acceptance*, a parameter expressed as the ration of the percentage of good steps (i.e. steps that increase the probability value of the next point) over the sum of good and bad steps together (i.e. steps that decrease the probability value of the next point). When using the Metropolis-Hastings algorithm a value of acceptance between 20 and 30 % ought to be preferred for a successful exploration of the parameter space, as derived by several applications of the criterion in the case of large numbers of free parameters employed (e.g. see Chapter 7).

5.6 Model Comparison

As announced already in the introduction of this chapter, we now focus to some other very interesting and useful features of Bayesian statistics in this second part of the discussion. In fact, it is important to stress that the parameters inference may sometimes fall into a degeneracy problem, that is when a framework of more than one possible model for explaining a given dataset, is obtained. This issue is usually known as the *model selection problem* and it rises whenever there are several competing theoretical models for the interpretation of the same dataset. In such case the Bayesian statistics offers another powerful tool, usually known in astrophysics as the *Bayesian Evidence*, which provides a formal way to evaluate the degree of reliability of each model in the light of the information related to the choice of the priors and the features of the dataset used. The model selection problem, which is typically wrongly assessed by means of the intuition level of the experimenter, can then therefore solved in an objective manner. The guiding principle of the Bayesian model selection states that the 'best' model is the one that strikes an optimum balance between quality of fit and predictivity[12]. In this part of the chapter we intend to provide a fulfilling discussion concerning the *complexity* notion of a model and the method used for its selection, together with the introduction of some analytical tools that allow for the development of the Bayesian evidence technique.

[12]The term predictivity suggests the degree of simplicity whereby a model is capable of predicting the results derived from observations.

5.6.1 Occam's Razor

The name *Occam's razor*, deriving from the english philosopher and Franciscan friar William of Ockham (c. 1285 - c. 1349), establishes that among two or more possible theories for resolving some problem the one able to provide an explanation in the simplest (or most economic) manner ought to be preferred. As argued above, the Bayesian evidence relies on this principle by assessing the complexity of a model in the light of its degree of simplicity for reproducing the observed phenomenon, though in general the chosen model will be the one representing the best trade-off between complexity and quality of the fit over the experimental data. Obviously, by increasing the number of free parameters involved the fit to the data will always be better (or at least equal), despite of the bulk of information we want to derive, which will consequently increase due to the introduction of less significative free parameters. Thus, when too many free parameters are introduced for explaining the observed data, an excess of unnecessary information is derived, which clearly compromises our level of comprehension[13]. For instance, when referring to the planetary motions one can consider the model comparison of the Ptolemaic model of epicycles versus the heliocentric model based on Newtonian gravity. Occam's razor would favor the latter because of its simplicity and ability to explain the planetary motions in a more economic fashion than the baroque construction of epicycles.

In the Bayesian mindset it appears pointless to reject a theory unless an alternative explanation of the event that fits the data better is provided. This means that, although data might be improbable, we cannot discard a model if no alternatives are given, unless it is impossible to yield observations for the model itself. Such aspect is in contrast with the frequentist goodness-of-fit tests (e.g. χ^2 minimum criterion, Maximum Likelihood Estimator, etc.). So for example, perturbations to the motion of Uranus led the French astronomer U. J. J. Leverrier and the English scholar J.C. Adams to formulate the prediction, based on Newtonian theory, that a further planet ought to exist beyond the orbit of Uranus. The discovery of Neptune in 1846 within 1 degree of the predicted position thus should strengthen our belief in the correctness of Newtonian gravity. However the change in the plausibility of Newton's theory following the discovery of Uranus crucially depends on the alternative we are considering. If the alternative theory is Einstein gravity, then obviously the two theories make the same predictions as far as the orbit of Uranus is concerned, hence their relative plausibility is unchanged by the discovery. Hence the alternative "Newton theory is false" is not useful in Bayesian model comparison, and we are forced to put on the table a more specific model than that before, assessing how much the new observation changes our relative degree of belief between an alternative theory and Newtonian gravity.

In the context of model comparison it is appropriate to think of a model $\mathcal{M}$ as a specification of a set of parameters ξ and of their prior distribution, $p(\xi \mid \mathcal{M})$. As shown below, it is the number of free parameters *and* their prior range that control the strength of the Occam's razor effect within the Bayesian model comparison. Models that have many free parameters ranging within a wide interval of values but that are not needed in the light of the data are hence penalized for their unwarranted complexity. Therefore, we can assert that *the prior choice ought to reflect the available parameter space under the model* $\mathcal{M}$*, independently of experimental constraints we might already be aware of.* This important statement justifies the fact we are trying to asses the economy (or simplicity) of the model itself by selecting priors on the basis of the model under consideration. Often these will take the form of a range of values that are

[13] We refer to the example mentioned in Section 5.5. In fact, when the number of free parameters increases, a longer computational time is required for mapping the posterior probability density function by means of the MCMC methods. This is caused by the amount of "wasted" parameter space, which clearly increases by increasing the number of unnecessary dimensions.

deemed *intuitively* plausible, or *nature*, as already discussed in Section 5.2.2. It is then clear that prior specification is inherent in the model comparison approach.

5.6.2 The Bayesian Evidence

The prime tool for model selection allowing for the evaluation of a model's performance in the light of the data, is represented by the *marginal likelihood* or *model likelihood*, best known as the *Beyesian evidence* in the field of astrophysics. As announced in Sections 5.2.1 and 5.3, the evidence is the normalization term appearing on the denominator of the right-hand side of Eq. (5.5), which can be expressed in the continuous parameter space $\Omega_{\mathcal{M}}$ by means of Eq. (5.6) as

$$\begin{aligned} p\left(d \mid \mathcal{M}\right) &\equiv \int_{\Omega_{\mathcal{M}}} p\left(d \mid \boldsymbol{\xi}, \mathcal{M}\right) \pi(\boldsymbol{\xi} \mid \mathcal{M}) d\boldsymbol{\xi} \\ &= \int_{\Omega_{\mathcal{M}}} \mathcal{L}(\boldsymbol{\xi}) \pi(\boldsymbol{\xi} \mid \mathcal{M}) d\boldsymbol{\xi} \end{aligned} \tag{5.29}$$

where again $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_m\}$ is the parameter vector and $\Omega_{\mathcal{M}}$ the m-dimensional parameter space. Thus the Bayesian evidence is the average of the likelihood function under the prior for a specific model choice, or equivalently the normalization factor to the volume of the posterior PDF[14]. By means of the Bayes' Theorem it is possible to invert the condition and thus evaluate the *model posterior probability*, given as

$$p\left(\mathcal{M} \mid d\right) = \frac{\pi\left(\mathcal{M}\right) p\left(d \mid \mathcal{M}\right)}{p(d)}\,, \tag{5.30}$$

where we can drop the constant $p\left(d\right)$, depending only on the data, since it is irrelevant within the model comparison problem. The term $\pi\left(\mathcal{M}\right)$ is the prior probability assigned to the model $\mathcal{M}$, which can also be expressed. Such model prior has not to be confused with the priors discussed in Section 5.2.2 and it is usually taken to be non-committal and equal to $1/N_{\mathcal{M}}$, with $N_{\mathcal{M}}$ the number of different models considered. When comparing two models, $\mathcal{M}_0$ versus $\mathcal{M}_1$, the ratio of their posterior probabilities (or *posterior odds*) represents the interesting quantity. By adopting Eq. (5.30), it is given as

$$\frac{p\left(\mathcal{M}_0 \mid d\right)}{p\left(\mathcal{M}_1 \mid d\right)} = B_{01} \frac{\pi(\mathcal{M}_0)}{\pi(\mathcal{M}_1)}\,, \tag{5.31}$$

where $\pi(\mathcal{M}_0) = \pi(\mathcal{M}_1) = 1/2$ and B_{01} is denoted as the *Bayes factor*, that is, the ratio of the models' evidences:

$$B_{01} \equiv \frac{p\left(d \mid \mathcal{M}_0\right)}{p\left(d \mid \mathcal{M}_1\right)}\,. \tag{5.32}$$

When $B_{01} > 1$ then we have an increase of the support in favor of the model $\mathcal{M}_0$ versus the model $\mathcal{M}_1$ given the observed data. Conversely, that is when $B_{01} < 1$, the model $\mathcal{M}_1$ will be the favorite one and a new Bayes factor given as the reciprocal of Eq. (5.32) can be then considered, that is

$$B_{10} \equiv \frac{p\left(d \mid \mathcal{M}_1\right)}{p\left(d \mid \mathcal{M}_0\right)}\,, \tag{5.33}$$

where we will clearly have $B_{10} > 1$ when the more complex model is preferred. As we will show in the forthcoming Chapter 6, it is sometimes more convenient to adopt the natural logarithm of the Bayes factor, which therefore reads

$$\ln B_{01} = \ln p(d \mid \mathcal{M}_0) - \ln p(d \mid \mathcal{M}_1)\,. \tag{5.34}$$

[14] In Chapter 6 a particular case of interest that simplifies the computation of the Bayesian evidence, is described.

Example

In order to gain some intuition about the application of Occam's razor through the Bayes factor, we can consider a clarifying example, described in [303], Section 4.2. Assume to have two competing models $\mathcal{M}_0$ and $\mathcal{M}_1$ which provide their explanation about a physical meaningful parameter ξ. In particular we suppose $\mathcal{M}_0$ to predict that $\xi = 0$ with no other free parameters, and $\mathcal{M}_1$ assigning ξ a Gaussian prior distribution with 0 mean and variance Σ^2 (hence showing one extra parameter with respect to the simpler model). Hence the two prior distributions will be

$$\begin{aligned} p\left(\xi \mid \mathcal{M}_0\right) &= \delta\left(\xi = 0\right) , \\ p\left(\xi \mid \mathcal{M}_1\right) &= \frac{1}{\sqrt{2\pi}\Sigma} \exp\left(-\frac{\xi^2}{2\Sigma^2}\right) , \end{aligned} \tag{5.35}$$

where $\delta\left(\xi = 0\right)$ is the Delta of Dirac peaked at $\xi = 0$. Assume we perform a measurement of ξ described by the following normal likelihood

$$\mathcal{L}\left(\xi\right) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{\left(\xi - \lambda\sigma\right)^2}{2\sigma^2}\right] \tag{5.36}$$

of variance σ^2, and whose maximum lies λ standard deviations away from 0, that is $|\xi_{max}/\sigma| = \lambda$. By adopting Eq. (5.29) and Eq. (5.32) the Bayes factor B_{01} can be formally written as

$$B_{01} = \frac{\int_{\Omega_{\mathcal{M}_0}} \mathcal{L}(\xi) p\left(\xi \mid \mathcal{M}_0\right) d\xi}{\int_{\Omega_{\mathcal{M}_1}} \mathcal{L}(\xi) p\left(\xi \mid \mathcal{M}_1\right) d\xi} , \tag{5.37}$$

and substituting the above functions, we obtain

$$\begin{aligned} B_{01} &= \frac{\sqrt{2\pi}\Sigma \int_{-\infty}^{\infty} \exp\left[-\frac{\left(\xi-\lambda\sigma\right)^2}{2\sigma^2}\right] \delta(\xi = 0) d\xi}{\int_{-\infty}^{\infty} \exp\left[-\frac{\left(\xi-\lambda\sigma\right)^2}{2\sigma^2}\right] \exp\left(-\frac{\xi^2}{2\Sigma^2}\right) d\xi} \\ &= \frac{\sqrt{2\pi}\Sigma}{\int_{-\infty}^{\infty} \exp\left[-\frac{\xi^2 - 2\lambda\sigma\xi}{2\sigma^2}\right] \exp\left(-\frac{\xi^2}{2\Sigma^2}\right) d\xi} , \end{aligned} \tag{5.38}$$

where the parameter spaces have been taken as the real axis, $\Omega_{\mathcal{M}_{0,1}} \equiv \mathbb{R}$, since the one-dimensional case is considered. By defining the quantity

$$\alpha = \sqrt{1 + \sigma^2/\Sigma^2} ,$$

one can rewrite the denominator as

$$\exp\left(\frac{\lambda^2}{2\alpha^2}\right) \int_{-\infty}^{\infty} \exp\left[-\frac{\left(\alpha\xi - \lambda\sigma/\alpha\right)^2}{2\sigma^2}\right] d\xi ,$$

whose integral can be solved by means of the Gaussian integral

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} ,$$

thus yielding to the final expression

$$B_{01} = \sqrt{1 + (\sigma/\Sigma)^{-2}} \exp\left[-\frac{\lambda^2}{2(1 + (\sigma/\Sigma)^2)}\right] . \tag{5.39}$$

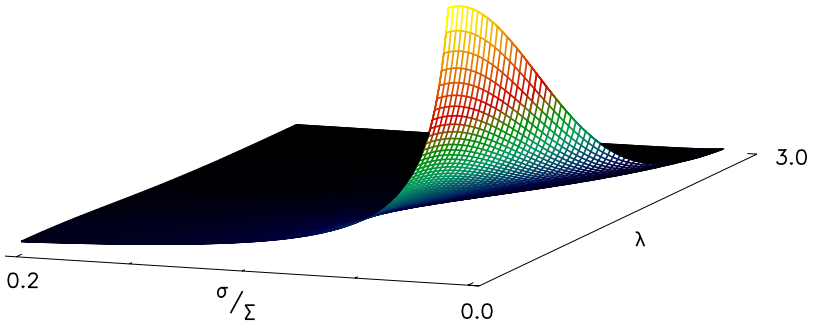

Figure 5.7: *The Bayes factor as a function of parameters λ and σ/Σ of the given example. For $\lambda \gg 1$ the exponential term of Eq. (5.39) dominates and the Bayes factor flattens to the value $B_{01} \ll 1$. Conversely, if $\lambda \lesssim 1$ and we have informative data ($\sigma/\Sigma \ll 1$) then $B_{01} \simeq \Sigma/\sigma$, otherwise if $\sigma/\Sigma \gg 1$ the Bayes factor $B_{01} \to 1$, i.e. model comparison is prior dominated.*

Table 5.1: *Empirical scale (so–called "Jeffreys' scale" [172]) for evaluating the strength of evidence when comparing two models, $\mathcal{M}_0$ versus $\mathcal{M}_1$. It is an empirically calibrated scale where odds of about 3 : 1, 12 : 1 and 150 : 1 correspond to weak, moderate and strong evidence respectively.*

$\lvert\ln B_{01}\rvert$	Odds	Strength of evidence
< 1.0	$\lesssim 3:1$	Inconclusive
1.0	$\sim 3:1$	Weak evidence
2.5	$\sim 12:1$	Moderate evidence
5.0	$\sim 150:1$	Strong evidence

For $\lambda \gg 1$, that is when the detection occurs at many sigma away from the value 0, the exponential term dominates and $B_{01} \ll 1$, favoring the model $\mathcal{M}_1$ as intuitively in agreement with the usual conclusion. If $\lambda \lesssim 1$ and $\sigma/\Sigma \ll 1$, meaning the likelihood is much more informative than the prior and in proximity of 0, then the prediction of the simpler model $\mathcal{M}_0$ is confirmed. This allows the Occam's razor term to dominate, with $B_{01} \simeq \Sigma/\sigma$, that is the evidence increases in favor of the simpler model proportionally to the volume of "wasted" parameter space. Lastly, when $\sigma/\Sigma \gg 1$ instead, i.e. when the likelihood is less informative than the prior, the Bayes factor $B_{01} \to 1$, hence the data have not changed our relative belief between the two models (see Figure 5.7).

Table 5.1 shows the natural logarithm of the Bayes factor against the Jeffreys' scale for the strength of the evidence [172]. It is an empirically calibrated scale where odds of about 3 : 1, 12 : 1 and 150 : 1 correspond to weak, moderate and strong evidence respectively. Another case with odds $\lesssim 3:1$ is also reported, leading to an inconclusive comparison. It appears from the natural logarithm of the Bayes factor that evidence accumulates slowly and that indeed moving up a level in the evidence strength scale requires about an order of magnitude more support then the level before. As we shall argue in Chapter 6, the natural logarithm of the Bayes factor provides great advantages for numerical computations.

Finally it is important to stress that while the parameter inference, discussed in Section 5.3, deals with the assessment of hypotheses within the chosen model, the Bayesian evidence *extends* (and does not substitute) the assessment in the light of the available data to the space of theoretical models. This feature is clearly represented by Eq. (5.31), which is the equivalent expression for models to Eq. (5.10).

5.6.3 The Bayes Factor

A wider overview concerning the use of the Bayes factor is here discussed in order to provide the reader with a deeper understanding of the model comparison problem[15]. We can refer again to the simple example mentioned in the above section and by considering the natural logarithm of Eq. (5.39) the Bayes factor can be expressed as

$$\ln B_{01} = \ln \sqrt{1 + (\sigma/\Sigma)^{-2}} - \frac{\lambda^2}{2\,(1 + (\sigma/\Sigma)^2)} \tag{5.40}$$

where λ gives once again the number of sigma away from $\xi = 0$ (it is also known as the *significance* of the measurement). Assuming the data are more informative than the prior, that is if $\sigma/\Sigma \ll 1$ we obtain

$$\ln B_{01} \approx \ln (\Sigma/\sigma) - \lambda^2/2, \tag{5.41}$$

where the first term on the left-hand side is approximately the logarithm of the ratio of the prior to posterior volume or the "wasted" parameter space, as already announced. Thus, on one hand the first term can be interpreted as the constraining power of the information content of the data, which gives the factor of reduction of the parameter space volume in going from the prior to the posterior. It is obviously positive for informative data, i.e. the likelihood is more sharply peaked than the prior. On the other hand, the second term appearing in Eq. (5.41) is always negative and it favors the more complex model if the measurement gives a result many sigmas away from the prediction (i.e., for $\lambda \gg 1$).

By taking the reciprocal of B_{01}, given by Eq. (5.33) of the previous section, we obtain the Bayes factor referred to the more complex model, B_{10}. Therefore, its logarithm will be

$$\ln B_{10} = -\ln \sqrt{1 + (\sigma/\Sigma)^{-2}} + \frac{\lambda^2}{2\,(1 + (\sigma/\Sigma)^2)}. \tag{5.42}$$

One can refer to the Jeffreys' scale of strength reported in Table 5.1 also when B_{10} is taken into account. Therefore, the strength of evidence will be related to the more complex model $\mathcal{M}_1$.

For a deeper investigation of the features of the Bayesian model comparison we can finally analyze the behavior of the Bayes factor with respect to the significance and the volume of "wasted" parameter space. Following the example described in [303], we define the quantity $I_{10} \equiv \log_{10}(\Sigma/\sigma)$, known as the *Information gain*, because a base 10 logarithm is closer to our intuition being the order of magnitude of our information increase. By plotting Eq. (5.41) and Eq. (5.42) as functions of λ (significance of the fit) and I_{10} (the state of knowledge acquired by inferring our model prediction with the information provided by the data) for the odds listed in the Jeffreys' scale of strength of Table 5.1, we obtain a clear contouring of the different significance regions, shown in Figure 5.8. For values of $I_{10} > 0$ we have a parameter space dominated by data information, since $\Sigma/\sigma > 1$. The orange shaded region represents the case of simpler model favored, where the values $|\ln B_{01}| = 1.0, 2.5, 5.0$ for contours have been plotted. For moderately informative data, $I_{10} \sim 1 - 2$ the measured mean has to lie at least about 4σ away from 0 in order to robustly disfavor the simpler model (i.e., $\lambda \gtrsim 4$). On the other side, for $\lambda \lesssim 3$ and highly informative data, $I_{10} > 2$, the simpler model furnishes the best prediction for our measurement. Hence, in general, *a large information content favors the simpler model because Occam's razor penalizes the large volume of "wasted" parameter space of the extended model*, being $\Sigma \gg \sigma$. When $\lambda \gtrsim 4$ another free parameter is required for a better prediction of the result, since we are obtaining a measure many sigma away from the predicted value 0. Then

[15] We suggest the reader to refer to [173], for a detailed discussion about the use of the Bayes factor.

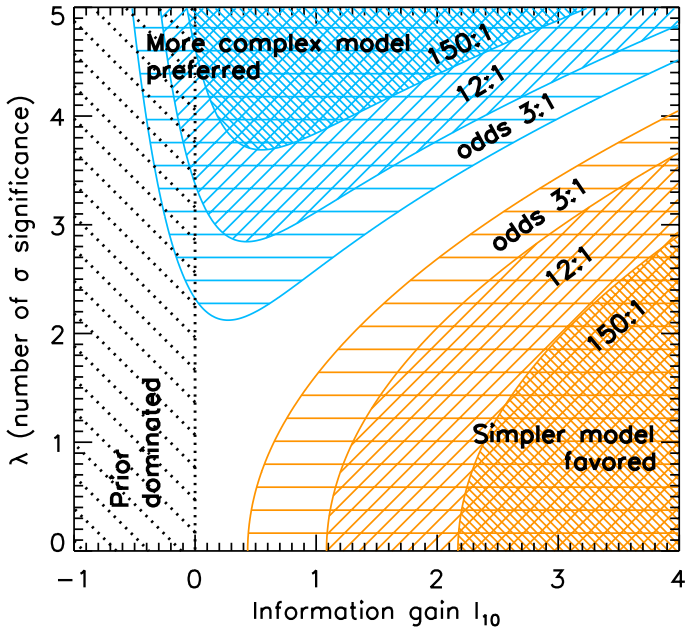

Figure 5.8: *The Bayes factors* $|\ln B_{01}|$ *(orange) and* $|\ln B_{10}|$ *(blue) as functions of significance* λ *and information gain* I_{10} *of the example shown in Section 5.6.2. Inconclusive evidence is shown (white region) so as for prior dominated evidence (hatched region). Contours for values of* $1.0, 2.5, 5.0$ *of the Jeffreys' scale of strength for both* B_{01} *and* B_{10} *are also marked.*

the more complex model is favored, as shown by the blue shaded region with contours plotted at the values $|\ln B_{10}| = 1.0, 2.5, 5.0$. The white region represents an inconclusive evidence, i.e. odds $\lesssim 3:1$, while the hatched region in the left side of the plot, occurring at $I_{10} \leq 0$, shows a prior dominated evidence, meaning that no more information than the one contained in the prior is provided by data. Furthermore an useful property of Figure 5.8 allows to quantify the impact of a change of prior within the analysis. In fact, a different choice of prior width (i.e. Σ) amounts to a *horizontal shift* across the plot, at least as long as $I_{10} > 0$. Picking more restrictive priors (lower Σ, reflecting more predictive theoretical models) corresponds to shift the result of the model comparison toward the left side of the plot, resulting in an inconclusive result or a prior dominated outcome.

Bayesian model comparison is usually *conservative* when it admits a new quantity in our model, even in the case when the prior is chosen incorrectly. Consider the following two possibilities:

- If the prior range is too small, the model comparison result will be non-committal (white region in Figure 5.8), or even prior dominated (hatched region, where the posterior is dominated by the prior). Hence in this case we have to hold judgement until better data come along.

- Too wide a prior will instead unduly favor the simpler model (orange, shaded regions). However, as new, better data come along the result will move to the right (for a fixed prior width, as the likelihood becomes narrower) but eventually also upwards, towards a larger number of sigma significance, if the true model really has a non–zero extra parameter. Eventually, our initial "poor" prior choice will be overridden as the number of sigma becomes large enough to take the result into the blue, shaded region.

In both cases the result of the model comparison will at least override the "wrong" prior choice

exactly as it happens for parameter inference, as long as better datasets can be provided (see the example given in Section 5.2.2).

5.6.4 Model Complexity

As widely argued already, the advantage of the Bayesian model selection approach is that the Bayesian evidence tells us whether the increased "complexity" of a model (i.e. a model with more free parameters) is justified by the data. However, sometimes the number of free parameters is not a satisfying description of this concept since it might be reduced by the data we are using under particular conditions, and a more refined definition of *model complexity* is desirable to be introduced. For instance, consider we are trying to measure a periodic signal in a time series and assume we have a model to fit the data that looks like

$$f(t) = A\left[1 + \beta\cos(t+\delta)\right], \tag{5.43}$$

where A, β, δ are free parameters we wish to constrain. Nevertheless if $\beta \ll 1$ and the noise is large compared to β, then the oscillatory term is left unconstrained by the data, and only the normalization term A could effectively be measured. Thus the parameters β, δ should not amount among the total number of free parameters as they cannot be constrained given the data we have, hence the effective model complexity is closer to 1 rather than to 3. From this example it follows that the very notion of *free parameter* is not absolute, but it depends on both what our expectations are under the model, i.e. on the prior, and on the constraining power of the data at hand.

For the sake of completeness, it is therefore plausible to define an appropriate measure of complexity[16] taking into account the number of parameters that the data can support. For a given model $\mathcal{M}$ and a vector of parameters $\boldsymbol{\xi}$ formalizing model's hypotheses, consider first a new notion of the information gain obtained when upgrading the prior to the posterior, as measured by the Kullback-Leibler (KL) divergence [197] between the posterior, p, and the prior, π:

$$D_{KL}(p,\pi) \equiv \int_{\Omega_{\mathcal{M}}} p(\boldsymbol{\xi} \mid d, \mathcal{M}) \ln\left[\frac{p(\boldsymbol{\xi} \mid d, \mathcal{M})}{\pi(\boldsymbol{\xi} \mid \mathcal{M})}\right] d\boldsymbol{\xi}\,. \tag{5.44}$$

In virtue of the Bayes' Theorem given by Eq. (5.10), $p(\boldsymbol{\xi} \mid d, \mathcal{M}) = \mathcal{L}(\boldsymbol{\xi})\pi(\boldsymbol{\xi} \mid \mathcal{M})/p(d \mid \mathcal{M})$ and since

$$\int_{\Omega_{\mathcal{M}}} p(\boldsymbol{\xi} \mid d, \mathcal{M}) d\boldsymbol{\xi} = 1\,, \tag{5.45}$$

being the integral of the posterior PDF, the definition of the KL divergence can be rewritten as

$$D_{KL}(p,\pi) = -\ln p(d \mid \mathcal{M}) + \int_{\Omega_{\mathcal{M}}} p(\boldsymbol{\xi} \mid d, \mathcal{M}) \ln \mathcal{L}(\boldsymbol{\xi}) d\boldsymbol{\xi}\,, \tag{5.46}$$

that is, the sum between the negative natural logarithm of the evidence and the expectation value of the log-likelihood under the posterior.

Example 1

To gain some feeling about the meaning of the KL divergence let us consider a one-dimensional case for the measure of a physical meaningful parameter ζ. As in a previous example we assume

[16] We adopt the notion of *Bayesian complexity* defined in [282].

a Gaussian prior around $\zeta = 0$, of variance Σ^2 and a Gaussian likelihood centered at ζ_{max} and variance σ^2:

$$\pi(\zeta \mid \mathcal{M}) = \frac{1}{\sqrt{2\pi}\Sigma} \exp\left(-\frac{\zeta^2}{2\Sigma^2}\right), \tag{5.47}$$

$$\mathcal{L}(\zeta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(\zeta - \zeta_{max})^2}{2\sigma^2}\right]. \tag{5.48}$$

Since $p(\xi \mid d, \mathcal{M}) = \mathcal{L}(\xi)\pi(\xi \mid \mathcal{M})/p(d \mid \mathcal{M})$ Eq. (5.46) reads

$$D_{KL}(p, \pi) = -\ln p(d \mid \mathcal{M}) + \frac{1}{p(d \mid \mathcal{M})} \int_{-\infty}^{\infty} \mathcal{L}(\zeta)\pi(\zeta \mid \mathcal{M}) \ln \mathcal{L}(\zeta) d\zeta\,, \tag{5.49}$$

where $\Omega_{\mathcal{M}} \equiv \mathbb{R}$ is the real axis, and the evidence $p(d \mid \mathcal{M})$ given by Eq. (5.29) gives

$$p\left(d \mid \mathcal{M}\right) = \int_{-\infty}^{\infty} \mathcal{L}(\zeta)\pi\left(\zeta \mid \mathcal{M}\right) d\zeta\,, \tag{5.50}$$

which can be solved similarly to the example described in Section 5.6.2, hence obtaining

$$p\left(d \mid \mathcal{M}\right) = \frac{1}{\sqrt{2\pi}\Sigma\alpha} \exp\left[\frac{\zeta_{max}^2(1-\alpha^2)}{2\alpha^2\sigma^2}\right], \tag{5.51}$$

with $\alpha = \sqrt{1 + \sigma^2/\Sigma^2}$. The integral on the right-hand side of Eq. (5.49) can be evaluated as the sum of three terms:

$$\int_{-\infty}^{\infty} \mathcal{L}(\zeta)\pi(\zeta \mid \mathcal{M}) \ln \mathcal{L}(\zeta) d\zeta = \left(\ln \frac{1}{\sqrt{2\pi}\sigma} - \frac{\zeta_{max}^2}{2\sigma^2}\right) p(d \mid \mathcal{M}) + \frac{\zeta_{max}}{\sigma^2} I_1 - \frac{I_2}{2\sigma^2} \tag{5.52}$$

where we defined the integrals I_1 and I_2 as proportional to the mean of ζ, $<\zeta>$, and the mean of ζ^2, $<\zeta^2>$, under the posterior probability, namely

$$I_1 \equiv \int_{-\infty}^{\infty} \mathcal{L}(\zeta)\pi(\zeta \mid \mathcal{M})\zeta d\zeta \quad \text{and} \quad I_2 \equiv \int_{-\infty}^{\infty} \mathcal{L}(\zeta)\pi(\zeta \mid \mathcal{M})\zeta^2 d\zeta\,. \tag{5.53}$$

We can now proceed in evaluating the integrals I_1 and I_2 by means of the integration by parts of the Gaussian Integral, that is

$$\int_{-\infty}^{\infty} x e^{-x^2} dx = 0 \qquad \text{and} \qquad \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \sqrt{\pi}/2\,, \tag{5.54}$$

being the integral on the left side equal to 0 since its integrand is the product of an even to an odd function. We obtain after a short calculation

$$I_1 = \frac{\zeta_{max}}{\sqrt{2\pi}\Sigma\alpha^3} \exp\left[\frac{\zeta_{max}^2(1-\alpha^2)}{2\alpha^2\sigma^2}\right], \tag{5.55}$$

$$I_2 = \left(\frac{\sigma^2}{\sqrt{2\pi}\Sigma\alpha^3} + \frac{\zeta_{max}^2}{\sqrt{2\pi}\Sigma\alpha^4}\right) \exp\left[\frac{\zeta_{max}^2(1-\alpha^2)}{2\alpha^2\sigma^2}\right]. \tag{5.56}$$

Finally Eq. (5.49) becomes

$$D_{KL}(p, \pi) = -\frac{1}{2} - \ln\frac{\sigma}{\Sigma} + \frac{1}{2}\left[\left(\frac{\sigma}{\Sigma}\right)^2 \left(\frac{\zeta_{max}^2}{\sigma^2} - 1\right)\right], \tag{5.57}$$

in which the second term on the right-hand side gives the reduction in volume in going from the prior to the posterior PDF. For informative data, $\sigma/\Sigma \ll 1$, this term is positive and grows as the logarithm of the volume ratio. Conversely, in the same regime the third term is small, unless $\zeta_{max}/\sigma \gg 1$, i.e. the maximum likelihood is many standard deviations away from the expected prevision under the prior. In this case we say the maximum likelihood is "surprising" since it is far from what our prior led us to expect. Therefore the KL divergence can be interpreted as a summary of the amount of information, or "surprise", contained in the data, or equivalently the entropy generated in the parameter space when going from prior to posterior distributions.

Let us now assume that we have a parameter ξ, formalizing the hypotheses of the model $\mathcal{M}$, that is a normal random variable, i.e. it is χ^2 distributed with one degree of freedom. Hence we can define an effective χ^2 through the likelihood (see also Appendix C) as $\mathcal{L}(\xi) \equiv \exp\left[-\chi^2(\xi)/2\right]$. Then Eq. (5.46) gives

$$D_{KL}(p,\pi) = -\frac{1}{2}\overline{\chi^2(\xi)} - \ln p(d \mid \mathcal{M})\,, \tag{5.58}$$

where the bar denotes a mean taken over the posterior distribution, namely

$$\overline{\chi^2(\xi)} = \int_{\Omega_\mathcal{M}} p(\xi \mid d, \mathcal{M})\chi^2(\xi)d\xi\,. \tag{5.59}$$

The posterior average of the effective chi-square can be easily evaluated through MCMC techniques (see Section 5.4) and it represents the "expected surprise". We then define the "estimated surprise" in the data after we have actually fitted the model parameters, as the following quantity

$$\widehat{D_{KL}} \equiv -\frac{1}{2}\chi^2(\hat{\xi}) - \ln p(d \mid \mathcal{M})\,, \tag{5.60}$$

where the first term on the right-hand side is the effective chi-square at the estimated value of the parameters, indicated by a hat. This will usually be the posterior mean of the parameters, as defined by Eq. (5.24), but other possible estimators are the maximum likelihood point or the posterior median, depending on the details. Thus, $\widehat{D_{KL}}$ represents the KL divergence evaluated at the single point $\hat{\xi}$. In general the *Bayesian complexity* can be defined by

$$\mathcal{C}_b \equiv -2\left(D_{KL}(p,\pi) - \widehat{D_{KL}}\right)\,, \tag{5.61}$$

in which, by taking the difference, we compare the effective information gain, namely the KL divergence, to the maximum information gain we can expect under the model, that is $\widehat{D_{KL}}$. For a given likelihood $\mathcal{L}(\xi)$, we have from Eq. (5.46) that

$$\mathcal{C}_b = -2\int_{\Omega_\mathcal{M}} p(\xi \mid d, \mathcal{M}) \ln \mathcal{L}(\xi)d\xi + 2\ln \mathcal{L}(\hat{\xi})\,, \tag{5.62}$$

and substituting Eqs. (5.58) and (5.60) we have

$$\mathcal{C}_b = \overline{\chi^2(\xi)} - \chi^2(\hat{\xi})\,, \tag{5.63}$$

where the Bayesian evidence term $p(d \mid \mathcal{M})$ disappears as it does not depend on the parameters (i.e. it has no weight within the complexity). Therefore we can assert that *the Bayesian complexity gives the effective number of parameters as a measure of the constraining power of the data as compared to the predictivity of the model, namely the prior.* Hence $\mathcal{C}_b$ depends both on the data and on the prior available parameter space.

Example 2

Let consider the likelihood and prior introduced in the previous example. After a short calculation [198], employing the results already obtained, we derive the expression for the Bayesian complexity

$$\mathcal{C}_b = \frac{1}{1 + (\sigma/\Sigma)^2} \,. \tag{5.64}$$

Thus, on one hand if $\sigma/\Sigma \ll 1$, that is when the data are more informative than the prior, then $\mathcal{C}_b \approx 1$ and the model has one effective, well constrained, parameter. On the other hand, when the data are not informative, i.e. $\sigma/\Sigma \gg 1$, then the experiment is not providing any constrains on our model belief and therefore the Bayesian complexity will be $\mathcal{C}_b \to 0$.

Bayesian complexity can be an useful diagnostic tool when the evidence for two competing models is about the same. In fact, since the evidence does not take into account the effective number of parameters and hence does not penalize the parameters that are unmeasured within a model, the evidence alone is not able to establish if we are in the situation where the extra parameters are unconstrained and hence irrelevant (e.g. for $\beta \ll 1$ in the model of Eq. (5.43)), or if they improve the quality-of-fit just enough to offset the Occam's razor penality term and produce the same evidence of the simpler model. The Bayesian complexity breaks this degeneracy in the evidence allowing to distinguish between the two following cases:

- $p(d \mid \mathcal{M}_0) \approx p(d \mid \mathcal{M}_1)$ and $\mathcal{C}_b(\mathcal{M}_1) > \mathcal{C}_b(\mathcal{M}_0)$: the quality of the data is sufficient to measure the additional parameters of the more complicated model ($\mathcal{M}_1$), but they do not improve its evidence by much. Since the Occam's razor rules the selection in model comparison, the model $\mathcal{M}_0$ with less parameters ought to be preferred.
- $p(d \mid \mathcal{M}_0) \approx p(d \mid \mathcal{M}_1)$ and $\mathcal{C}_b(\mathcal{M}_1) \approx \mathcal{C}_b(\mathcal{M}_0)$: both models have a comparable evidence and the number of effective parameters is about the same. In this case the data is not good enough to measure the additional parameters of the more complicated model and we cannot draw any conclusion as to whether the extra parameter is needed.

5.6.5 Information Criteria

As a last point we shall introduce some practical usages of the Bayesian complexity, which can be enclosed in the adoption of the so called information criteria. In fact, sometimes it might be useful to employ methods that aim at an approximate model selection under some simplifying assumptions with a given penalty term for more complicated models, which replaces the Occam's razor term coming from the different prior volumes in the Bayesian evidence. While this is an obviously appealing feature, on closer examination it has the drawback of being meaningful only in fairly specific cases, which are not always met in astrophysical applications. In particular, it can be argued that the Bayesian evidence tool (ideally coupled with an analysis of the Bayesian complexity) ought to be preferred in model selection because it is precisely the lack of predictivity of more complex models, as enclosed in the physically motivated range of the prior, that ought to penalize them.

With this *caveat* in mind, we list below three types of information criteria that have been widely used in several astrophysical contexts.

1. **Akaike Information Criterion (AIC):** Introduced by Akaike [2], it is an essentially frequentist criterion that sets the penalty term equal to twice the number of free parameters in the model, denoted as k:

$$\mathrm{AIC} \equiv -2 \ln \mathcal{L}_{\max} + 2k \tag{5.65}$$

where $\mathcal{L}_{\max} \equiv p(d \mid \xi_{\max}, \mathcal{M})$ is the maximum likelihood value. The derivation of the AIC follows from an approximate minimization of the KL divergence between the true model distribution and the distribution being fitted to the data.

2. **Bayesian Information Criterion (BIC):** It is also known as "Schwarz Information Criterion" (from the name of its proposer [273]), the BIC follows from a Gaussian approximation to the Bayesian evidence in the limit of large sample size:

$$\mathrm{BIC} \equiv -2 \ln \mathcal{L}_{\max} + k \ln N \tag{5.66}$$

where k is again the number of fitted parameters and N is the number of data points. The best model is once more the one that minimizes the BIC.

3. **Deviance Information Criterion (DIC):** Introduced by Spiegelhalter [282], the DIC can be written as

$$\mathrm{DIC} \equiv -2\widehat{D_{KL}} + 2\mathcal{C}_b. \tag{5.67}$$

In this form, the DIC is reminiscent of the AIC, with the $\ln \mathcal{L}_{\max}$ replaced by the estimated KL divergence and the number of free parameters by the effective number of parameters, $\mathcal{C}_b$, from Eq. (5.61). Indeed, for well-constrained parameters, the AIC is recovered from Eq. (5.67), but the DIC has the advantage of accounting for unconstrained directions in the parameters space.

The information criteria ought to be interpreted with care when applied to real situations. By comparing Eq. (5.66) with Eq. (5.65) it is noticeable that for $N > 7$ the BIC penalizes models with more free parameters more harshly than the AIC. Furthermore, both criteria penalize extra parameters regardless of whether they are constrained by data or not, unlike the Bayesian evidence. The BIC represents a simple and valuable approximation to Bayesian evidence in the light of highly informative data and of large sample sizes, as we will discuss in Chapter 6 for a practical astrophysical study.

Corsaro, E., et al., *A Bayesian approach to scaling relations for amplitudes of solar-like oscillations in Kepler stars*, 2012c, arXiv:1212.1156v1 [astro-ph.SR] [88]

6 Amplitude Scaling Relations

The NASA's *Kepler* spacecraft is providing a very large amount of high quality light curves, with a very high duty cycle, as already discussed in Chapter 3. These will become longer in the *Kepler* extended mission phase [295]. The asteroseismic studies done with *Kepler* have recently led to the birth of the ensemble asteroseismology [71], which is showing great potential for a thorough understanding of stellar evolution theory. The success of ensemble asteroseismology relies mainly on adopting scaling relations: generally simple empirical laws that allow for the derivation of fundamental stellar properties for stars different from the Sun by scaling their asteroseismic quantities from the solar values (see also Chapter 3, Section 3.4).

Among the most challenging asteroseismic quantities to measure and model one can certainly mention the oscillation amplitude. This is due to both the difficulty in estimating the background level in the power spectrum and the rather complicated physics involved in the driving and damping mechanisms of the modes (e.g. see [190, 191]). Different scaling relations aimed at predicting amplitudes by scaling from the Sun's values have been derived and discussed by several authors, both theoretically [187, 158, 159, 160, 269, 33, 188, 271] and observationally [291, 167, 244]. A variety of amplitude scaling relations has been used extensively in literature in both ensemble studies [290, 72, 310, 167, 244, 36] and detailed analyses of single stars from main sequence to the subgiant phase of evolution [48, 270, 166, 216, 63, 86].

The underlying physical meaning of these various amplitude scaling relations is still not properly understood (e.g. see the discussion by [269, 310, 167, 271, 36]). Testing them with observational data is vital for assessing the competing relations and for improving our understanding of stellar oscillations, i.e. the driving and damping mechanisms that produce the observed amplitudes (see also [72]). In this context, Bayesian methods can be of great use (see e.g. [303]) because they allow us to measure physical quantities of interest in a rigorous manner. Moreover, Bayesian statistics provides an efficient solution to the problem of model comparison, which is the most important feature of the Bayesian approach (see also [37, 144, 141]).

In the present paper we analyze amplitude measurements of a sample of 1640 stars observed with *Kepler*, together with temperature estimates derived from SDSS photometry, which we introduce in Section 6.1. In Section 6.2 we discuss the different scaling relations for predicting the oscillation amplitude per radial mode. The results obtained from a Bayesian parameter estimation for the different scaling relations are shown in Section 6.3 and the model comparison is presented in Section 6.4. Lastly, discussion and conclusions about the results of our analysis are drawn in Section 6.5 and 6.6, respectively.

6.1 Observations and Data

We use amplitude measurements and their uncertainties, obtained by [167] for a sample of 1673 stars spanning from main sequence (MS) to red giant stars (RGs) observed with *Kepler*

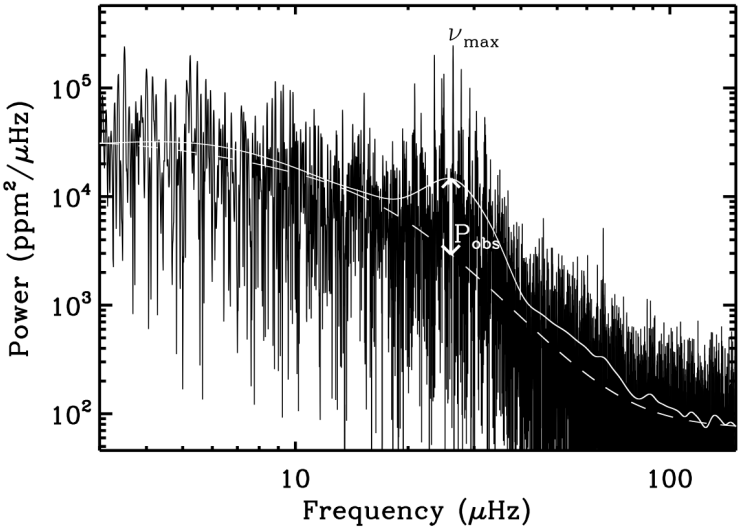

Figure 6.1: *Example PSD of a typical star studied by [291] where the method for measuring oscillation amplitudes described by [190, 191] was used. The smoothed spectrum (solid white line) and the fit to the stellar granulation background (dashed line) are shown. The oscillation power, P_{obs}, is evaluated at the frequency of maximum power, ν_{max}. Image copied from [291].*

in short cadence (SC; mostly MS stars but also some subgiants and low luminosity RGs) and long cadence (LC; all RGs) modes ([127, 174], respectively). Most of the 542 stars observed in SC have photometry for one month, while the 1131 stars observed in LC have light curves spanning from *Kepler*'s observing quarters 0 to 6. All amplitudes were derived according to the method described by [190, 191] (see also Chapter 2, Section 2.3.6 for more practical details about this technique), where the background level was fitted by using Harvey models [146], see also Figure 6.1 for an example of the method applied on the PSD of a typical red giant star studied by [291]. The amplitudes are so derived per radial mode (see [168] for more details). Values of the frequency of maximum power, ν_{max}, the large frequency separation, $\Delta\nu$, and their uncertainties for all the stars were also taken from [167], who used the SYD pipeline [168].

It is important to have accurate temperature estimates for the stars of our sample. Unfortunately, those provided by the KIC [60] are known to suffer from significant systematic effects (see [256] for a detailed discussion of the problem). We used revised T_{eff} derived by [256] for a total of 161977 KIC stars from Sloan Digital Sky Survey (SDSS) *griz* filters, which were corrected using temperature estimates from infrared flux method (IRFM) $(J - K_s)$ color index for hot stars (e.g. see [66]). The revised effective temperatures are available in the online catalog [257].

By cross-matching the stars of our sample with the temperature estimates provided by [256], we arrived at a final sample of 1640 stars with an accurate T_{eff} (1111 observed in LC and 529 in SC), which will be used for our investigation. Total uncertainties on temperature, as derived by [256], include both random and systematic contributions. The amplitudes of the final sample are plotted against ν_{max} and $\Delta\nu$ in in Figure 6.2 (top left and right panels, respectively). The bottom panel shows an asteroseismic HR diagram of our sample of stars (amplitudes against T_{eff} from [256]), similar to the one introduced by [290]. 1σ error bars are overlaid on both quantities for each panel. The average relative uncertainty in amplitude is $\langle\sigma_A/A\rangle = 9.2\,\%$ for the entire sample, and $\langle\sigma_A/A\rangle_{SC} = 11.7\,\%$ and $\langle\sigma_A/A\rangle_{LC} = 8.1\,\%$, for SC and LC targets, respectively.

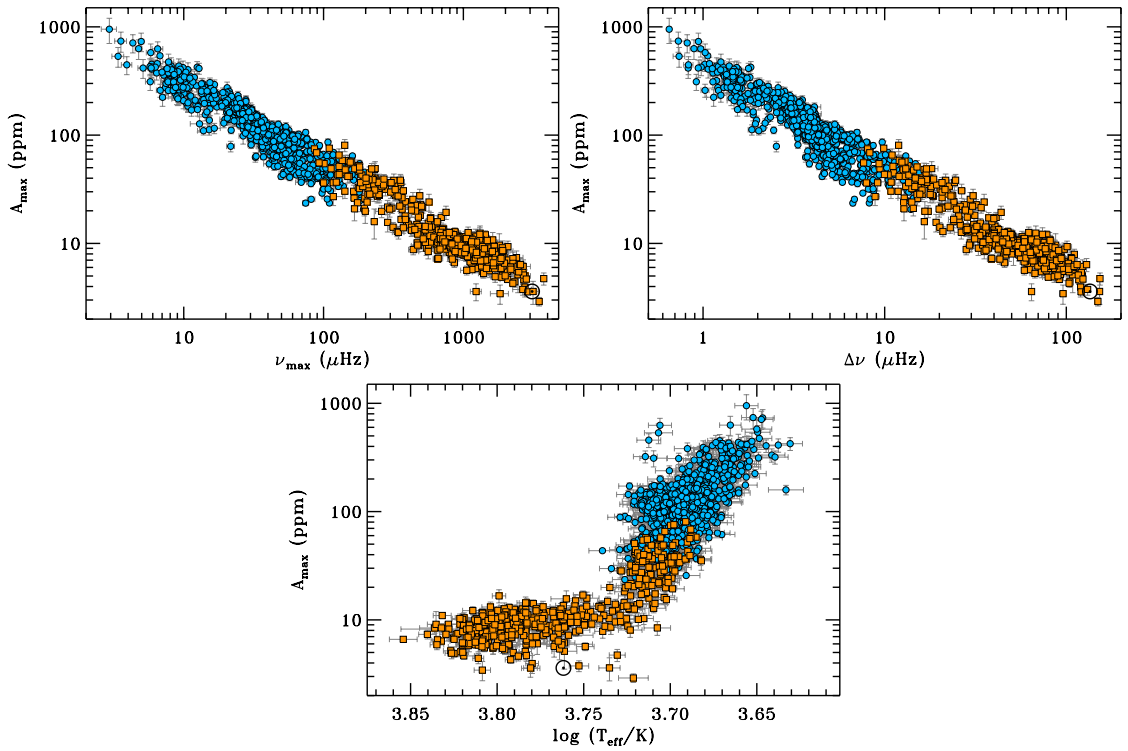

Figure 6.2: *Oscillation amplitudes for 1640 stars observed with Kepler in SC (orange squares) and LC (blue circles) modes and plotted against the frequency of maximum power $\nu_{\max}$ (top left) and the large frequency separation $\Delta\nu$ (top right) of the stars in a log-log scale. Amplitudes against the effective temperature $T_{\rm eff}$ are shown in the bottom panel, representing an asteroseismic HR diagram for our sample of stars. 1σ error bars are shown on both quantities for all the plots. The Sun is shown with its usual symbol ($\odot$).*

6.2 Amplitude scaling relations

Several scaling relations for oscillation amplitudes have been proposed so far. We will briefly introduce them in the following (see [167] for further discussion).

6.2.1 The L/M scaling relation

The first scaling relation for amplitudes was introduced by [187] for radial velocities, based on theoretical models by [74]. It is given by

$$v_{\rm osc} \propto \left(\frac{L}{M}\right)^{s}, \tag{6.1}$$

where $v_{\rm osc}$ represents the prediction for the amplitude in radial velocity, L is the luminosity and M the mass of the star, and $s = 1$ (see also Chapter 2, Section 2.4 were it was used for predicting the oscillation amplitude in β Aquilae). [187] also showed that the corresponding photometric amplitude A_λ, observed at a wavelength λ, is related to $v_{\rm osc}$ by

$$A_\lambda \propto \frac{v_{\rm osc}}{\lambda T_{\rm eff}^{r}}. \tag{6.2}$$

For adiabatic oscillations, the exponent r is 1.5 (see also [160]), but [187] found the observed value for classical p-mode pulsators to be 2.0 (see also [269]). By combining the two equations,

one obtains

$$A_\lambda \propto \left(\frac{L}{M}\right)^s \frac{1}{\lambda T_{\rm eff}^r}\,. \tag{6.3}$$

We are interested in a sample of stars observed with *Kepler*, whose bandpass has a central wavelength $\lambda = 650\,\mathrm{nm}$. By scaling Eq. (6.3) to our Sun, we have

$$\frac{A_\lambda}{A_{\lambda,\odot}} = \left(\frac{L/L_\odot}{M/M_\odot}\right)^s \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{-r}\,, \tag{6.4}$$

where $A_{650,\odot} = 3.98\,\mathrm{ppm}$ is the Sun's photometric amplitude observed at the *Kepler* wavelength (e.g. see [291]). The exponent s has been examined both theoretically (e.g. see [158, 159, 269]) and observationally [126, 104, 290, 310, 19], and found to be $0.7 < s < 1.5$. The exponent r has usually been chosen to be 2.0 (e.g. see the discussion by [291]), which implies that solar-like oscillations are not fully adiabatic. This was also shown in the case of *CoRoT* red giant stars by [271]. Nonetheless, some authors (e.g. see [229, 241]) have chosen to adopt $r = 1.5$ (see also the discussion by [188]).

We can also use the results introduced by [58], which suggest that the frequency of maximum power scales with the cut-off frequency of the star. Hence, $\nu_{\rm max} \propto g/\sqrt{T_{\rm eff}}$, and by considering that L/M scales as $T_{\rm eff}^4/g$, one obtains

$$\frac{L}{M} \propto \frac{T_{\rm eff}^{3.5}}{\nu_{\rm max}}\,. \tag{6.5}$$

Thus, Eq. (6.4) can be rewritten as

$$\frac{A_\lambda}{A_{\lambda,\odot}} = \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{-s} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{3.5s-r}\,, \tag{6.6}$$

where $\nu_{\rm max,\odot} = 3100\,\mu\mathrm{Hz}$ and $T_{\rm eff,\odot} = 5777\,\mathrm{K}$. The functional form of the amplitude scaling relation given by Eq. (6.6) has the advantage of simplifying the inference problem presented in Section 6.3 with respect to the one of Eq (6.4). This is because Eq. (6.6) is based on a set of independent observables, namely $\nu_{\rm max}$ and $T_{\rm eff}$, which represent the input data used for this work. An analogous argument has been applied to the other scaling relations described in the following sections.

Thus, Eq. (6.6) represents the first model to be investigated, which we will refer to as $\mathcal{M}_1$. For this model, both the exponents s and r are set to be free parameters. In parallel, an extended version of Eq. (6.6) given by

$$\frac{A_\lambda}{A_{\lambda,\odot}} = \beta \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{-s} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{3.5s-r}\,, \tag{6.7}$$

is also considered, where the factor β allows the model not to necessarily pass through the solar point, as we will discuss in more detail in Section 6.3.2. Eq. (6.7) is treated as a separate model and will be denoted as $\mathcal{M}_{1,\beta}$. Clearly, $\mathcal{M}_{1,\beta}$ depends upon the additional free parameter, β.

6.2.2 The bolometric amplitude

The stars considered span over a wide range of temperatures, from about 4000 K to more than 7000 K. Hence, following the discussion by [167], a more valuable expression for the photometric amplitude could be represented by the bolometric amplitude $A_{\rm bol}$, which is related to A_λ by (see [187]):

$$A_{\rm bol} \propto \lambda\, A_\lambda\, T_{\rm eff} \propto \frac{v_{\rm osc}}{T_{\rm eff}^{r-1}}\,. \tag{6.8}$$

By using Eq (6.1) we thus have

$$A_{\rm bol}^{(1)} \propto \left(\frac{L}{M}\right)^{s} \frac{1}{T_{\rm eff}^{r-1}}\,, \tag{6.9}$$

which by scaling to the Sun and adopting Eq. (6.5) yields

$$\frac{A_{\rm bol}^{(1)}}{A_{\rm bol,\odot}} = \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{-s} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{3.5s-r+1}\,, \tag{6.10}$$

where $A_{\rm bol,\odot} = 3.6\,$ppm represents the Sun's bolometric amplitude, determined by [230] and also adopted by [167]. Eq. (6.10) is the second model, $\mathcal{M}_2$, to be investigated in Section 6.3, with the exponents s and r set again to be the free parameters. We also consider the new model $\mathcal{M}_{2,\beta}$, which again includes the proportionality term β playing the same role as in Eq. (6.7).

6.2.3 The *Kepler* bandpass-corrected amplitude

[12] have recently established a bolometric correction for amplitude of radial modes observed with *Kepler*, which translates into a correction factor for effective temperatures falling within the range 4000–7500 K. Again following the approach by [167], we consider a revised version of Eq. (6.9), which reads

$$A_{\rm bol}^{(2)} \propto \left(\frac{L}{M}\right)^{s} \frac{1}{T_{\rm eff}^{r-1} c_K(T_{\rm eff})}\,, \tag{6.11}$$

where

$$c_{\rm K}(T_{\rm eff}) = \left(\frac{T_{\rm eff}}{5934}\right)^{0.8} \tag{6.12}$$

is the bolometric correction expressed as a power law of the effective temperature. By scaling once again to the Sun and applying Eq. (6.5), we obtain

$$\frac{A_{\rm bol}^{(2)}}{A_{\rm bol,\odot}} = \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{-s} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{3.5s-r+0.2}\,, \tag{6.13}$$

which we will refer to as $\mathcal{M}_3$. As for the other models, we introduce the model $\mathcal{M}_{3,\beta}$ with the proportionality term β included.

6.2.4 The mass-dependent scaling relation

A mass dependence of the oscillation amplitudes was suggested for the first time by [165], and later on studied in detail by [291] for cluster RGs with the introduction of a new scaling relation. It was also tested afterwards by [167] for a wider sample of field stars. According to [167], an obvious way to modify Eq. (6.11) is given by

$$A_{\rm bol}^{(3)} \propto \frac{L^s}{M^t} \frac{1}{T_{\rm eff}^{r-1} c_K(T_{\rm eff})}\,, \tag{6.14}$$

where now the mass varies with the independent exponent t. For this case, the dependence upon the quantities $\nu_{\rm max}$ and $T_{\rm eff}$ becomes slightly more complicated because the simple proportionality expressed by Eq. (6.5) can no longer be adopted. Therefore, the first step to derive the functional form based on our set of observables ($\nu_{\rm max}, \Delta\nu$, $T_{\rm eff}$), is to consider the scaling relations for the large frequency separation $\Delta\nu$ given by Eq. (2.43) with $\Delta\nu_\odot = 135\,\mu$Hz (see

Chapter 2, Section 2.4), and for the frequency of maximum power $\nu_{\rm max}$, given Eq. (2.45), which we now rewrite as

$$\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}} = \left(\frac{M}{M_\odot}\right)\left(\frac{R}{R_\odot}\right)^{-2}\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{-0.5} . \tag{6.15}$$

Both scaling relations are expressed in terms of the fundamental stellar properties M, R, and $T_{\rm eff}$. As already shown in Chapter 3, by combining them one can derive the expression for the seismic radius of a star, given by Eq. (3.4). As a second step, we express L^s/M^t in terms of R, $\nu_{\rm max}$, and $T_{\rm eff}$ and scale from the Sun's values, yielding

$$\left(\frac{L}{L_\odot}\right)^s\left(\frac{M}{M_\odot}\right)^{-t} = \left(\frac{R}{R_\odot}\right)^{2s-2t}\left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{-t} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{4s-0.5t} . \tag{6.16}$$

Finally, by combining Eqs. (6.12), (6.14), (**??**) and (6.16) we arrive at

$$\frac{A^{(3)}_{\rm bol}}{A_{\rm bol,\odot}} = \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{2s-3t}\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{4t-4s} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{5s-1.5t-r+0.2} . \tag{6.17}$$

This represents the model for the mass-dependent scaling relation for amplitudes, hereafter denoted as $\mathcal{M}_4$. In this case, we have the three free parameters s, r, and t and the set of observables now includes also our measurements of $\Delta\nu$. The corresponding model $\mathcal{M}_{4,\beta}$ has the largest number of free parameters among those investigated in this work. Note that models $\mathcal{M}_4$ and $\mathcal{M}_{4,\beta}$ reduce to models $\mathcal{M}_3$ and $\mathcal{M}_{3,\beta}$, respectively, for $t = s$.

6.2.5 The lifetime-dependent scaling relation

In their study of oscillation amplitudes and granulation [188] have recently provided physical arguments to propose a new scaling relation for predicting the amplitudes of solar-like oscillations observed in radial velocities. Their relation arises by postulating that the amplitudes depend on both the stochastic excitation (given by the granulation power, see [188] for details) and the damping rate (given by the mode lifetime). It reads

$$v_{\rm osc} \propto \frac{L\tau_{\rm osc}^{0.5}}{M^{1.5}T_{\rm eff}^{2.25}} , \tag{6.18}$$

where $\tau_{\rm osc}$ is the average mode lifetime of radial modes. By means of Eq. (6.8), and with the bolometric correction introduced in Section 6.2.3, the corresponding relation for the bolometric amplitude is given by (see also [167])

$$A^{(4)}_{\rm bol} \propto \frac{L\tau_{\rm osc}^{0.5}}{M^{1.5}}\frac{1}{T_{\rm eff}^{1.25+r}c_K(T_{\rm eff})} . \tag{6.19}$$

In order to obtain the expression for the model $\mathcal{M}_5$, we use similar arguments to those adopted in Section 6.2.4, arriving at

$$\frac{A^{(4)}_{\rm bol}}{A_{\rm bol,\odot}} = \left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)^{-2.5}\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{2}\left(\frac{\tau_{\rm osc}}{\tau_{\rm osc,\odot}}\right)^{0.5} \left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)^{2.3-r} , \tag{6.20}$$

where $\tau_{\mathrm{osc},\odot} = 2.88\,\mathrm{d}$, as adopted by [188]. For our computations we assume that the mode lifetime is a function of the effective temperature of the star alone (e.g. see [69, 19, 6, 35, 87]). We used the empirical law found by [87], which relates the mode linewidths Γ (see Chapter 2, Section 2.1.3 for a definition) of the radial modes ($\ell = 0$) to the effective temperature of the star within the range 4000–7000 K. In particular, they found that

$$\Gamma = \Gamma_0 \exp\left(\frac{T_{\mathrm{eff}} - T_{\mathrm{eff},\odot}}{T_0}\right), \tag{6.21}$$

where $\Gamma_0 = 1.39 \pm 0.10\,\mu\mathrm{Hz}$ and $T_0 = 601 \pm 3\,\mathrm{K}$ (see also Chapter 4, Section 4.3.5 and Appendix A for more details). This relation was calibrated using *Kepler* RGs in the open clusters NGC 6791 and NGC 6819, and a sample of MS and subgiant *Kepler* field stars. Given $\tau = (\pi\Gamma)^{-1}$, we obtain

$$\tau_{\mathrm{osc}} = \tau_0 \exp\left(\frac{T_{\mathrm{eff},\odot} - T_{\mathrm{eff}}}{T_0}\right), \tag{6.22}$$

with $\tau_0 = 2.65 \pm 0.19\,\mathrm{d}$. The mode lifetimes were computed for all the stars of our sample by means of Eq. (6.22), together with their corresponding uncertainties from the error propagation (see Appendix A). As for the other scaling relations, we also introduce model $\mathcal{M}_{5,\beta}$.

6.2.6 A new scaling relation

Following similar arguments to those adopted by [291] for introducing a new scaling relation for amplitudes of cluster RGs, and the discussion by [167] about the mass dependence of the amplitudes, we introduce a slightly modified version of the amplitude relation given by Eq. (6.19), where we set the mass to vary with an independent exponent t, thus yielding

$$A^{(5)}_{\mathrm{bol}} \propto \frac{L\tau^{0.5}_{\mathrm{osc}}}{M^t} \frac{1}{T^{1.25+r}_{\mathrm{eff}} c_K(T_{\mathrm{eff}})}. \tag{6.23}$$

By adopting again Eq. (6.16) and rearranging, we finally obtain

$$\frac{A^{(5)}_{\mathrm{bol}}}{A_{\mathrm{bol},\odot}} = \left(\frac{\nu_{\mathrm{max}}}{\nu_{\mathrm{max},\odot}}\right)^{2-3t} \left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)^{4t-4} \left(\frac{\tau_{\mathrm{osc}}}{\tau_{\mathrm{osc},\odot}}\right)^{0.5} \left(\frac{T_{\mathrm{eff}}}{T_{\mathrm{eff},\odot}}\right)^{4.55-r-1.5t}, \tag{6.24}$$

hereafter marked as model $\mathcal{M}_6$. As done for the other scaling relations, the model $\mathcal{M}_{6,\beta}$ is also included in our inference. Clearly, models $\mathcal{M}_6$ and $\mathcal{M}_{6,\beta}$ reduce to models $\mathcal{M}_5$ and $\mathcal{M}_{5,\beta}$, respectively, for $t = 1.5$.

6.3 Bayesian inference

We now use Bayesian inference for the free parameters of the models described above. The Bayes' theorem tells us that

$$p(\boldsymbol{\xi} \mid A, \mathcal{M}) = \frac{p(A \mid \boldsymbol{\xi}, \mathcal{M})\pi(\boldsymbol{\xi} \mid \mathcal{M})}{p(A \mid \mathcal{M})}, \tag{6.25}$$

where $\boldsymbol{\xi} = \xi_1, \xi_2, \ldots, \xi_k$ is the vector of the k free parameters that formalize the hypotheses of the model $\mathcal{M}$, considered for the inference, and A is the set of amplitude measurements (see

Chapter 5 for a general introduction to Bayesian probability theory and statistics). The term $p(A \mid \boldsymbol{\xi}, \mathcal{M})$ is now identified with the likelihood $\mathcal{L}(\boldsymbol{\xi})$ of the parameters ξ_i given the measured oscillation amplitudes:

$$p(A \mid \boldsymbol{\xi}, \mathcal{M}) = \mathcal{L}(\boldsymbol{\xi} \mid A, \mathcal{M}) \,. \tag{6.26}$$

Thus, the left-hand side of Eq. (6.25) is the posterior probability density function (PDF), while the right-hand side is the product of the likelihood function $\mathcal{L}(\boldsymbol{\xi})$, which represents our manner of comparing the data to the predictions by the model, and the prior PDF $\pi(\boldsymbol{\xi} \mid \mathcal{M})$, which represents our knowledge of the inferred parameters before any information from the data is available. The term $p(A \mid \mathcal{M})$ is a normalization factor, known as the Bayesian evidence, which we do not consider for the inference problem because it is a constant for a model alone. As we will argue in Section 6.4, the Bayesian evidence is essential for solving the problem of model comparison (see also Chapter 5, Section 5.6 for more details).

For our inference problem, we adopt the common Gaussian likelihood function, which presumes that the residuals arising from the difference between observed and predicted logarithms of the amplitudes are Gaussian distributed, i.e. the amplitudes themselves are presumed to be log-normal distributed (see also [5]). Therefore, we have

$$\mathcal{L}(\boldsymbol{\xi}) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\widetilde{\sigma}_i} \exp\left[-\frac{1}{2}\left(\frac{\Delta_i(\boldsymbol{\xi})}{\widetilde{\sigma}_i}\right)^2\right], \tag{6.27}$$

where N is the total number of data points (the number of stars, in our case), while

$$\Delta_i(\boldsymbol{\xi}) = \ln A_i^{\mathrm{obs}} - \ln A_i^{\mathrm{th}}(\boldsymbol{\xi}) \tag{6.28}$$

is the difference between the observed logarithmic amplitude for the i-th star and the predicted one (which depends on the adopted model, i.e. on the parameters vector $\boldsymbol{\xi}$). The term $\widetilde{\sigma}_i$ appearing in the leading exponential term of Eq. (6.27) is the total uncertainty in the predicted logarithmic amplitude, namely the relative uncertainty of the amplitude enlarged by the relative errors of the independent variables $\nu_{\max}$, $\Delta\nu$, and T_{eff}. This means that we are not considering error-free variables in our models (see e.g. [95, 4] for more details). For simplifying the computations a modified version of the likelihood function, known as the *log-likelihood*, is preferred. The log-likelihood function is defined as $\Lambda(\boldsymbol{\xi}) \equiv \ln \mathcal{L}(\boldsymbol{\xi})$, which yields

$$\Lambda(\boldsymbol{\xi}) = \Lambda_0 - \frac{1}{2}\sum_{i=1}^{N}\left[\frac{\Delta_i(\boldsymbol{\xi})}{\widetilde{\sigma}_i}\right]^2, \tag{6.29}$$

where

$$\Lambda_0 = -\sum_{i=1}^{N} \ln \sqrt{2\pi}\widetilde{\sigma}_i \,. \tag{6.30}$$

Further details about the method used to compute the final likelihood are given in Appendix C, Section C.3.

The choice of reliable priors is important in the Bayesian approach. For our purpose, uniform priors represent a sensible choice for most of the free parameters. This means one has no assumptions about the inferred parameters before any knowledge coming from the data, with equal weight being given to all values of each of the parameters considered. In particular, we use standard uniform priors for the exponents s, r and t of the models described above, letting the parameters vary within a limited range in order to make the priors proper, i.e. normalizable to unity. For the proportionality term β introduced in Eq. (6.7), we adopt the Jeffreys' prior

Table 6.1: *Maximum and minimum values of the free parameters, adopted for all the models and samples.*

ξ_j	$[\xi_j^{\min}, \xi_j^{\max}]$
s	$[0.2, 1.2]$
r	$[-6.5, 11.0]$
t	$[1.0, 2.0]$
$\ln \beta$	$[-1.0, 1.0]$

$\propto \beta^{-1}$ [184], a class of uninformative prior that results in a uniform prior for the natural logarithm of the parameter (see also Chapter 5, Section 5.2.2). In this manner, the parameter of interest is represented by the offset $\ln \beta$ (see below), whose prior is uniform distributed and also limited in range. Hence, uniform priors are included in the inference problem as a simple constant term depending on the intervals adopted for the inferred parameters

$$\pi(\boldsymbol{\xi} \mid \mathcal{M}) = \prod_{j=1}^{4} \left[\xi_j^{\max} - \xi_j^{\min}\right]^{-1} \tag{6.31}$$

with $\xi_1 = s$, $\xi_2 = r$, $\xi_3 = t$ and $\xi_4 = \ln \beta$, and $\xi_j^{\min}, \xi_j^{\max}$ the minimum and maximum values defining the interval of the j-th parameter (see also Appendix C, Section **??** for more details). The intervals that we adopt are listed in Table 6.1. These ranges are used for both the Bayesian parameter estimation and the model comparison.

A note of caution concerns the treatment of the uncertainties. In fact, by using the natural logarithm of the equations that describe the models (see also Section 6.3.2), we ensure that we are not favoring for example frequencies upon periods, amplitudes upon power, temperatures upon surface brightness, etc., which has the advantage of making the error propagation law fully correct. Thus, the corresponding uncertainties to be considered in Eq. (6.27) will be the relative uncertainties.

As already explained in Chapter 5, Section 5.3 for a general case, the inference problem for a given parameter, e.g. ξ_1, is then performed by integrating (marginalizing) the posterior distribution function $p(\boldsymbol{\xi} \mid d, \mathcal{M})$ over the remaining $k-1$ parameters $\xi_2, \xi_3, \ldots, \xi_k$. We obtain the corresponding marginal PDF of the parameter ξ_1

$$p(\xi_1 \mid A, \mathcal{M}) = \int p(\boldsymbol{\xi} \mid A, \mathcal{M}) d\xi_2, d\xi_3, \ldots, d\xi_k \,, \tag{6.32}$$

whose statistical moments and credible intervals (i.e. Bayesian confidence intervals) are the quantities of interest for our work. In particular, the confidence intervals are introduced in Appendix C, Section C.2, together with a detailed description of their computation for the work described in this chapter. Since the dimensionality of our problem is not higher than $k = 4$, all the integrations can be computed by direct numerical summation of the posterior distribution over the remaining parameters (only in case the observables, like $T_{\rm eff}$, were error-free all integrations could be computed analytically).

The results presented in the coming sections are derived in three cases: for the entire sample of stars, and for SC (dominated by MS stars) and LC (RGs) targets separately. Analyzing the two subsets separately allows us to test whether the fitted parameters of the scaling relations are sensitive to the evolutionary stage of the stars (see also [167]).

The mean values (or expectation values) of the free parameters of the models, together with their corresponding 68.3 % Bayesian credible intervals, are listed in Table 6.2 for the case of the entire sample, and in Tables 6.3 and 6.4 for the subsets of SC and LC targets. We also computed

a weighted Gaussian rms, $\sigma^w_{\rm rms}$, of the residuals $\Delta^2_i(\boldsymbol{\xi})$, where we adopted the weights $w_i = \widetilde{\sigma}_i^{-2}$, $\widetilde{\sigma}_i$ being the total uncertainty used in Eq. (6.27). The maximum of the log-likelihood function, $\Lambda_{\rm max}$, and $\sigma^w_{\rm rms}$, used as an estimate of the fit quality, are also listed in the same tables. In addition, we derived correlation coefficients for each pair of free parameters by means of principal component analysis using Singular Value Decomposition (SVD) from the posterior PDFs (see Appendix D for an introduction). The results are shown in Tables 6.5, 6.6, 6.7 for the entire sample, and SC and LC targets, respectively, with -1 meaning total anti-correlation and 1 total correlation. The effects of the correlations will be discussed in Section 6.5.

6.3.1 Statistically independent models

Before getting to a description on how to correctly include the models in the inference problem, it is useful to highlight that the models $\mathcal{M}_{1,\beta}$, $\mathcal{M}_{2,\beta}$, and $\mathcal{M}_{3,\beta}$ on the one hand, and the models $\mathcal{M}_2$, and $\mathcal{M}_3$ on the other hand, are statistically identical to one another (but not identical in the general sense, since their underlying physical assumptions are different). In case the intervals of their free parameters are the same for all the models when performing the Bayesian parameter estimation (as it is for the analysis presented in this work) this implies that a statistical inference for these models would lead to identical (or directly related) values of these free parameters. In particular, according to Eqs. (6.7), (6.10), and (6.13) we have that

$$\begin{aligned} s(\mathcal{M}_{1,\beta}) &= s(\mathcal{M}_{2,\beta}) = s(\mathcal{M}_{3,\beta}) \\ s(\mathcal{M}_2) &= s(\mathcal{M}_3) \end{aligned} \tag{6.33}$$

for the exponent s,

$$\begin{aligned} r(\mathcal{M}_{1,\beta}) &= r(\mathcal{M}_{2,\beta}) - 1 = r(\mathcal{M}_{3,\beta}) - 0.2 \\ r(\mathcal{M}_2) &= r(\mathcal{M}_3) - 0.8\,, \end{aligned} \tag{6.34}$$

for the exponent r, and

$$\begin{aligned} \ln\beta(\mathcal{M}_{1,\beta}) &= \ln\beta(\mathcal{M}_{2,\beta}) + \ln\left(\frac{A_{\rm bol,\odot}}{A_{\lambda,\odot}}\right) \\ &= \ln\beta(\mathcal{M}_{3,\beta}) + \ln\left(\frac{A_{\rm bol,\odot}}{A_{\lambda,\odot}}\right) \end{aligned} \tag{6.35}$$

for the offset $\ln\beta$, where the term $\ln\left(A_{\rm bol,\odot}/A_{\lambda,\odot}\right)$ arises from the difference in considering the amplitudes to be either observed at $\lambda = 650\,{\rm nm}$ ($\mathcal{M}_{1,\beta}$) or bolometric ($\mathcal{M}_{2,\beta}$ and $\mathcal{M}_{3,\beta}$). As a consequence, from now on the entire analysis for the models $\mathcal{M}_{1,\beta}$, $\mathcal{M}_{2,\beta}$, and $\mathcal{M}_{3,\beta}$ on one side, and for the models $\mathcal{M}_2$ and $\mathcal{M}_3$ on the other side, will be reduced to that of the two models $\mathcal{M}_{1,\beta}$ and $\mathcal{M}_2$, respectively. The reader can derive the corresponding parameters for the other dependent models using Eqs. (6.33), (6.34), and (6.35).

6.3.2 Models $\mathcal{M}_1$ and $\mathcal{M}_{1,\beta}$

The first model to be investigated is given by Eq. (6.6). As argued above, we need to consider the natural logarithm in order to treat the observables independently of the function adopted (see the discussion in Section 6.3 and in Chapter 5, Section 5.2.2). Hence the model reads

$$\begin{aligned} \ln\left(\frac{A_\lambda}{A_{\lambda,\odot}}\right) = &- s\ln\left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right) \\ &+ (3.5s - r)\ln\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right)\,. \end{aligned} \tag{6.36}$$

Table 6.2: *Expectation values of the inferred parameters for all the models described in Section 6.2 in the case of the entire sample (both LC and SC targets), having $N = 1640$ stars. 68.3 % Bayesian credible intervals are added. The maximum value for the log-likelihood function $\Lambda_{\max}$ and a weighted Gaussian rms of the residuals, $\sigma^{w}_{\rm rms}$, are also reported.*

Model	s	r	t	$\ln\beta$	$\Lambda_{\max}$	$\sigma^{w}_{\rm rms}$
$\mathcal{M}_1$	0.680 ± 0.002	4.31 ± 0.04	–	–	-3533.1	0.23
$\mathcal{M}_{1,\beta}$	0.524 ± 0.004	5.51 ± 0.05	–	0.400 ± 0.010	-2491.3	0.24
$\mathcal{M}_2$	0.722 ± 0.002	4.96 ± 0.04	–	–	-4159.9	0.24
$\mathcal{M}_4$	0.822 ± 0.003	3.93 ± 0.06	1.58 ± 0.02	–	-948.5	0.23
$\mathcal{M}_{4,\beta}$	0.643 ± 0.005	4.46 ± 0.06	1.36 ± 0.02	0.528 ± 0.012	163.4	0.18
$\mathcal{M}_5$	–	-5.71 ± 0.03	–	–	-925.5	0.31
$\mathcal{M}_{5,\beta}$	–	-5.07 ± 0.05	–	-0.122 ± 0.007	-786.9	0.29
$\mathcal{M}_6$	–	-5.04 ± 0.05	1.80 ± 0.02	–	-727.1	0.30
$\mathcal{M}_{6,\beta}$	–	-4.97 ± 0.06	1.75 ± 0.02	-0.035 ± 0.012	-722.8	0.29

Table 6.3: *Same description as for Table 6.2 but in the case of SC targets, having $N = 529$ stars.*

Model	s	r	t	$\ln\beta$	$\Lambda_{\max}$	$\sigma^{w}_{\rm rms}$
$\mathcal{M}_1$	0.775 ± 0.003	3.23 ± 0.05	–	–	-321.4	0.20
$\mathcal{M}_{1,\beta}$	0.624 ± 0.010	3.68 ± 0.06	–	0.241 ± 0.016	-199.3	0.19
$\mathcal{M}_2$	0.838 ± 0.003	4.05 ± 0.05	–	–	-443.8	0.22
$\mathcal{M}_4$	$0.984^{+0.009}_{-0.010}$	2.79 ± 0.09	$1.66^{+0.04}_{-0.05}$	–	-94.9	0.23
$\mathcal{M}_{4,\beta}$	0.748 ± 0.015	3.47 ± 0.09	1.27 ± 0.04	0.321 ± 0.020	18.0	0.20
$\mathcal{M}_5$	–	-2.78 ± 0.09	–	–	-83.1	0.24
$\mathcal{M}_{5,\beta}$	–	-2.75 ± 0.09	–	0.020 ± 0.008	-80.2	0.24
$\mathcal{M}_6$	–	$-2.80^{+0.10}_{-0.09}$	1.56 ± 0.02	–	-79.6	0.24
$\mathcal{M}_{6,\beta}$	–	-2.75 ± 0.10	1.72 ± 0.04	$0.087^{+0.014}_{-0.015}$	-59.4	0.24

Table 6.4: *Same description as for Table 6.2 but in the case of LC targets, having $N = 1111$ stars.*

Model	s	r	t	$\ln\beta$	$\Lambda_{\max}$	$\sigma^{w}_{\rm rms}$
$\mathcal{M}_1$	$0.464^{+0.007}_{-0.006}$	$9.53^{+0.15}_{-0.16}$	–	–	-799.2	0.28
$\mathcal{M}_{1,\beta}$	0.548 ± 0.009	$9.67^{+0.15}_{-0.16}$	–	-0.35 ± 0.03	-737.4	0.27
$\mathcal{M}_2$	$0.491^{+0.007}_{-0.006}$	$10.51^{+0.15}_{-0.16}$	–	–	-769.0	0.28
$\mathcal{M}_4$	$0.666^{+0.006}_{-0.005}$	6.99 ± 0.12	1.28 ± 0.02	–	207.4	0.18
$\mathcal{M}_{4,\beta}$	0.602 ± 0.008	5.87 ± 0.14	1.31 ± 0.02	0.45 ± 0.03	301.0	0.16
$\mathcal{M}_5$	–	-6.08 ± 0.04	–	–	-357.9	0.24
$\mathcal{M}_{5,\beta}$	–	-4.38 ± 0.16	–	$-0.27^{+0.02}_{-0.03}$	-309.2	0.24
$\mathcal{M}_6$	–	-5.94 ± 0.08	1.55 ± 0.03	–	-356.0	0.24
$\mathcal{M}_{6,\beta}$	–	-4.39 ± 0.16	$1.45^{+0.02}_{-0.03}$	-0.30 ± 0.03	-307.1	0.24

Table 6.5: *Correlation coefficients for pairs of free parameters for each model in the case of the entire sample.*

Model	s vs r	s vs b	s vs t	r vs b	r vs t	b vs t
$\mathcal{M}_1$	−0.90	–	–	–	–	–
$\mathcal{M}_{1,\beta}$	−0.85	−0.92	–	0.66	–	–
$\mathcal{M}_2$	−0.94	–	–	–	–	–
$\mathcal{M}_4$	−0.74	–	0.59	–	0.01	–
$\mathcal{M}_{4,\beta}$	−0.74	−0.76	0.44	0.31	−0.25	0.04
$\mathcal{M}_5$	–	–	–	–	–	–
$\mathcal{M}_{5,\beta}$	–	–	–	−0.74	–	–
$\mathcal{M}_6$	–	–	–	–	0.74	–
$\mathcal{M}_{6,\beta}$	–	–	–	−0.41	0.18	0.71

Table 6.6: *Same description as for Table 6.5 but in the case of SC targets.*

Model	s vs r	s vs b	s vs t	r vs b	r vs t	b vs t
$\mathcal{M}_1$	0.22	–	–	–	–	–
$\mathcal{M}_{1,\beta}$	−0.40	−0.94	–	0.49	–	–
$\mathcal{M}_2$	0.20	–	–	–	–	–
$\mathcal{M}_4$	−0.29	–	0.88	–	−0.27	–
$\mathcal{M}_{4,\beta}$	−0.48	−0.85	0.71	0.44	−0.37	−0.31
$\mathcal{M}_5$	–	–	–	–	–	–
$\mathcal{M}_{5,\beta}$	–	–	–	0.12	–	–
$\mathcal{M}_6$	–	–	–	–	−0.07	–
$\mathcal{M}_{6,\beta}$	–	–	–	0.03	−0.01	0.76

Table 6.7: *Same description as for Table 6.5 but in the case of LC targets.*

Model	s vs r	s vs b	s vs t	r vs b	r vs t	b vs t
$\mathcal{M}_1$	−0.98	–	–	–	–	–
$\mathcal{M}_{1,\beta}$	−0.49	−0.71	–	−0.25	–	–
$\mathcal{M}_2$	−0.98	–	–	–	–	–
$\mathcal{M}_4$	−0.92	–	0.50	–	−0.19	–
$\mathcal{M}_{4,\beta}$	−0.31	−0.66	0.30	−0.46	−0.30	0.18
$\mathcal{M}_5$	–	–	–	–	–	–
$\mathcal{M}_{5,\beta}$	–	–	–	−0.98	–	–
$\mathcal{M}_6$	–	–	–	–	0.88	–
$\mathcal{M}_{6,\beta}$	–	–	–	−0.88	0.04	0.37

At this stage we briefly describe how the uncertainties have been included in our analysis. The new uncertainties on the scaled amplitude to be considered in Eq. (6.27) are clearly given by σ_{A_i}/A_i, hereafter $\widetilde{\sigma}_{A_i}$ for simplicity, where A_i is the observed amplitude for the i-th star and σ_{A_i} its corresponding uncertainty as derived by [167]. However, Eq. (6.36) suggests that the uncertainty on amplitude is not the only one affecting the predicted amplitude A_λ. In fact, uncertainties on both $\nu_{\rm max}$ (derived by [167]) and $T_{\rm eff}$ (from [256]) have to be included in our computations. The total uncertainty to be used in Eq. (6.27) is given by the Gaussian error propagation law, which gives

$$\widetilde{\sigma}_i^2(s,r) = \widetilde{\sigma}_{A_i}^2 + s^2\widetilde{\sigma}_{\nu_{\rm max},i}^2 + (3.5s-r)^2\widetilde{\sigma}_{T_{\rm eff},i}^2 \tag{6.37}$$

where we defined $\widetilde{\sigma}_{\nu_{\rm max},i} \equiv \sigma_{\nu_{\rm max},i}/\nu_{{\rm max},i}$ and $\widetilde{\sigma}_{T_{\rm eff},i} \equiv \sigma_{T_{\rm eff},i}/T_{{\rm eff},i}$, similarly to what was done for the amplitudes. Eq. (6.37) only holds in case of uncorrelated uncertainties and linear relations (see also [95, 4] for a demonstration). Intuitively, Eq. (6.37) is the quadratic sum of the relative uncertainties over the physical quantities considered, according to Eq. (6.36). A variation of Eq. (6.36) is represented by the model $\mathcal{M}_{1,\beta}$ given by Eq. (6.7), whose natural logarithm reads

$$\begin{aligned}\ln\left(\frac{A_\lambda}{A_{\lambda,\odot}}\right) = &- s\ln\left(\frac{\nu_{\rm max}}{\nu_{{\rm max},\odot}}\right) \\ &+ (3.5s-r)\ln\left(\frac{T_{\rm eff}}{T_{{\rm eff},\odot}}\right) + \ln\beta\,.\end{aligned} \tag{6.38}$$

which differs from Eq. (6.36) by the additional term $\ln\beta$. As already argued before, the offset $\ln\beta$ allows the model not to necessarily pass through the solar point $(A_\odot, \nu_{{\rm max},\odot}, T_{{\rm eff},\odot})$. Its introduction in the inference is of importance if one wants to assess whether or not the Sun is a good reference star for the sample considered. This choice is also motivated by the fact that the Sun is falling at the edge of the sample of stars when plotting amplitudes against $\nu_{\rm max}$ and $\Delta\nu$ (Figure 6.2, top left and right panels). This peculiar position is also evident from our asteroseismic HR diagram (Figure 6.2, bottom panel), and is caused by the lack of solar twins in our sample of stars (see the discussion by [71] and the results shown in Chapter 3, Section 3.4). In fact, in the case of $\ln\beta \neq 0$, by replacing the solar values in Eq. (6.38) (or alternatively Eq. (6.7)), the predicted amplitude for the Sun would be resized by a factor β. This means that the best reference amplitude for scaling the amplitudes of our sample of stars would be represented by $\beta A_{\lambda,\odot}$. According to Eq. (6.38), the total uncertainty to be considered in building the likelihood function for the model $\mathcal{M}_{1,\beta}$ is given again by Eq. (6.37) because the offset does not play any role in the total contribute of the uncertainties.

A representative sample of the resulting marginal PDFs is plotted in Figure 6.3 for the case of the entire sample, where 68.3 % Bayesian credible regions (shaded bands) and expectation values (dashed lines) are also marked. The comparison between the predicted and the observed amplitudes is shown in Figure 6.4 for the three cases considered (top panels for model $\mathcal{M}_1$, bottom panels for model $\mathcal{M}_{1,\beta}$), together with a plot of the residuals arising from the difference between the models and the observations.

6.3.3 Model $\mathcal{M}_2$

The model $\mathcal{M}_2$ given by Eq. (6.10) deserves a similar description to that presented in Section (6.3.2) for model $\mathcal{M}_1$, where the natural logarithm is now given by

$$\ln\left(\frac{A_{\rm bol}^{(1)}}{A_{\rm bol,\odot}}\right) = -\,s\ln\left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right) + (3.5s - r + 1)\ln\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right), \tag{6.39}$$

with a total uncertainty for the i-th star expressed as

$$\widetilde{\sigma}_i^2(s,r) = \widetilde{\sigma}_{A_i}^2 + s^2\widetilde{\sigma}_{\nu_{{\rm max},i}}^2 + (3.5s - r + 1)^2\widetilde{\sigma}_{T_{{\rm eff},i}}^2 \tag{6.40}$$

to be included in Eq. (6.27). The resulting models are shown in Figure 6.5, with similar descriptions as those adopted for Figure 6.4.

6.3.4 Models $\mathcal{M}_4$ and $\mathcal{M}_{4,\beta}$

The models $\mathcal{M}_4$ and $\mathcal{M}_{4,\beta}$ (see Section 6.2.4) are clearly the most complex ones among those investigated in this work because the largest number of free parameters is involved, and measurements of $\Delta\nu$ are also needed. We note that, although a tight correlation between $\nu_{\rm max}$ and $\Delta\nu$ has been found in previous studies (e.g. see [286]), we choose not to adopt the $\nu_{\rm max}$–$\Delta\nu$ relation to express model $\mathcal{M}_4$ in terms of $\nu_{\rm max}$ only (or alternatively $\Delta\nu$) because additional uncertainties arising from the scatter around this relation would affect the results of our inference. This is also motivated by recent results by [167] who found that the $\nu_{\rm max}$-$\Delta\nu$ relation changes as a function of $T_{\rm eff}$ between dwarf and giant stars.

Therefore, by considering the natural logarithm of Eq. (6.17) one obtains

$$\ln\left(\frac{A_{\rm bol}^{(3)}}{A_{\rm bol,\odot}}\right) = (2s - 3t)\ln\left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right) + (4t - 4s)\ln\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right) + (5s - 1.5t - r + 0.2)\ln\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right), \tag{6.41}$$

for model $\mathcal{M}_4$, and with the additional term $\ln\beta$ for $\mathcal{M}_{4,\beta}$. According to Eq. (6.41), the total uncertainty to be considered in Eq. (6.27) reads

$$\widetilde{\sigma}_i^2(s,r,t) = \widetilde{\sigma}_{A_i}^2 + (2s - 3t)^2\widetilde{\sigma}_{\nu_{{\rm max},i}}^2 + (4t - 4s)^2\widetilde{\sigma}_{\Delta\nu_i}^2 + (5s - 1.5t - r + 0.2)^2\widetilde{\sigma}_{T_{{\rm eff},i}}^2, \tag{6.42}$$

with $\widetilde{\sigma}_{\Delta\nu_i} \equiv \sigma_{\Delta\nu_i}/\Delta\nu_i$, as done for the other quantities. As one can intuitively expect, the new total uncertainty depends on the three free parameters of the model. The resulting models are shown in Figure 6.6.

6.3.5 Models $\mathcal{M}_5$ and $\mathcal{M}_{5,\beta}$

The models described in Section 6.2.5 are derived with a quite different approach, which requires an estimate of the mode lifetime for each star considered in our sample. We note that model $\mathcal{M}_5$ was also investigated by [167], who however did not take into account mode lifetimes.

The natural logarithm of model $\mathcal{M}_5$ lead us to

$$\begin{aligned}\ln\left(\frac{A_{\rm bol}^{(4)}}{A_{\rm bol,\odot}}\right) = &-2.5\ln\left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)\\ &+2\ln\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)+0.5\ln\left(\frac{\tau_{\rm osc}}{\tau_{\rm osc,\odot}}\right)\\ &+(2.3-r)\ln\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right).\end{aligned} \tag{6.43}$$

Thus, the total uncertainty for the i-th star of the sample is given by

$$\begin{aligned}\widetilde{\sigma}_i^2(r) = &\widetilde{\sigma}_{A_i}^2+6.25\widetilde{\sigma}_{\nu_{{\rm max},i}}^2+4\widetilde{\sigma}_{\Delta\nu_i}^2\\ &+0.25\widetilde{\sigma}_{\tau_{{\rm osc},i}}^2+(2.3-r)^2\widetilde{\sigma}_{T_{{\rm eff},i}}^2,\end{aligned} \tag{6.44}$$

with $\widetilde{\sigma}_{\tau_{{\rm osc},i}}^2 \equiv \sigma_{\tau_{{\rm osc},i}}/\tau_{{\rm osc},i}$, and is the same for both the models here considered, as model $\mathcal{M}_{5,\beta}$ differs only by the additional term $\ln\beta$. The resulting models are shown in Figure 6.7, with similar descriptions as those adopted for Figures 6.3 and 6.4.

6.3.6 Models $\mathcal{M}_6$ and $\mathcal{M}_{6,\beta}$

Following the same arguments used for the other models, we obtain

$$\begin{aligned}\ln\left(\frac{A_{\rm bol}^{(5)}}{A_{\rm bol,\odot}}\right) = &(2-3t)\ln\left(\frac{\nu_{\rm max}}{\nu_{\rm max,\odot}}\right)\\ &+(4t-4)\ln\left(\frac{\Delta\nu}{\Delta\nu_\odot}\right)+0.5\ln\left(\frac{\tau_{\rm osc}}{\tau_{\rm osc,\odot}}\right)\\ &+(4.55-r-1.5t)\ln\left(\frac{T_{\rm eff}}{T_{\rm eff,\odot}}\right).\end{aligned} \tag{6.45}$$

for the model $\mathcal{M}_6$. The new uncertainties to be considered will depend on the free parameters r and t, thus we have

$$\begin{aligned}\widetilde{\sigma}_i^2(r,t) = &\widetilde{\sigma}_{A_i}^2+(2-3t)^2\widetilde{\sigma}_{\nu_{{\rm max},i}}^2+(4t-4)^2\widetilde{\sigma}_{\Delta\nu_i}^2\\ &+0.25\widetilde{\sigma}_{\tau_{{\rm osc},i}}^2+(4.55-r-1.5t)^2\widetilde{\sigma}_{T_{{\rm eff},i}}^2.\end{aligned} \tag{6.46}$$

with the same quantities adopted in Eq. (6.44). An analogous discussion to that used for other scaling relations has to be applied for model $\mathcal{M}_{6,\beta}$. The results of the inference for both models are plotted in Figure 6.8.

6.4 Bayesian Model Comparison

As mentioned in Section 6.3, the term $p(A \mid \mathcal{M})$ appearing in Eq. (6.25) (Bayesian evidence) is the one of interest for solving the problem of model comparison in the context of Bayesian statistics (e.g. see [303, 37, 144, 141] and Chapter 5, Section 5.6). The Bayesian evidence is given by integrating the numerator appearing in the right-hand side of Eq. (6.25) over all the possible values of the free parameters $\xi_1, \xi_2, \ldots, \xi_k$. Thus we have

$$\mathcal{E}_{\mathcal{M}} \equiv p(A \mid \mathcal{M}) = \int_{\Omega_{\mathcal{M}}} \mathcal{L}(\boldsymbol{\xi})\pi(\boldsymbol{\xi} \mid \mathcal{M})d\boldsymbol{\xi}\,, \tag{6.47}$$

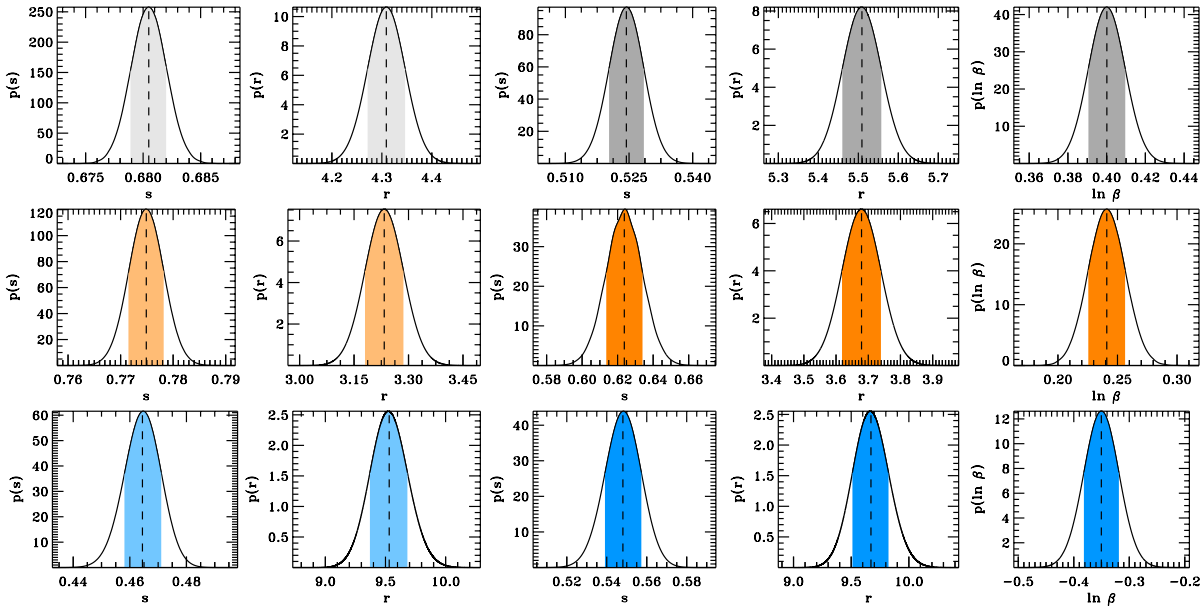

Figure 6.3: *Top panels: Marginal PDFs for the free parameters of the models $\mathcal{M}_1$ and $\mathcal{M}_{1,\beta}$, where expectation values listed in Table 6.2 (dashed lines) and 68.27 % Bayesian credible regions (light gray for $\mathcal{M}_1$ and dark gray for $\mathcal{M}_{1,\beta}$) have been marked. Middle panels: Same description of the top panels but for the case of SC targets only (light orange for $\mathcal{M}_1$ and dark orange for $\mathcal{M}_{1,\beta}$). Bottom panels: Same description of the top panels but for the case of LC targets only (light blue for $\mathcal{M}_1$ and dark blue for $\mathcal{M}_{1,\beta}$).*

where $\Omega_{\mathcal{M}}$ represents the parameter space, defined by the intervals of variation of the free parameters that formalize the hypotheses of the model $\mathcal{M}$, and having volume given by Eq. (6.31). The Bayesian evidence given by Eq. (6.47) basically represents the integral of the likelihood function "averaged" by the prior distribution. As the prior $\pi(\boldsymbol{\xi} \mid \mathcal{M})$ has to be normalized, the evidence depends on the parameter space. Thus, the intervals $[\xi_j^{\min}, \xi_j^{\max}]$ of the free parameters ξ_j used for computing Eq. (6.47) are those listed in Table 6.1. A detailed discussion about the computation of the evidences is given in Appendix C, Section C.3, where the values of all the evidences are also listed.

Since a measure of $\mathcal{E}_{\mathcal{M}}$ alone does not carry any meaningful information, to solve the problem of model comparison it is useful to take into account the ratios (or odds) of the evidences, namely the so-called Bayes factor, which is given as

$$B_{ij} = \frac{p(A \mid \mathcal{M}_i)}{p(A \mid \mathcal{M}_j)} = \frac{\mathcal{E}_{\mathcal{M}_i}}{\mathcal{E}_{\mathcal{M}_j}} \,. \tag{6.48}$$

We recall from Chapter 5 that in case $B_{ij} > 1$ the model $\mathcal{M}_i$ is the favored one, while conversely if $B_{ij} < 1$ the model $\mathcal{M}_j$ ought to be preferred. The resulting natural logarithms of the Bayes factor, which are computed according to Eq. (6.48) are listed in Tables 6.8, 6.9, and 6.10, for the cases of the entire sample, and SC and LC targets, respectively. Therefore, if $\ln B_{ij} > 0$ the model M_i is preferred over M_j and vice versa if $\ln B_{ij} < 0$.

It is sometimes useful to consider so-called Information Criteria, which may offer a simpler alternative to the Bayesian evidence, whose numerical computation in some cases can be very time demanding. In particular, we adopted the Bayesian Information Criterion (BIC), also known as Schwarz Information Criterion [273], which follows from a Gaussian approximation

Table 6.8: *Natural logarithms of the Bayes factor* $\ln B_{ij}$ *for each pair of models* $\mathcal{M}_i$, $\mathcal{M}_j$ *as derived by means of Eq. (6.48) for the case of the entire sample.*

Model	$\mathcal{M}_1$	$\mathcal{M}_{1,\beta}$	$\mathcal{M}_2$	$\mathcal{M}_4$	$\mathcal{M}_{4,\beta}$	$\mathcal{M}_5$	$\mathcal{M}_{5,\beta}$	$\mathcal{M}_6$
$\mathcal{M}_1$	–							
$\mathcal{M}_{1,\beta}$	1038.2	–						
$\mathcal{M}_2$	−626.7	−1664.9	–					
$\mathcal{M}_4$	2582.3	1544.1	3209.0	–				
$\mathcal{M}_{4,\beta}$	3690.6	2652.4	4317.3	1108.3	–			
$\mathcal{M}_5$	2613.9	1575.7	3240.6	31.6	−1076.6	–		
$\mathcal{M}_{5,\beta}$	2748.4	1710.2	3375.2	166.2	−942.1	134.5	–	
$\mathcal{M}_6$	2809.1	1770.9	3435.9	226.8	−881.4	195.2	60.7	–
$\mathcal{M}_{6,\beta}$	2810.0	1771.8	3436.7	227.7	−880.6	196.1	61.6	0.9

Table 6.9: *Same description as for Table 6.8 but in the case of the sample of SC targets.*

Model	$\mathcal{M}_1$	$\mathcal{M}_{1,\beta}$	$\mathcal{M}_2$	$\mathcal{M}_4$	$\mathcal{M}_{4,\beta}$	$\mathcal{M}_5$	$\mathcal{M}_{5,\beta}$	$\mathcal{M}_6$
$\mathcal{M}_1$	–							
$\mathcal{M}_{1,\beta}$	118.8	–						
$\mathcal{M}_2$	−122.4	−241.2	–					
$\mathcal{M}_4$	225.2	106.4	347.6	–				
$\mathcal{M}_{4,\beta}$	334.7	215.8	457.0	109.5	–			
$\mathcal{M}_5$	243.7	124.8	366.0	18.5	−91.0	–		
$\mathcal{M}_{5,\beta}$	242.7	123.9	365.1	17.5	−91.9	−0.9	–	
$\mathcal{M}_6$	244.3	125.5	366.7	19.1	−90.3	0.7	1.6	–
$\mathcal{M}_{6,\beta}$	261.4	142.6	383.8	36.2	−73.2	17.8	18.7	17.1

Table 6.10: *Same description as for Table 6.8 but in the case of the sample of LC targets.*

Model	$\mathcal{M}_1$	$\mathcal{M}_{1,\beta}$	$\mathcal{M}_2$	$\mathcal{M}_4$	$\mathcal{M}_{4,\beta}$	$\mathcal{M}_5$	$\mathcal{M}_{5,\beta}$	$\mathcal{M}_6$
$\mathcal{M}_1$	–							
$\mathcal{M}_{1,\beta}$	59.2	–						
$\mathcal{M}_2$	30.2	−29.0	–					
$\mathcal{M}_4$	1003.2	943.9	973.0	–				
$\mathcal{M}_{4,\beta}$	1094.2	1035.0	1064.0	91.1	–			
$\mathcal{M}_5$	445.7	386.4	415.5	−557.5	−648.6	–		
$\mathcal{M}_{5,\beta}$	491.5	432.3	461.3	−511.6	−602.7	45.9	–	
$\mathcal{M}_6$	444.9	385.6	414.7	−558.3	−649.4	−0.8	−46.7	–
$\mathcal{M}_{6,\beta}$	490.9	431.6	460.7	−512.3	−603.4	45.2	−0.7	46.0

Table 6.11: *Bayesian Information Criterion (BIC) computed for all the models in three cases considered: all targets (second column), SC targets (third column), LC targets (fourth column).*

Model	BIC	$\mathrm{BIC}^{(\mathrm{SC})}$	$\mathrm{BIC}^{(\mathrm{LC})}$
$\mathcal{M}_1$	7081	655	1612
$\mathcal{M}_{1,\beta}$	5004	417	1497
$\mathcal{M}_2$	8335	901	1552
$\mathcal{M}_4$	1920	209	−393
$\mathcal{M}_{4,\beta}$	−296	−11	−574
$\mathcal{M}_5$	1859	172	723
$\mathcal{M}_{5,\beta}$	1589	173	632
$\mathcal{M}_6$	1469	173	726
$\mathcal{M}_{6,\beta}$	1468	137	635

to the Bayesian evidence in the limit of a large sample size, as it can be represented by our sample of stars ($N \gg 1$, see also Chapter 5, Section 5.6.5). Thus, the BIC reads

$$\mathrm{BIC} \equiv -2\Lambda_{\max} + k \ln N \,, \tag{6.49}$$

where k is the number of free parameters of the model considered (i.e. the dimension of the corresponding parameter space), and N the number of data points. Since $\Lambda_{\max}$ is known, the BIC can be computed straightforwardly. The resulting values of the BIC are listed in Table 6.11 for the cases of the entire sample (second column) and of SC and LC targets separately (third and fourth columns, respectively). According to the Occam's razor principle on which Bayesian model comparison relies, the most eligible model is the one that minimizes the BIC.

As highlighted by the shaded rows and columns, the model $\mathcal{M}_{4,\beta}$ is largely the favored one for all the samples considered because its evidence is always greater than those of any other model investigated in this work. In addition, the BIC confirms the result computed through the evidences.

6.5 Discussion

The analysis described in Section 6.3 and in Section 6.4 lead us to interesting results about the use of the amplitude scaling relations in asteroseismology. The main aspects of the work presented here can be divided into two groups, whose contents we discuss in the following.

6.5.1 Results from Bayesian parameter estimation

The results coming from the inference described in Section 6.3 show that:

1. for models from $\mathcal{M}_1$ to $\mathcal{M}_{4,\beta}$, the expectation values of exponent r, which we remind was introduced for converting radial velocity amplitudes into photometric ones, differ from the value $r = 2$ adopted for non-adiabatic oscillations [187, 291, 271]. On one hand, this outcome is even more evident in the case of RGs (namely the LC targets), where we found $r \simeq 10.5$ for model $\mathcal{M}_2$ (Table 6.4). Although such a high value for r could be partly explained by a very tight anti-correlation with the exponent s (see Tables 6.5 and 6.7), this seems to support the recent findings by [271] who suggest that for RGs the non-adiabatic effects become significantly important in the driving mechanism of solar-like oscillations and that, in general, the excitation model is underestimating the true amplitudes. On the other hand, in the case of SC targets, which are essentially dominated by MS stars, the mean r is not as correlated with s as for the LC targets (see

Table 6.6), and its estimated values are considerably lower ($r \approx 3$), and much closer to the value $r = 2$ adopted in previous works (Table 6.3). This again could suggest that solar-like oscillations are more adiabatic in early stages of stellar evolution.

Conversely, r becomes negative for the models that take into account the granulation power and the lifetime of the modes ($\mathcal{M}_5$ to $\mathcal{M}_{6,\beta}$), which give a different contribution to the predicted amplitudes from effective temperature (see Sections 6.2.5 and 6.2.6). The reason our estimates of r are much lower than the theoretical value is mainly related to the fact that predicted amplitudes based on the scaling relation by [188] are considerably greater than the observed ones, even for the most adiabatic case corresponding to $r = 1.5$ (e.g. see the discussions by [291, 167]). This overestimation could be explained by a missing term containing the mode masses, as proved by recent theoretical calculations by [271] for a sample of RGs observed by *CoRoT*.

Correlations among the free parameters are very tight in many cases and can be partially responsible of decreasing r, except for the case of SC targets, where no significant correlations are found

2. for the models that include the exponent s ($\mathcal{M}_1$ to $\mathcal{M}_{4,\beta}$), the expectation values derived here are fairly compatible with previous results (e.g. see [187, 126, 104, 290, 310, 19, 291, 167]). However, it is worth mentioning that when moving from the main sequence to the red giant phase, s decreases, which is apparent for all the models mentioned above. This effect can be partially explained by a compensation of the exponent r

3. the expectation values of the exponent t found for models $\mathcal{M}_4, \mathcal{M}_6$, and $\mathcal{M}_{6,\beta}$ are not far from the value $t = 1.7 \pm 0.1$ derived by [291] using a sample of cluster RGs, although correlation effects with the exponents s and r cannot be neglected and are, in some cases, quite pronounced (see Table 6.5). For model $\mathcal{M}_{4,\beta}$ instead, we found t to be very close to the value $t = 1.32 \pm 0.02$ derived by [167], who used a very similar sample of stars but with effective temperatures from KIC. These results are confirmed for both the entire sample and the sample of SC targets. The RGs sample is instead behaving differently, showing values of t significantly lower than the other two samples. This result is certainly dominated by correlations, which are stronger than those of the MS-dominated sample (see below)

4. as a qualitative result we can state that, in general, all the expectations of the free parameters involved in the amplitude scaling relations investigated here are on average rather different when comparing the MS-dominated sample to the RGs one. The correlations among the free parameters are enhancing the differences between dwarfs and giants and are likely to be stronger in the sample of more evolved stars because for red giants there is a large degree of degeneracy in the stellar fundamental properties as stars with different mass all converge along the red giant branch spanning only a small temperature range. As a consequence, this could suggest that not a single of the amplitude scaling relations discussed can be adopted to both MS and RGs simultaneously because the driving and damping mechanisms responsible of generating the observed amplitudes encounter a substantial change as the stars evolve

5. the introduction of a set of models that take into account the offset $\ln \beta$ allowed us to produce new outcomes that provide better fits for the entire set of scaling relations adopted, as it appears clear by looking at the comparison of different fits shown in Figures 6.4, 6.6, 6.7, and 6.8, together with the hint provided by our estimates of a weighted rms of the residuals listed in Tables 6.2, 6.3, and 6.4. In fact, in almost all cases, the offset

$\ln \beta$ differs from zero significantly. In addition, we note that our estimates of $\sigma^{w}_{\rm rms}$ are fairly consistent with the total relative uncertainty adopted in Eq. (6.27) for models $\mathcal{M}_4$, $\mathcal{M}_{4,\beta}$, $\mathcal{M}_6$, and $\mathcal{M}_{6,\beta}$, which are the only ones that have a separate dependence upon the mass of the stars (accounted for in the additional free parameter t). Models $\mathcal{M}_1$ to $\mathcal{M}_2$ and models $\mathcal{M}_5$ and $\mathcal{M}_{5,\beta}$ have instead a scatter in the residuals that is from 1.5 to 2 times larger than the total relative uncertainties. This suggests that the average relative uncertainty on the observed amplitudes alone is considerably smaller than the intrinsic scatter of the residuals given by Eq. (6.28), for any of the models investigated. Thus, we suppose that additional contributions to the total uncertainty are missing for models that do not take into account a separate dependence upon the mass of the stars. It is however important to note that this outcome only holds for the set of measurements and uncertainties adopted in the inference.

6.5.2 Results from Bayesian model comparison

From the model comparison applied to the amplitude scaling relations here investigated, we can say that:

1. models $\mathcal{M}_1$, $\mathcal{M}_{1,\beta}$ and $\mathcal{M}_2$ are not performing well at predicting the amplitudes along the entire range of stars considered. This is clear from the fact that their corresponding Bayes factors are the lowest among the models considered (see Tables 6.8, 6.9, 6.10), a result that is confirmed by our computation of the BIC from Eq. (6.49), which reaches its highest value for these models (see Table 6.11 in Section 6.4). The larger scatter of the residuals arising from these models (up to a factor of 2) compared to the total relative uncertainties derived from Eqs. (6.37) and (6.40) are supporting our conclusion from the model selection described in Section 6.4

2. models from $\mathcal{M}_4$ to $\mathcal{M}_{6,\beta}$, according to Bayesian principles of model selection, are certainly more favored than the others (Tables 6.8, 6.9, 6.10, and 6.11). In particular, model $\mathcal{M}_{4,\beta}$, which includes a separate dependence upon the mass of the stars (see Section 6.2.4), is strongly dominant over all the other scaling relations investigated, by at least a factor of $\sim \exp(880)$ if one takes into account the Bayes factor. Thus, model $\mathcal{M}_{4,\beta}$ is the best one according to the computation of both the evidences and the BIC for all the samples used. This suggests that the spread observed in the amplitudes is likely to be caused by a spread in mass of the stars. This is also confirmed by the consistency between $\sigma^{w}_{\rm rms}$ and the total relative uncertainties computed from Eq. (6.42).

   The models having an evidence weaker than that of $\mathcal{M}_{4,\beta}$, but still stronger than that of models from $\mathcal{M}_1$ to $\mathcal{M}_2$, are the models from $\mathcal{M}_5$ to $\mathcal{M}_{6,\beta}$, which instead include the effects of mode lifetimes and granulation power (see Section 6.2.5). These results are again confirmed for all the three samples considered from both the evidences and the BIC

3. all the models that take the proportionality term β into account, are preferred to their counterparts without this extra free parameter, with the only exception of models $\mathcal{M}_6$ and $\mathcal{M}_{6,\beta}$ for which the model comparison is inconclusive in the case of the entire sample of stars (Table 6.8). These results are confirmed from the computation of both the evidences and the BIC regardless of the evolutionary state of the stars. As a consequence, according to the set of measurements and uncertainties adopted in this work, we can conclude that the Sun is not necessarily a good reference star for this *Kepler* sample, and that only models having $\beta \neq 0$ should be considered. Equivalently, one could consider models

that do not include β as a free parameter, but having a reference amplitude for scaling given by $\beta A_{\mathrm{bol},\odot}$ (for models $\mathcal{M}_2$ to $\mathcal{M}_{6,\beta}$) or $\beta A_{\lambda,\odot}$ (for models $\mathcal{M}_1$ and $\mathcal{M}_{1,\beta}$), with β given according to the results shown in Tables 6.2, 6.3, and 6.4. Such result could also be supported by the effect of the stellar activity (occurring mostly for dwarfs), which is responsible of reducing the amplitudes of solar-like oscillations (see the discussion by [73, 167]). In fact, as noted by [167], the Sun shows on average a higher activity level than the other stars of the sample.

6.6 Conclusions

A Bayesian approach to test the set of amplitude scaling relations discussed in this work allowed us to draw new interesting conclusions on the free parameters that describe all the models.

First, as evident from the results derived in Section 6.4, our analysis strongly recommends the use of Eq. (6.17) for predicting amplitudes of solar-like oscillations, for all the stars spanning from MS to RGs. This result, together with the consistency found between our estimates of the fit quality, $\sigma^{w}_{\mathrm{rms}}$, arising from the corresponding models and the total relative uncertainties discussed in Section 6.3.4, supports the idea that a mass difference from star to star is among the main effects that produce the observed spread in the oscillation amplitudes (see also [165, 167, 291]).

Second, one should keep in mind that according to the Bayesian parameter estimation described in Section 6.3 the results arising from the inference change considerably from MS to RGs. This suggests that, in general, particular attention has to be paid when using amplitude scaling relations for samples containing a large range of stars. This behavior is already apparent from our plot of the asteroseismic HR diagram (Figure 6.2, bottom panel), which shows that, assuming the logarithmic amplitudes change linearly with the logarithm of the temperatures — as done by all the models considered — the two samples have on average a quite different slope, which is reflected mainly in a different value of the exponent r of the two samples. As a consequence, we strongly recommend further investigations of this aspect for different sets of measurements and uncertainties and also for samples of stars different than the ones used here.

Moreover, as shown in Tables (6.5), (6.6), and (6.7), the free parameters of the models are, in most of the cases, highly correlated. This may suggest to adopt different priors (e.g. non-uniform priors) for some of them in future analyses. The correlation is much more evident for RGs, possibly due to the larger range of the fundamental stellar properties in the LC sample considered.

Finally, as it is evident from many of the models investigated, the Sun is not necessarily a good average star from which to scale amplitudes for this sample of *Kepler* stars, neither for the MS-dominated sample nor for the RGs. This outcome could be partially explained by the fact that the Sun is placed at the edge of our sample of measurements, as visible from Figure 6.2 (see also the discussion in Section 6.3.2), and that its activity level is on average higher than that of the other stars [73, 167]. Nevertheless, we stress that this result also relies on our choice of the set of uncertainties adopted for computing the likelihood. In fact, as discussed at the end of Section 6.5.1, the uncertainties used are underestimated in most of the cases, even when considering errors on the variables of the models. This aspect may in fact produce a misleading solution to the inference problem and to the model selection. Since both the inference and the model comparison presented here rely on the assumption that these uncertainties are themselves error-free, in future work, it will be worth investigating how uncertainties affect the results by performing uncertainty-independent analyses that do not necessarily assume the adoption of a Gaussian likelihood function (see e.g. the Bayesian inference discussed in Chapter 7).

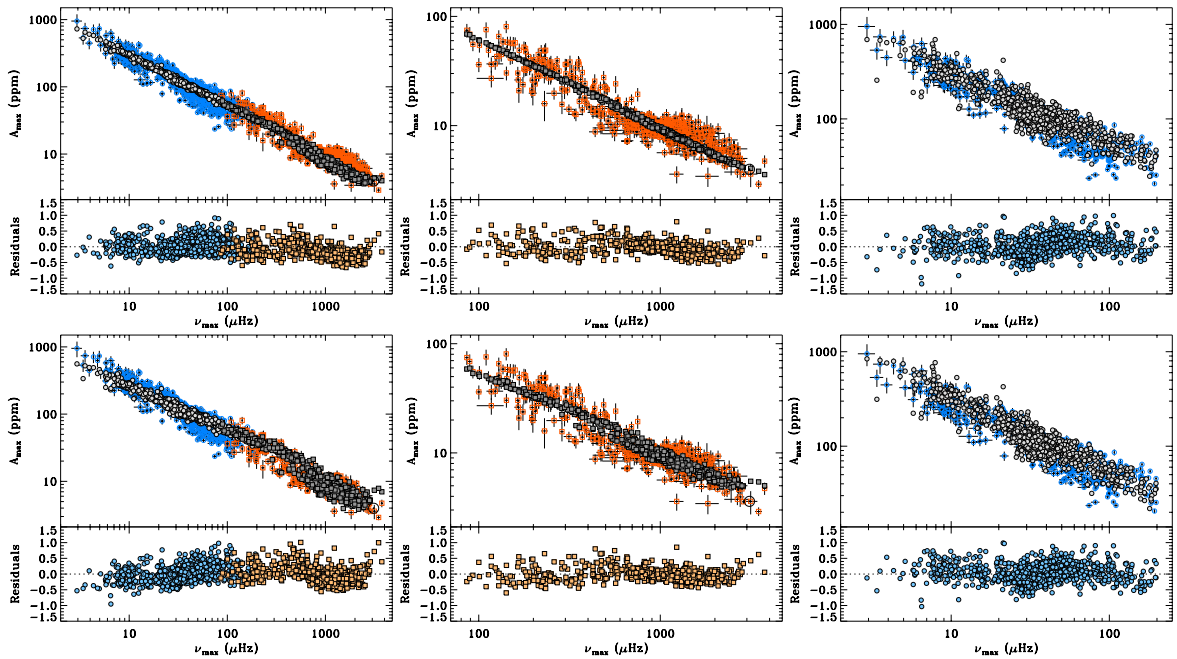

Figure 6.4: *Top panels: Predicted amplitudes (filled light-gray circles for LC targets and filled dark-gray squares for SC targets) for model $\mathcal{M}_1$ plotted against $\nu_{\rm max}$ in the three cases considered (all sample in left panel, SC targets only in middle panel, LC targets only in right panel). The expectation values of the free parameters reported in Tables 6.2, 6.3, 6.4, have been adopted for plotting the predicted amplitudes. Observed amplitudes are shown in background for both SC (open orange squares) and LC (open blue circles) targets, together with 1-σ error bars shown on both quantities. The Sun's symbol ($\odot$) is added for comparison, where $A_{650,\odot} = 3.98\,ppm$ ($\lambda = 650\,nm$). The residuals arising from the difference between the logarithms of observed and predicted amplitudes, according to Eqs. (6.36) and (6.38), are also plotted (same symbols). Bottom panels: Same description of the top panels but for the model $\mathcal{M}_{1,\beta}$.*

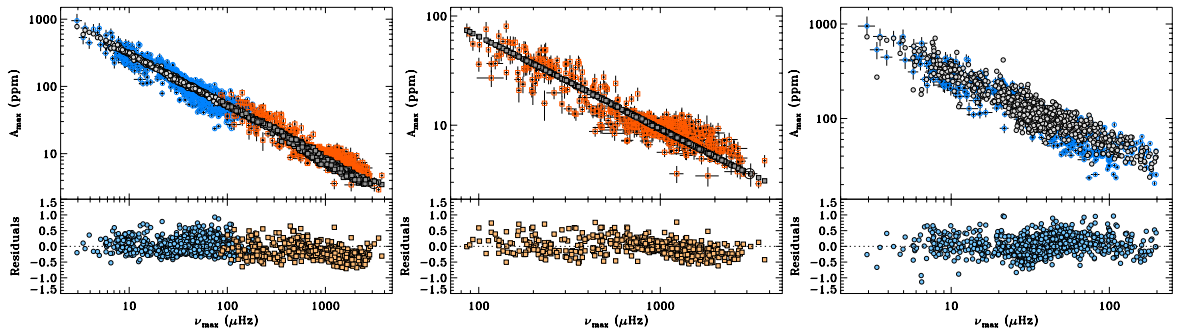

Figure 6.5: *Same description as for Figure 6.4 but for the model $\mathcal{M}_2$.*

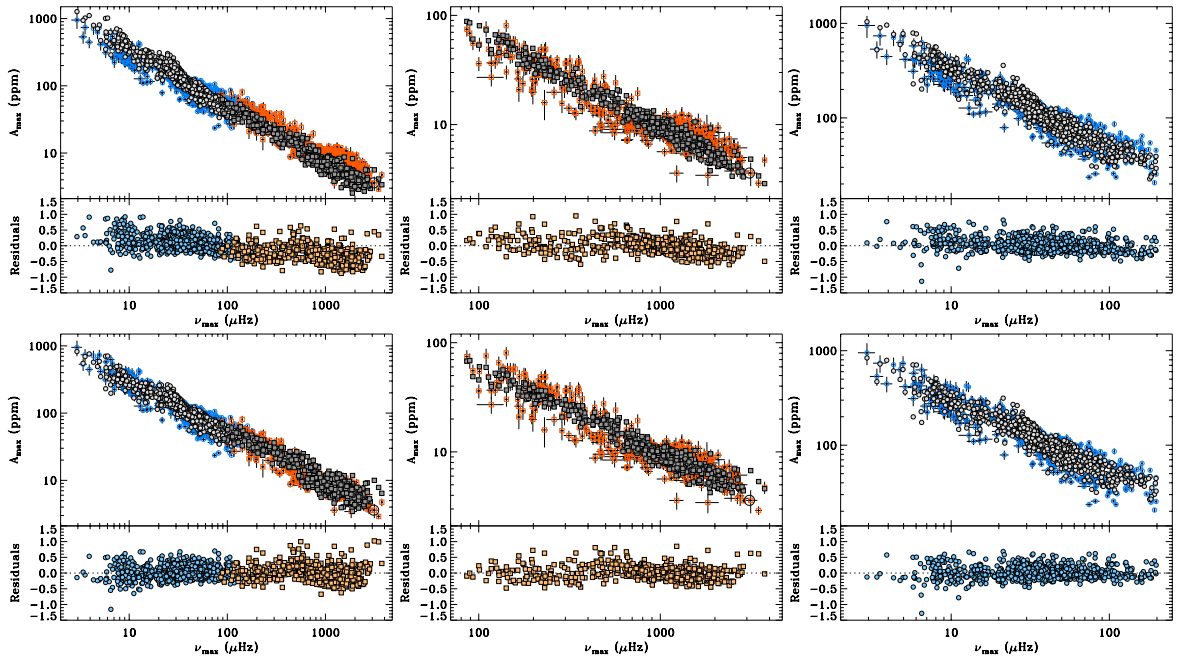

Figure 6.6: *Same description as for Figure 6.4 but for the models $\mathcal{M}_4$ (top panels) and $\mathcal{M}_{4,\beta}$ (bottom panels).*

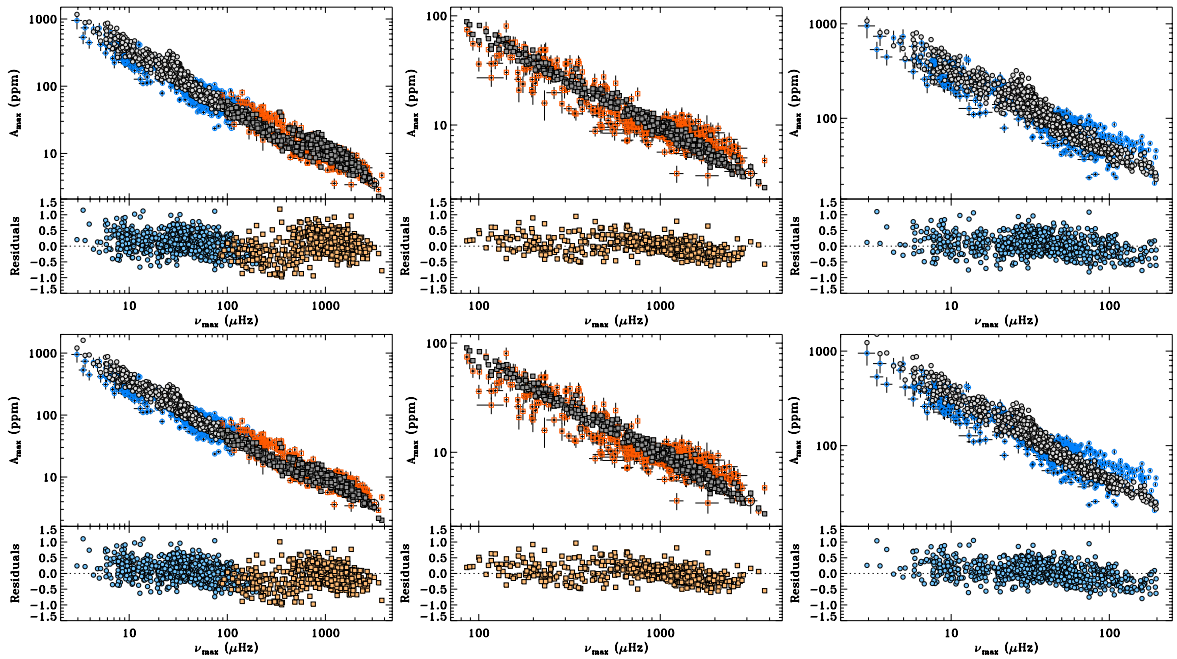

Figure 6.7: *Same description as for Figure 6.4 but for the models $\mathcal{M}_5$ (top panels) and $\mathcal{M}_{5,\beta}$ (bottom panels).*

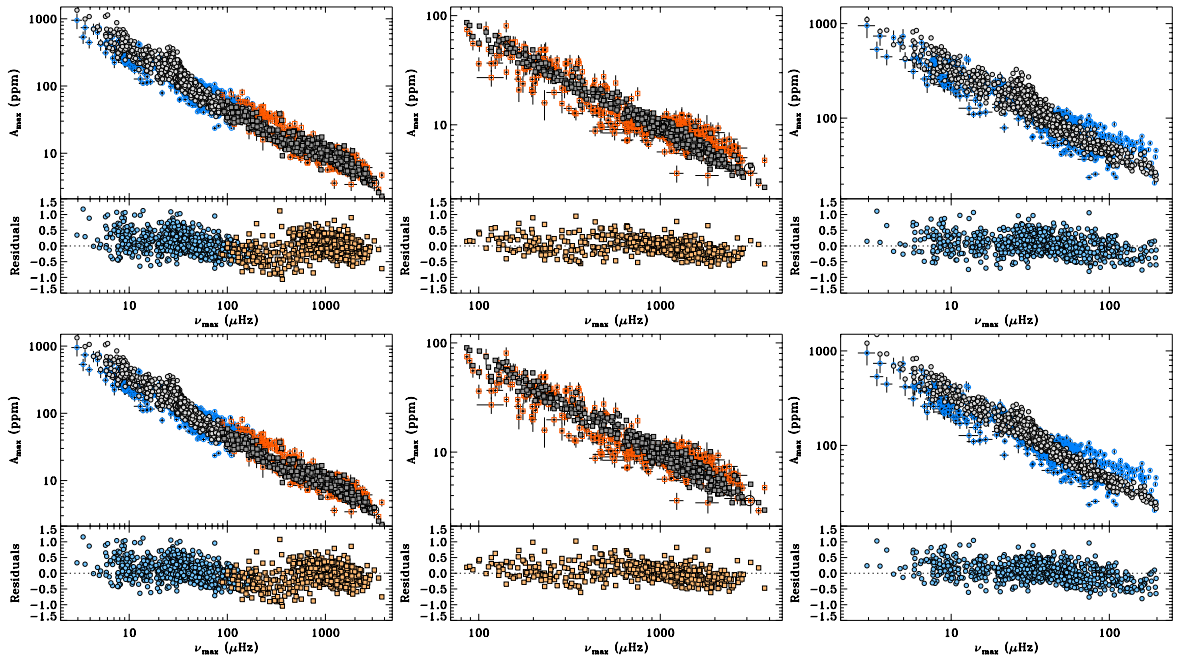

Figure 6.8: *Same description as for Figure 6.4 but for the models* $\mathcal{M}_6$ *(top panels) and* $\mathcal{M}_{6,\beta}$ *(bottom panels).*

Fröhlich, H.-E., et al., *Magnetic activity and differential rotation in the young Sun-like stars KIC 7985370 and KIC 7765135*, 2012, A&A, 543, 146 [118]

7 Differential Rotation in Sun-like Active Stars

In the Sun, magnetic activity is thought to be produced by a global-scale dynamo action arising from the coupling of convection and rotation [251, 283]. Young Sun-like stars are rotating faster than the Sun and display a much higher level of magnetic activity at all atmospheric layers, which is likely due to a stronger dynamo action. They also show different manifestations of activity compared to the Sun, such as bigger and long-living spots in their photospheres, active longitude belts, absence or a different behavior of activity cycles, highly energetic flares, etc. These differences are likely related to the dynamo mechanism, which is operating in rather different conditions in young stars mainly regarding their rotation rate and internal structure. Understanding the properties of young Sun-like stars and, particularly, their activity and rotation, is crucial to trace the Sun and its environment back to the first evolutionary stages.

In fact the properties of the magneto-convection in these stars seem to be strongly influenced by the Ω-effect that produces characteristic "wreaths" on large scales [248]. Although strong latitudinal variations of the differential rotation can be obtained by means of the combined role of thermal wind balance and geostrophy, the results of numerical simulations seem to be strongly dependent on the Reynolds number of the flow. The situation is at variance with the dynamo action in main sequence solar-type stars, where the role of the tachocline is instead essential in producing the α effect [99, 46, 49].

It is still unclear whether a strong latitudinal differential rotation is common among fast-rotating stars. The study by [214] reports values of the absolute differential rotation $\mathrm{d}\Omega$ in the range 0.08–0.45 rad d^{-1} for a sample of rapid rotators similar to and slightly more massive than the Sun. Despite the spread of values, it seems that $\mathrm{d}\Omega$ is in any case larger than in the Sun. The measures of absolute differential rotation in a large sample of F- and early G-type stars through the Fourier transform technique [264, 263] show no indication of the decrease in this parameter with the rotation period, rather the highest values of $\mathrm{d}\Omega$ are encountered for periods between 2 and 3 days. On the other hand, some recent calculations predict a moderate differential rotation, comparable to that of the Sun, also for a Sun-like star rotating 20 times faster (e.g. [196]).

Moreover, $\mathrm{d}\Omega$ seems also to be a function of the stellar mass for main-sequence stars, increasing with their effective temperature, as shown, e.g., by [15]. One of the largest values of differential rotation for a star noticeably cooler than the Sun was found by [115] for KIC 8429280, a 50 Myr-old K2-type star, from the analysis of the light curve collected by the NASA *Kepler* spacecraft.

The highly precise photometry of *Kepler* [50, 193] (Chapter 3) coupled with the long and virtually uninterrupted coverage makes these data unique for the study of photospheric activity and differential rotation in late-type stars, as it has been shown by [115] based on the analysis of spotted stars from *Kepler* data. However, whether star spots are indeed the best tracers of

the surface rotation or not is still a matter of debate (see [194] for a different point of view).

In this last chapter we shall briefly describe the study done by [118] who used a Bayesian approach to the *Kepler* light curves of the stars KIC 7985370 and KIC 7765135, which are introduced in Section 7.1. In fact, the spot modeling analysis of these targets done by means of Bayesian methods represents part, although a little one, of the research addressed in this dissertation. The Bayesian inference for the spot modeling applied to the *Kepler* light curves is described in Section 7.2 while the results are shown in Section 7.3, with relevance to the problem of differential rotation. Lastly, a discussion and conclusions about the results are drawn in Section 7.4.

7.1 Data and Observations

As for KIC 8429280 [115], the two new targets KIC 7985370 (HD 189210, 2MASS J19565974+4345083, TYC 3149-1571-1) and KIC 7765135 (2MASS J19425057+4324486 = TYC 3148-2163-1) were selected as active stars from their optical variability and from the cross-correlation of the ROSAT All-Sky Survey (RASS; [311, 312]) with Tycho and Hipparcos catalogues [254]. With $V = 10^m.0$ and $11^m.8$, respectively, both stars are relatively bright ones in the *Kepler* field of view. Both of them were recently reported as variable stars by [255], who searched for bright variable stars in the *Kepler* field of view with ASAS3-North station. The variability of KIC 7985370 could be due to a rotational modulation according to [305] that rely on the first two quarters of *Kepler* data. The *Kepler* light curves readily show these stars as rotationally variable with a period of about 2–3 days, which is typical of G-type stars in the Pleiades cluster (age ~ 130 Myr, [16]). The estimates of their atmospheric parameters reported in the Kepler Input Catalog (KIC) [60], which are based on Sloan photometry (for a revised temperature scale cf. [256]), suggested to us that these objects were similar to the Sun.

The analysis of the optical spectra collected by us confirmed that the stars are nearly identical to the Sun, but much younger and as such deserving a detailed investigation. Applying the same techniques used by [115], the basic stellar parameters, the chromospheric activity, and elements abundances were derived for these stars (see [118] for details). The kinematics of these stars was also discussed by the same authors.

7.2 Bayesian Spot Modeling

All the available public *Kepler* LC light curve ($\Delta t \sim 30$ min, [174]), spanning from 2009 May 2 to 2009 December 16, was analyzed. It covers altogether 229 days and corresponds to the observing quarters 0–3 (Q0–Q3), with the largest gap, about 4.5 days, appearing between Q1 and Q2.

To remove systematic trends in the *Kepler* light curves associated with the spacecraft, detector, and environment, and to prepare them for the analysis of star spots that we will describe below, we used the software KEPCOTREND[1]. This procedure is based on Cotrending Basis Vectors (CBV), which are calculated (and ranked) through singular value decomposition and describe the systematic trends present in the ensemble flux data for each CCD channel. We used the first two basis vectors for Q0 data, while from three to five CBV were adopted for the correction of longer data sets such as Q1, Q2, and Q3.

In order to check how the data rectification accomplished with KEPCOTREND is reflected in the outcome, the spot modelling has been done twice: with the rectified data (Case A) as well as with the original data (Case B).

[1] http://keplergo.arc.nasa.gov/ContributedSoftwareKepcotrend.shtml.

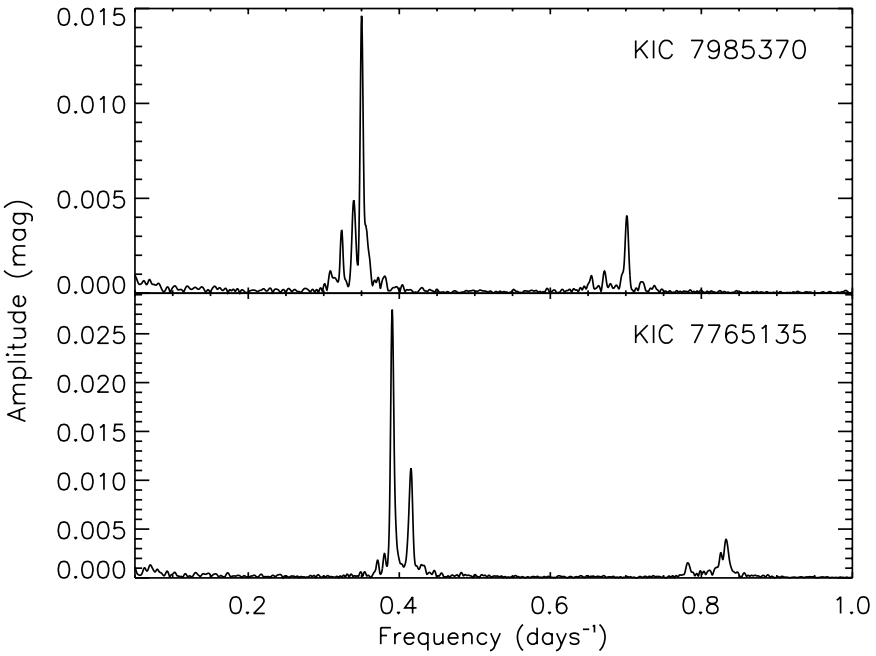

Figure 7.1: *Cleaned periodograms of the Kepler Q0+Q1+Q2+Q3 time-series for KIC 7985370 (upper panel) and KIC 7765135 (lower panel). Image courtesy of Antonio Frasca.*

The power spectra of the *Kepler* time-series, cleaned by the spectral window according to [266], are displayed in Figure 7.1. The lower panel of Figure 7.1 clearly shows two main peaks for KIC 7765135, which are close in frequency (0.391 and 0.414 d^{-1}). The corresponding periods are 2.560 ± 0.015 and 2.407 ± 0.014 days, respectively. The period errors are from the FWHM of the spectral window. The low-amplitude peaks at frequency of $\approx 0.8\,\mathrm{d}^{-1}$ are overtones of the two main peaks. As visible from the upper panel of Figure 7.1, the structure of the peaks for KIC 7985370 is more complex, with the maximum corresponding to 2.856 ± 0.019 days and a second peak, blended with the first one on its low-frequency side, at 2.944 days. A third small peak corresponding to $P = 3.090$ days is also visible.

Such a double- or multiple-peaked periodogram hints at differential rotation. As [199] predicted, a photometric time series, if sufficiently accurate ($\Delta F/F = 10^{-5} - 10^{-6}$) as it is the case of *Kepler* photometry, may reveal a Sun-like latitudinal differential rotation.

An estimate of the inclination of the rotation axis with respect to the line of sight is very useful to constrain the spot model. With $v \sin i$, stellar radius R, and rotation period P known, the inclination of the rotation axis follows from

$$\sin i = \frac{(v \sin i) \cdot P}{2\pi R} \,. \tag{7.1}$$

In the absence of an accurate parallax value the stellar radius cannot be derived from the effective temperature and luminosity.

If we adopt the radius for a ZAMS star with the effective temperature of our targets ($T_{\mathrm{eff}} = 5800\,\mathrm{K}$), $R \approx 1.1\,R_{\odot}$, we get $\sin i = 0.967$ ($i = 75° \pm 15°$) for both stars. However, as stated in [118], Section 3.3, the lithium content cannot provide a firm lower limit for the ages of these stars, which could also be as young as a few 10 Myr (post-T Tauri phase). Thus, allowing for such a young age, a $T_{\mathrm{eff}} = 5800\,\mathrm{K}$ is reached by a star of $1.5\,M_{\odot}$ at 10 Myr with a

radius of about $2\,R_\odot$ according to the evolutionary tracks by [274]. In this case, an inclination of about 30° is deduced.

7.2.1 Bayesian photometric imaging

The method is basically that adopted by [115, 118]. However, as the light curves are now significantly longer, the introduction of further free parameters was inevitable. The latitudinal dependence of rotation frequency $\Omega(\beta)$, Eq. (7.2), now contains a $\sin^4\beta$-term and, more important, the prescription for spot area evolution is much more detailed. Furthermore, the likelihood function, Eq. (7.3), is generalized by taking into account an unknown linear trend in the data. To tackle the problem of strong correlations between some parameters, an essential new ingredient is the usage of an orthogonalized parameter space, where the steered random walk of the Markov chains is performed. For the sake of clarity, due to these new features, the basics of the method are explained in this section, although they can also found in [115]. A full account of the method will be described in a forthcoming paper [119].

A light curve fitting that represents spots as dark and circular regions has the advantage of reducing the dimensionality of the problem and to promptly provide us with average parameters (area, flux contrast, position, etc.) for each photospheric active region. Of course, there are other techniques, which are based on different assumptions, to reconstruct surface features photometrically as the inversion of *Kepler* light curves done by, e.g., [59].

Our aim is to present a low-dimensional spot model, with few spots only, that fits reasonably well the data regardless to very low-amplitude details that require a high degree of complexity. As widely discussed in Chapter 5 and also with the analysis presented in Chapter 6, in a Bayesian context this claim could be even quantified by ideally estimating the Bayesian evidence, i.e. the integral over the posterior probability distribution. It would provide a measure of the probability of a n-spot model and, therefore, allow one to constrain the number n of spots that are really needed. For numerical reasons we are compelled to resort instead to the less demanding Bayesian Information Criterion (BIC) by [273], which has been used also for solving the model comparison discussed in Chapter 6, Section 6.4. As discussed in Chapter 5, Section 5.6.5, the BIC or any other related criterion expresses Occam's razor principle in mathematical terms without the need to compute the evidence.

Unfortunately, we have to admit that – due to the unprecedented accuracy of the *Kepler* data – we could not reach this goal with only seven or nine spots. There is obviously more information in the data than our most elaborate model is able to account for.

Dorren's analytical star-spot model [101], generalized to a quadratic limb-darkening law, was used. The two coefficients are taken from the tables by [82] for a microturbulence velocity of $\xi = 2\,\mathrm{km\,s^{-1}}$ and are used for both the unperturbed photosphere and the spots.

Four parameters describe the star as a whole: one is the cosine of the inclination angle i, while the other three parameters (A, B and C) describe the latitudinal dependence of the angular velocity. With β being the latitude value, the angular velocity Ω is parameterized by a series expansion using Legendre polynomials:

$$\begin{aligned}\Omega(\beta) = A &+ 3B\left(5\sin^2\beta - 1\right)/2 + \\ &+ C\left(315\sin^4\beta - 210\sin^2\beta + 15\right)/8\,. \end{aligned} \tag{7.2}$$

The equatorial angular velocity is $\Omega_{\mathrm{eq}} = A - 3B/2 + 15C/8$ and the equator-to-pole differential rotation $\mathrm{d}\Omega = 15B/2 + 105C/8$. In the case of the equator rotating faster than the poles $\mathrm{d}\Omega$ is negative. In what follows the minus sign is suppressed, and only the absolute value $|\mathrm{d}\Omega|$ is given. Both stars are definitely rotating like the Sun.

As in the work done by [115], all star spots have the same intensity κ relative to the unspotted photosphere and are characterized by two position coordinates (latitude and initial longitude) and by their radius. All these are free parameters in the model. Other spot parameters are the rotation period, which defines the spot longitude at any time and is tied to the latitude via Eq. (7.2). The hemisphere, to which a spot belongs to, is to be found by trial and error. Further parameters describe the spot area evolution.

As our photometric analysis mainly aims at estimating the level of surface differential rotation, our focus is on *long-lived* spots. Longevity of star spots is at the heart of our approach. In order to obtain, in view of the extraordinary length of the time series, a satisfying fit, more freedom has been given to spot area evolution with respect to [115]. It is now parameterized by up to eight parameters.

Spot area is given in units of the star's cross-section. Area evolution is assumed to go basically linearly with time. The underlying physical reason is that then, at least in the case of a decaying spot, the slope of the area–time relation is somehow related to the turbulent magnetic diffusivity. With the aim of enhancing flexibility and to describe the waxing and waning of a spot, three consecutive slope values are considered. The time derivative of spot area is then a mere step function over time. Step height measures the increase/decrease of area per day. So, there are six free parameters: three slope values, two dates of slope change, and the logarithm of spot area at some point of the time series. We have done even a little bit more. In order to prevent sharp bends in spot area evolution, some smoothing is introduced. Each date where the slope changes is replaced by a time interval within which the slope is linearly interpolated between the two adjacent values. This makes the second time derivative of spot area a mere step function of time, described by six parameters. To get the integrated area itself as function of time two constants of integration enter, thus bringing the number of free parameters to describe a spot's area evolution to a total of eight.

In addition to the free parameters of the model there are derived ones, the marginal distributions of which are of interest. An example is the rotational period. It follows from the longitudes of the spot centre at the beginning of the time series and at its end. All parameters are estimated in a Bayesian manner, i.e. their mean values as well as the corresponding uncertainties follow straightforwardly from the data alone (see Chapter 5).

To maintain a flat prior distribution in parameter space, all dimensional parameters like periods or spot radii must actually be described by their logarithms (this has been done similarly in Chapter 6). Only then the posterior probability distribution for a period will be consistent with that of a frequency and likewise the posterior for a radius with that of an area, i. e. it does not matter whether one prefers periods to frequencies or radii to areas.

The likelihood function, Eq. (7.3) assumes that the measurement errors have a Gaussian distribution in the magnitude domain. This is justified as long as the signal-to-noise ratio does not vary with changing magnitude, as it is for our data that span a full variation range of less than 0.1 magnitudes. It has the invaluable advantage that the likelihood function can be analytically integrated over measurement error σ, offset c_0, and linear trend d_0. To perform the integration over σ one has to use Jeffreys' prior $\propto \sigma^{-1}$ [184]. The resulting *mean* likelihood depends on spot-modelling parameters $p_1 \dots p_M$ only. It takes into account all possible error values, offsets and linear trends. By multiplying it with the prior, assumed constant in parameter space, one gets the posterior density distribution. All interesting quantities, parameter averages and confidence intervals, are then obtained by marginalization.

With the N magnitude values d_i measured at times t_i, their standard deviations σ_i, the model magnitudes $f_0(t_i, p_1 \dots p_M)$, offset c_0, and trend d_0, the likelihood function is given by

$$\mathcal{L}\left(\sigma, c_0, d_0, p_1 \dots p_M; d_i\right) =$$

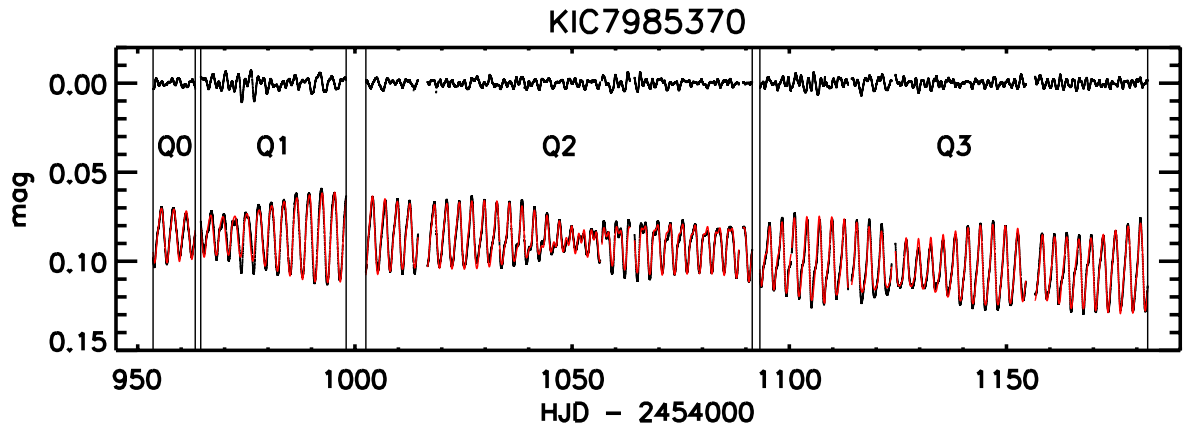

Figure 7.2: *Kepler light curve with best fit (solid red line, second Case-A solution of Table 7.1) overplotted. The residuals, shown at the top, are ±2.14 mmag. As for the case of KIC 7765135, the residuals are not homogeneous from one part of the light curve to another. Image courtesy of Antonio Frasca.*

$$\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{(d_i - f_0\,(t_i, p_{1\ldots M}) - c_0 - d_0 \cdot (t_i - t_0))^2}{2\sigma_i^2}\right]. \tag{7.3}$$

We set $\sigma_i = s_i \cdot \sigma$, with relative errors s_i being normalized according to

$$\sum_{i=1}^{N} \frac{1}{s_i^2} = N\,. \tag{7.4}$$

Parameter estimation by sampling the parameter space has been done by the Markov Chain Monte Carlo (MCMC) method [259] (see also Chapter 5).

Often parameter values are highly correlated. As MCMC performs best in an orthogonalized parameter space, all parameters have been converted by a principal component analysis using Singular Value Decomposition [259] (see Appendix D). Each parameter in this abstract space is linearly dependent on all of the original parameters. The reconstruction of the original parameter values can be done exploiting a subspace of that orthogonalized parameter space. The dimension of that subspace, the number of degrees of freedom, proves lower by roughly one third or even more than the number of original parameters.

7.3 Results on Differential Rotation

In this section we present the results derived by [118] by means of the Bayesian spot modeling introduced above. The description is divided into two parts, each one concerning one of the stars investigated.

KIC 7985370

We have identified eleven gaps longer than an hour and two additional small jumps in the light curve (Figure 7.2). The data set was accordingly divided into 14 parts. Each part has been assigned its individual error level, off-set and, in Case B (i.e. non-rectified data), linear trend. Hence, the likelihood, Eq. (6.27), is the product of 14 independent contributions.

KIC 7985370's inclination value i appears ill-defined by the spot model applied to the *Kepler* photometry. Indeed, with only six spots the MCMC results in very dark spots ($\kappa \approx 0$)

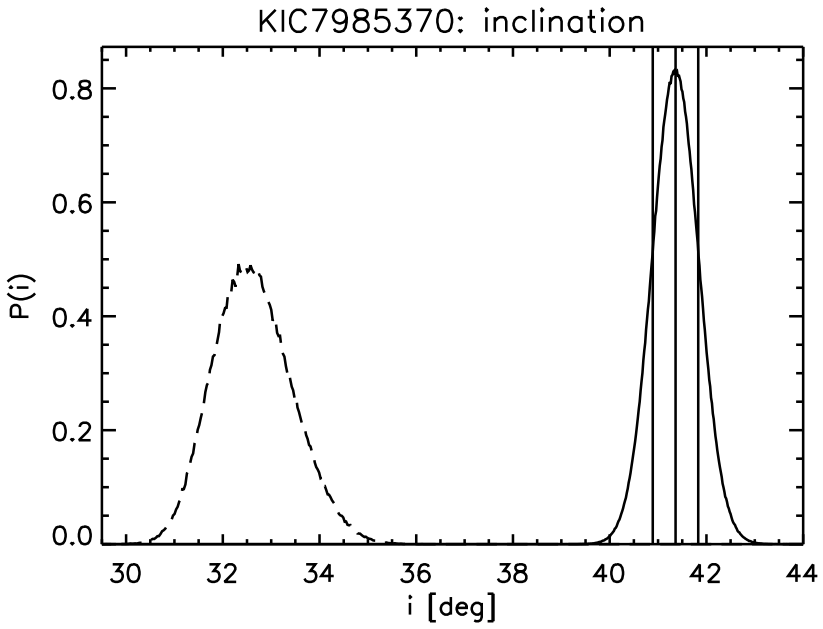

Figure 7.3: *Determination of the stellar inclination from Kepler photometry. Mean and 68-per-cent confidence region are marked by vertical lines (Case A only). Dashed: The corresponding marginal distribution for the original data with linear trends removed (Case B). Image courtesy of Antonio Frasca.*

at very low inclination ($i \approx 10°$). But even for these unrealistic solutions the equator-to-pole differential rotations was $0.18\,\mathrm{rad\,d^{-1}}$. Only with seven spots and allowing for enough spot evolution we arrived at acceptable inclination values (Figure 7.3) and spot intensities. If the inclination is fixed to the spectroscopically derived value of $i = 75°$ the residuals are rather high, $\pm 2.46\,\mathrm{mmag}$, exceeding the residuals of our best solution ($\pm 2.14\,\mathrm{mmag}$) by far. Nevertheless, details of the solution with fixed inclination are also included in Table 7.1, where the results are presented.

Improving the 7-spot solution by adding an eight spot leads formally to a better fit. As the new spot proves to be ephemeral, lasting only six rotations, it neither constrains the differential rotation nor adds any significant insight (one can always get a better result by adding short-lived features).

The marginal distributions of the seven spot frequencies (Case A only), combined into one plot, are shown in Figure 7.4.

From the three parameters describing the star's surface rotation, A, B and C, the equatorial rotational period and the equator-to-pole differential rotation (Figure 7.5) follow. The latter amounts to $0.1774^{+0.0004}_{-0.0005}\,\mathrm{rad\,d^{-1}}$ (Case A) and $0.1729 \pm 0.0002\,\mathrm{rad\,d^{-1}}$ (Case B), respectively. The difference is significant, considering the formal errors, albeit very small. In the case of fixed inclination ($i = 75°$) the differential rotation would be slightly enhanced, $0.1839 \pm 0.0002\,\mathrm{rad\,d^{-1}}$.

The spot area evolution is depicted in Figure 7.6. The sudden rise of spot # 7 seems to be an artifact. It falls into the gap between the end of Q2 data and the beginning of Q3 data. On the other hand, the sudden disappearance of spot # 3 is not related to any switching from one part of the light curve to the next one. The fall in area of spot # 1 at the end of the time series is somehow mirrored in an increase in the size of spot # 5. Maybe this indicates a flaw due to

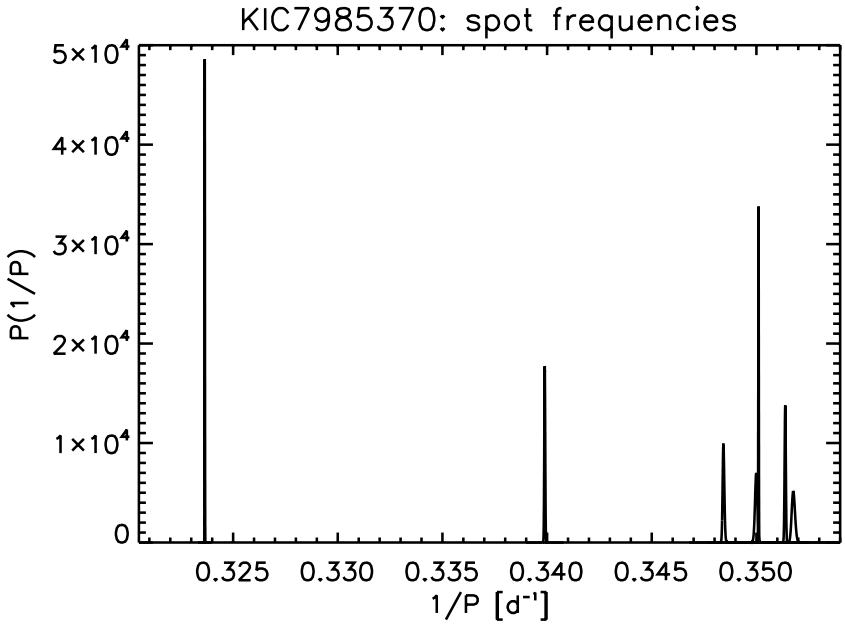

Figure 7.4: *All seven marginal PDFs (Case A) of the spot frequencies. The three frequencies (0.324, 0.340, and 0.350 d^{-1}) seen in the low-resolution Fourier spectrum (Figure 7.1) are confirmed by the results of the spot model performed by [118]. Image courtesy of Antonio Frasca.*

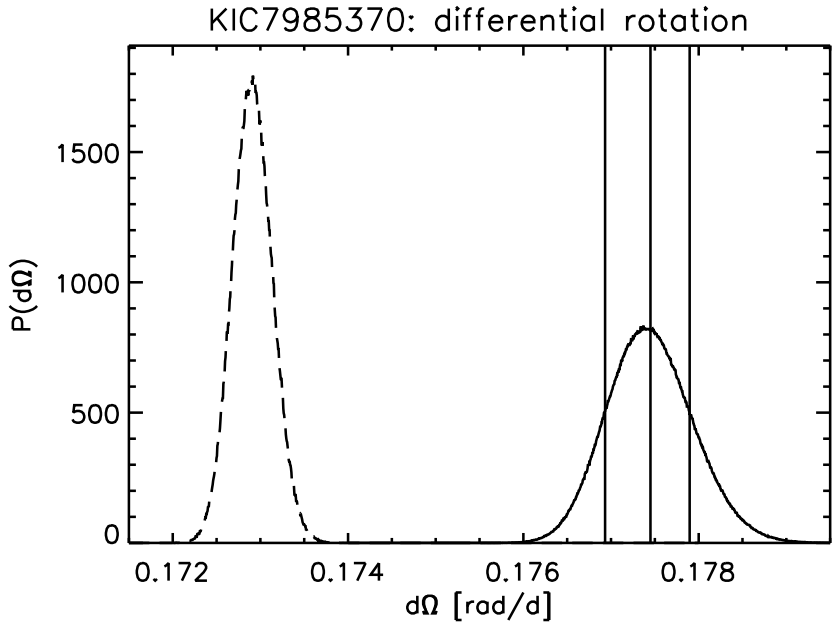

Figure 7.5: *Equator-to-pole differential rotation of the star. Mean and 68-per-cent confidence region are marked by vertical lines (Case A only). Dashed: The corresponding marginal distribution for the original data with linear trends removed (Case B). Image courtesy of Antonio Frasca.*

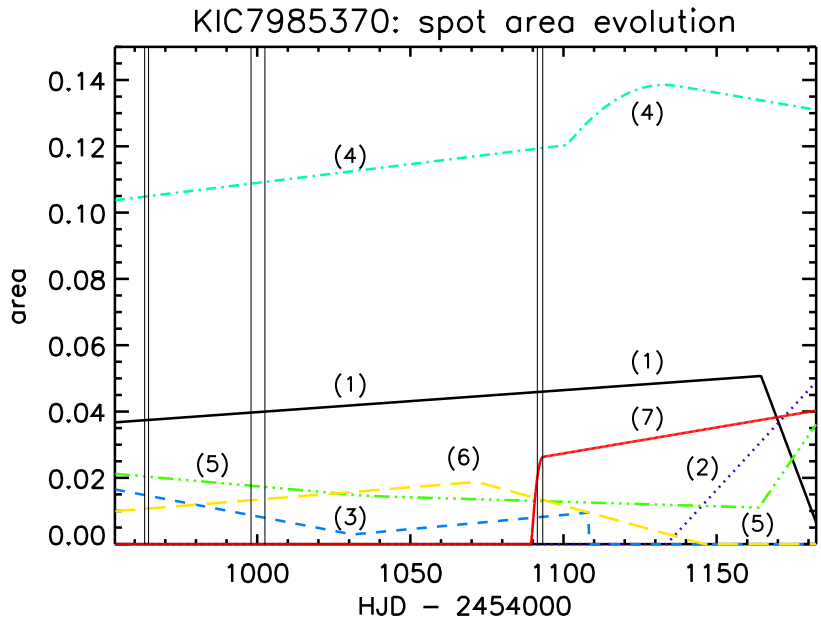

Figure 7.6: *Spot area evolution (Case A). Area is in units of the star's cross-section. Vertical lines mark the boundaries of the Q0 to Q3 quarters of data. A number in parenthesis indicates the spot number. Image courtesy of Antonio Frasca.*

too much freedom in describing spot area evolution. Expectation values with 1–σ confidence limits for various parameters are also quoted in Table 7.1.

One should be aware that there is more than one solution for each case. The second Case-A solution presented in Table 7.1 is the one that has the lowest residuals found so far. There is an other well-relaxed 7-spot solution with slightly larger residuals nearby in parameter space. In that solution the fastest spot (# 2), coming into existence near the end of the time series at JD∼2455135, is located at a more southern latitude of −21°, resulting in a slightly increased differential rotation. All other spots are virtually unaffected. Further details of this second solution are given in Table 7.2.

KIC 7765135

We have identified eleven gaps longer than an hour and three additional small jumps in the light curve (Figure 7.7). The data set was accordingly divided into 15 parts. Each part has been assigned its individual error level, off-set and, in Case B (i. e. non-rectified data), linear trend. Hence, the likelihood function (Eq. 7.3) is the product of 15 independent contributions.

As the inclination is photometrically ill-defined, we fixed it to the spectroscopically derived value of $i = 75°$.

Despite two spots more, the residuals, ±2.35 mmag, exceed those of the seven-spot model of KIC 7985370 (±2.14 mmag). This is not due to the fainter magnitude of KIC 7765135 compared to KIC 7985370, because the photometric uncertainties are typically 0.047 mmag for the former and 0.022 mmag for the latter. The reason may be that three of the nine spots are definitely short-lived with a life span as low as two months (cf. Figure 7.10), which is less than twice the lapping time of 38 days between the fastest and the slowest spot. We have to admit that dealing with nine spots goes already to the limit of the MCMC technique since the method's relaxation time becomes prohibitively long. The marginal distributions of the nine

Table 7.1: *Three 7-spot solutions for KIC 7985370. Listed are* expectation *values and 1-σ confidence limits. Latitudes β are derived from the assumed law of differential rotation, Eq. (7.2). Periods P are given in days, the spot intensity κ is in units of the intensity of the unspotted surface. The ratio C/B measures the deviation from a pure* $\sin^2$*-law of differential rotation. The differential rotation* $\mathrm{d}\Omega$ *(rad d*$^{-1}$*) is the equator-to-pole value of the shear. Residuals are in mmag. Case A refers to rectified data, Case B to non-rectified one. In order to get the Case-B solution the Case-A solution has been taken as a starting point for the MCMC parameter estimation. In the first Case-A solution the inclination is fixed to* $i = 75°$*. The second spot is near the equator, therefore, the hemisphere it belongs to is doubtful.*

parameter			Case A		Case A		Case B	
inclination		i	75°.0	*fixed*	41°.4	$^{+0.5}_{-0.5}$	32°.6	$^{+0.7}_{-0.9}$
1st	latitude	β_1	34°.0	$^{+0.1}_{-0.1}$	29°.4	$^{+0.4}_{-0.4}$	22°.6	$^{+0.5}_{-0.6}$
2nd	latitude	β_2	−10°.0	$^{+0.6}_{-0.9}$	−6°.9	$^{+0.9}_{-0.9}$	3°.3	$^{+0.1}_{-0.1}$
3rd	latitude	β_3	32°.2	$^{+0.2}_{-0.2}$	29°.9	$^{+0.3}_{-0.3}$	27°.7	$^{+0.4}_{-0.4}$
4th	latitude	β_4	86°.8	$^{+0.1}_{-0.1}$	87°.5	$^{+0.1}_{-0.1}$	87°.8	$^{+0.1}_{-0.1}$
5th	latitude	β_5	53°.8	$^{+0.1}_{-0.1}$	53°.6	$^{+0.1}_{-0.1}$	51°.1	$^{+0.2}_{-0.2}$
6th	latitude	β_6	35°.8	$^{+0.2}_{-0.2}$	35°.8	$^{+0.3}_{-0.2}$	33°.0	$^{+0.3}_{-0.3}$
7th	latitude	β_7	29°.6	$^{+0.3}_{-0.2}$	19°.9	$^{+0.9}_{-0.9}$	10°.0	$^{+0.5}_{-0.6}$
1st	period	P_1	2.8581	$^{+0.0001}_{-0.0001}$	2.8563	$^{+0.0001}_{-0.0001}$	2.8572	$^{+0.0001}_{-0.0001}$
2nd	period	P_2	2.8350	$^{+0.0003}_{-0.0003}$	2.8428	$^{+0.0007}_{-0.0006}$	2.8475	$^{+0.0002}_{-0.0002}$
3rd	period	P_3	2.8541	$^{+0.0004}_{-0.0004}$	2.8572	$^{+0.0004}_{-0.0005}$	2.8644	$^{+0.0007}_{-0.0005}$
4th	period	P_4	3.0895	$^{+0.0001}_{-0.0001}$	3.0898	$^{+0.0001}_{-0.0001}$	3.0888	$^{+0.0001}_{-0.0001}$
5th	period	P_5	2.9382	$^{+0.0002}_{-0.0002}$	2.9421	$^{+0.0002}_{-0.0002}$	2.9417	$^{+0.0003}_{-0.0002}$
6th	period	P_6	2.8629	$^{+0.0003}_{-0.0004}$	2.8700	$^{+0.0004}_{-0.0003}$	2.8754	$^{+0.0005}_{-0.0004}$
7th	period	P_7	2.8490	$^{+0.0003}_{-0.0002}$	2.8460	$^{+0.0002}_{-0.0002}$	2.8487	$^{+0.0002}_{-0.0002}$
spot intensity		κ	0.437	$^{+0.005}_{-0.004}$	0.396	$^{+0.006}_{-0.006}$	0.406	$^{+0.012}_{-0.012}$
equ. period		P_{eq}	2.8347	$^{+0.0003}_{-0.0003}$	2.8427	$^{+0.0007}_{-0.0006}$	2.8474	$^{+0.0003}_{-0.0002}$
deviation		C/B	0.28		0.28		0.21	
diff. rotation		$\mathrm{d}\Omega$	0.1839	$^{+0.0002}_{-0.0002}$	0.1774	$^{+0.0004}_{-0.0005}$	0.1729	$^{+0.0002}_{-0.0002}$
residuals			±2.46		±2.14		±2.12	

Table 7.2: *A second pair of 7-spot solutions for KIC 7985370. The meaning of the entries is the same as in Table 7.1, i.e. periods are in days, the differential rotation in rad d*$^{-1}$*, and the residuals in mmag.*

parameter		Case A		Case B	
equ. period	P_{eq}	2.8202	±0.0002	2.8209	±0.0002
diff. rotation	$\mathrm{d}\Omega$	0.1943	±0.0002	0.1933	±0.0002
residuals		±2.20		±2.21	

spot frequencies (Case A only), combined into one plot, are shown in Figure 7.8.

From the three parameters describing the star's surface rotation, A, B and C, the equatorial rotational period and the equator-to-pole differential rotation (Figure 7.9) follows. The latter amounts to 0.1760 ± 0.0003 rad d^{-1} (Case A) and $0.1774^{+0.0003}_{-0.0004}$ rad d^{-1} (Case B), respectively. As for KIC 7985370 the difference is small, but nevertheless significant.

The level of differential rotation does not depend on the number of spots considered. Neglecting the three short-lived spots, i.e. considering a six-spot model, would result in an equator-to-pole shear of 0.1777 ± 0.0006 rad d^{-1}.

Inclination does not significantly affect $\mathrm{d}\Omega$. Indeed, decreasing the inclination from the adopted value of $i = 75°$ to 45° would result in a marginally larger differential rotation, three to four per cent. This is quite understandable. Inclination affects latitudes, but hardly periods.

A cursory glance cast at the beating pattern (Figure 7.7) reveals a lapping time $P_{\mathrm{beat}} \sim 40$ days, which is nearly exactly the lapping time of 40.3 days from the two peaks of the cleaned periodogram (Figure 7.1). From these 40.3 days one already gets an estimate of the minimum

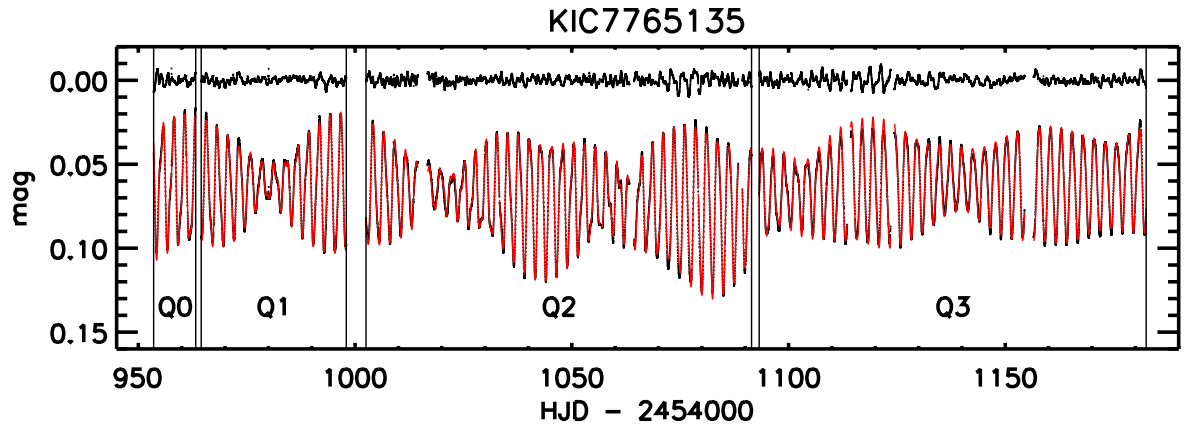

Figure 7.7: *Kepler light curve with best fit (solid red line, Case-A solution of Table 7.3) over-plotted. The residuals, shown at the top, are ±2.35 mmag. Obviously, the residuals are not homogeneous from one part of the light curve to another, hence, the value ±2.35 mmag has to be considered an overall average. Image courtesy of Antonio Frasca.*

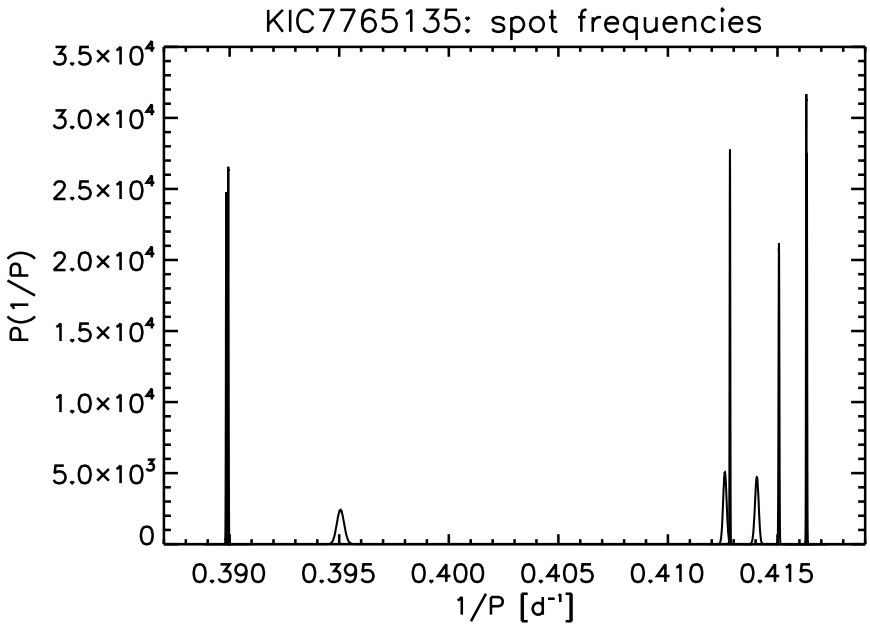

Figure 7.8: *Marginal PDFs of the frequency for all the nine spots. The frequency values group around the two principal frequencies (0.391 and 0.414 d^{-1}) seen already in the Fourier spectrum (Figure 7.1), which is the reason for the obvious "beating" phenomenon in Figure 7.7 with a period of 40 days. The shortest and the longest frequencies are a superposition of two frequencies. Image courtesy of Antonio Frasca.*

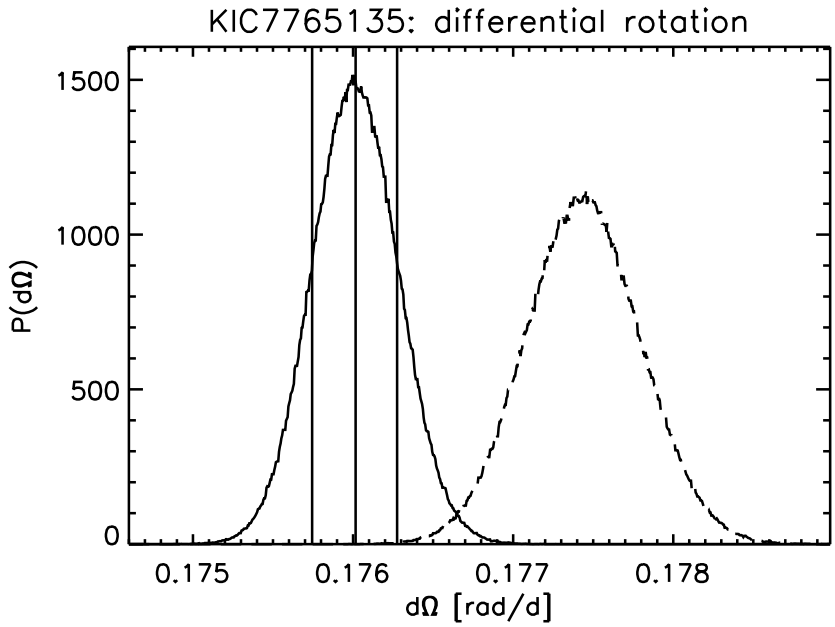

Figure 7.9: *Same as Figure 7.5, for KIC 7765135. Image courtesy of Antonio Frasca.*

value for the differential rotation as $2\pi/P_{\rm beat} \sim 0.156\,{\rm rad\,d^{-1}}$, which is not far from that derived by the model.

The spot area evolution (Case A) is depicted in Figure 7.10. The overwhelmingly large southern spot – at the beginning it fills to a large extent the southern hemisphere – may be an artifact. Because of its southern location its contribution to the light curve is rather modest. Perhaps it is actually a feature of the northern hemisphere, a non-circular extension of spot #2. To prevent spot overlapping, spot #8 had to be moved to the southern hemisphere. The reader should be aware that even in the case of a large spot the whole spot region has been assigned the angular velocity of its centre. Differential rotation is, to be exact, not compatible with a fixed circular shape. This is a shortcoming of our simple model.

In the case of KIC 7765135 it cannot be excluded that spot area evolution is partly driven by the need to avoid overlapping of spots. Expectation values with 1–σ confidence limits for various parameters are compiled in Table 7.3.

7.4 Discussion & Conclusions

Despite the fact that both stars are very active ones and exhibit filled-in absorption of several chromospheric activity indicators, our photometric analysis is in terms of dark surface features only. Allowing for bright ones too, would make the MCMC approach usually unstable. Anyway, photometry alone seems to be unable to discriminate even between dark and bright spots [210].

As a comparison of the two cases A and B reveals, the results of our spot modelling are hardly influenced by the rectification procedure. The smaller residuals for non-rectified data (Case B) are very likely due to the fact that the Case-B likelihood function, Eq. (7.3), also takes into account a linear trend in the data, individually for each part of the light curve. This allows for more freedom in fitting the data and results in a slightly better fit.

The reader should be aware that the estimated parameter values and their often surprisingly small errors are those of the model constrained by the data. Error bars indicate the "elbow

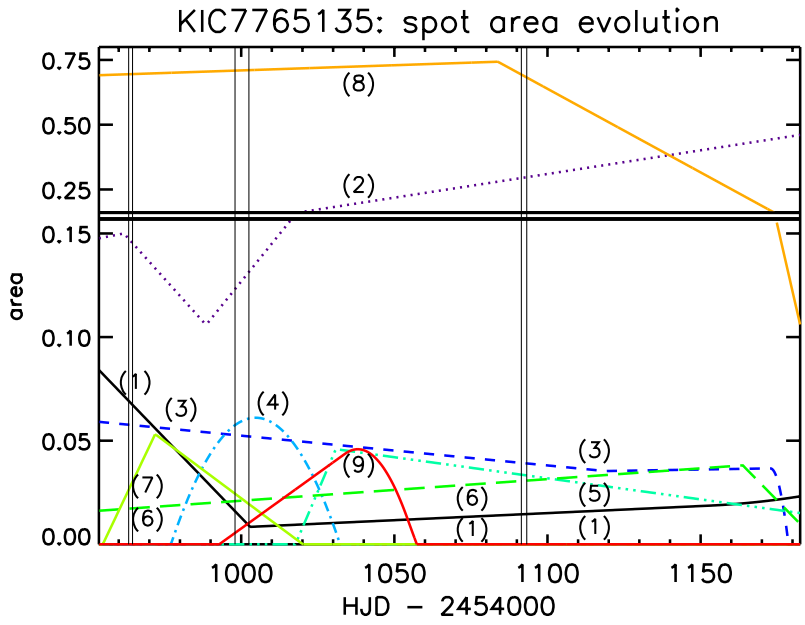

Figure 7.10: *Same as Figure 7.6, for KIC 7765135. Three of the nine spots (# 4, # 7, and # 9) are short-lived ones. Note the change in scale in the upper part! Image courtesy of Antonio Frasca.*

Table 7.3: *Two 9-spot solutions for KIC 7765135 with inclination being fixed to $i = 75°$. Listed are* expectation *values and 1-σ confidence limits. The meaning of the superscripts is the same as in Table 7.1.*

parameter			Case A		Case B	
inclination		i	$75\overset{\circ}{.}0$	*fixed*	$75\overset{\circ}{.}0$	*fixed*
1st	latitude	β_1	$20\overset{\circ}{.}4$	$^{+0.1}_{-0.1}$	$20\overset{\circ}{.}0$	$^{+0.1}_{-0.1}$
2nd	latitude	β_2	$76\overset{\circ}{.}3$	$^{+0.2}_{-0.2}$	$75\overset{\circ}{.}1$	$^{+0.2}_{-0.2}$
3rd	latitude	β_3	$1\overset{\circ}{.}9$	$^{+0.2}_{-0.1}$	$0\overset{\circ}{.}8$	$^{+0.2}_{-0.2}$
4th	latitude	β_4	$21\overset{\circ}{.}1$	$^{+0.3}_{-0.3}$	$20\overset{\circ}{.}5$	$^{+0.3}_{-0.3}$
5th	latitude	β_5	$0\overset{\circ}{.}9$	$^{+0.3}_{-0.4}$	$0\overset{\circ}{.}7$	$^{+0.3}_{-0.2}$
6th	latitude	β_6	$12\overset{\circ}{.}1$	$^{+0.1}_{-0.1}$	$11\overset{\circ}{.}7$	$^{+0.1}_{-0.1}$
7th	latitude	β_7	$16\overset{\circ}{.}3$	$^{+0.3}_{-0.3}$	$17\overset{\circ}{.}2$	$^{+0.3}_{-0.3}$
8th	latitude	β_8	$-75\overset{\circ}{.}9$	$^{+0.2}_{-0.2}$	$-74\overset{\circ}{.}4$	$^{+0.2}_{-0.2}$
9th	latitude	β_9	$60\overset{\circ}{.}2$	$^{+0.4}_{-0.4}$	$59\overset{\circ}{.}4$	$^{+0.4}_{-0.4}$
1st	period	P_1	2.4223	$^{+0.0001}_{-0.0001}$	2.4222	$^{+0.0001}_{-0.0001}$
2nd	period	P_2	2.5651	$^{+0.0001}_{-0.0001}$	2.5653	$^{+0.0001}_{-0.0001}$
3rd	period	P_3	2.4020	$^{+0.0001}_{-0.0001}$	2.4022	$^{+0.0001}_{-0.0001}$
4th	period	P_4	2.4237	$^{+0.0004}_{-0.0004}$	2.4231	$^{+0.0005}_{-0.0005}$
5th	period	P_5	2.4019	$^{+0.0001}_{-0.0001}$	2.4022	$^{+0.0001}_{-0.0001}$
6th	period	P_6	2.4092	$^{+0.0001}_{-0.0001}$	2.4092	$^{+0.0001}_{-0.0001}$
7th	period	P_7	2.4151	$^{+0.0004}_{-0.0005}$	2.4172	$^{+0.0005}_{-0.0005}$
8th	period	P_8	2.5645	$^{+0.0001}_{-0.0001}$	2.5641	$^{+0.0001}_{-0.0001}$
9th	period	P_9	2.5313	$^{+0.0010}_{-0.0010}$	2.5311	$^{+0.0009}_{-0.0009}$
spot intensity		κ	0.671	$^{+0.003}_{-0.003}$	0.700	$^{+0.003}_{-0.003}$
equ. period		$P_{\rm eq}$	2.4018	$^{+0.0001}_{-0.0001}$	2.4022	$^{+0.0001}_{-0.0001}$
deviation		C/B	-0.008		-0.008	
diff. rotation		$\mathrm{d}\Omega$	0.1760	$^{+0.0003}_{-0.0003}$	0.1774	$^{+0.0003}_{-0.0004}$
residuals			± 2.35		± 2.29	

room" of the model, nothing more.

7.4.1 Frequencies

It is remarkable that the frequencies that stand out in the power spectrum of the light curves (Figure 7.1) represent the distribution of spot frequencies (Figures 7.4 and 7.8) astonishingly well. The lapping time, as a measure of the lower limit of surface differential rotation, follows already from the periodogram analysis! However, in order to get an estimate of the full equator-to-pole span of the latitudinal shear, including its sign, one needs latitudinal information.

7.4.2 Inclination

Combining the inclination value from photometry, $i \approx 40°$, with the spectroscopically measured projected rotational velocity $v \sin i$ (see [118]) allows us to determine in the case of KIC 7985370 the star's radius. Taking the shortest rotational period ($P_{\rm eq}$), one arrives at $R = 1.42$–$1.44\,R_\odot$. This (minimal) radius is larger than the ZAMS value of $R \approx 1.1\,R_\odot$, but smaller than the $R \approx 2\,R_\odot$ for a star of $1.5\,M_\odot$ at 10 Myr. Hence, the photometrically derived radius is within the expected range. As stated in Section 7.3, in the case of KIC 7765135 the photometric inclination is badly defined. Therefore, the inclination has been fixed to $i = 75°$, assuming the radius to have its ZAMS value.

7.4.3 Spot contrast and spot longevity

Although both stars share the same spectral type and age, there are differences concerning the spots. The spots of KIC 7985370 seem to be darker and longer living than those of KIC 7765135.

We would like to remind that a "spot" may be in fact a group of smaller spots that all together form an active region, which could also include bright features.

For KIC 7765135, the spot contrast, $\kappa \approx 0.7$, looks rather normal. It is similar to the previously studied case of KIC 8429280 [115]. The corresponding temperature contrast, the ratio between spot and photospheric temperature $T_{\rm sp}/T_{\rm ph}$, is 0.9, assuming that the "white light" *Kepler* flux matches the bolometric conditions. The much darker spots, $\kappa \approx 0.4$, in the case of KIC 7985370 defy a simple explanation. There is no need for exceptionally small (and therefore dark) spots to prevent spot overlap.

Apart from a few late F-type stars with low or moderate activity observed by *CoRoT* [239] where spots seem to be short-lived, there is strong evidence that spots in very active stars like our targets have rather long lives compared to the star rotation. Active longitudes lasting for months or years have been observed in young stars (e.g. [84, 148, 14, 168, 201]) and in the evolved components of close binary systems, like II Peg (e.g. [267, 207]). This does not exclude that individual unresolved spots, which are composing the active region, have shorter evolution times, but the photospheric active region seen as an entity endures for a very long time in such cases.

Unlike KIC 7985370, in the case of KIC 7765135 mid-latitude spots (30–50°) are missing. This is reminiscent of the spot distribution of two fast-rotating early G dwarfs, He 520 and He 699, of the α Persei cluster studied by [14]. Despite the fact that there is a clear distinction between near-equator and near-pole spots, with regard to spot lifetimes, no correlation seems to exist between lifetime and latitude, which is contrary to the case of the rapidly-rotating young AB Dor [84], where only low- and intermediate-latitude spots are long-lived.

7.4.4 Differential rotation

Both stars exhibit low-latitude spots as well as high-latitude ones at the time of observation making them suitable for studying their latitudinal shear. The most robust and important result of the present work is the high degree of surface differential rotation found for both stars: $d\Omega = 0.18\,\mathrm{rad\,d^{-1}}$.This exceeds threefold the solar value.

This estimate is rather robust, because any spot model with a few long-lasting spots able to reproduce the beating of the light curve must provide a value of equator-to-pole differential rotation that exceeds the lower limit of $2\pi/P_{\mathrm{beat}}$, irrespective of the number of spots used. Inclination has a marginal effect since the periods found in the light curve do not depend on it.

We remark that the high value $d\Omega$ relies on the assumption of spot longevity. It is always possible to get an excellent fit with many short-lived spots even for rigid rotation.

Very different values of differential rotation have been found for HD 171488 (V889 Her), a young (∼50 Myr) Sun. For this star, which is rotating faster ($P = 1.33\,\mathrm{days}$) than our targets, a very high solar-type differential rotation $d\Omega \approx 0.4$–$0.5\,\mathrm{rad\,d^{-1}}$, with the equator lapping the poles every 12–16 days, was found by both [213] and [171]. Much weaker values ($d\Omega \approx 0.04\,\mathrm{rad\,d^{-1}}$) were derived instead for the same star by [169] and [195]. In the study by [168] the authors even claim their data being consistent with no differential rotation.

The analysis done by [214] reports values of $d\Omega$ in the range 0.08–$0.45\,\mathrm{rad\,d^{-1}}$ for a sample of stars similar to and slightly more massive than the Sun. Among these stars, HD 141943, a 1.3-$M_{\odot}$ star that is still in the PMS phase (age ∼ 17 Myr), displays values of $d\Omega$ ranging from about 0.23 to $0.44\,\mathrm{rad\,d^{-1}}$ in different epochs. A solar-type differential rotation, $d\Omega \approx 0.2\,\mathrm{rad\,d^{-1}}$, was also found by [313] for HD 106506, a G1 V-type star ($T_{\mathrm{eff}} = 5900\,\mathrm{K}$) that is very similar to our targets, but it is rotating faster ($P_{\mathrm{eq}} = 1.39\,\mathrm{days}$).

Moreover, the Fourier transform technique applied to high-resolution spectra of a large sample of F- and early G-type stars indicates that differential rotation is rather frequently found [264, 263]. In their data, there is no clear dependency on the rotation period, but the strongest differential rotation, up to $\sim 1.0\,\mathrm{rad\,d^{-1}}$, occurs for periods between 2 and 3 days and values as high as $\sim 0.7\,\mathrm{rad\,d^{-1}}$ are encountered down to $P \sim 0.5\,\mathrm{days}$.

From ground-based photometry, which is basically devoted to cooler stars, a different behavior, i.e. a differential rotation decreasing with the rotation period, seems to emerge [224]. However, the precision of the ground-based light curves does not allow to draw firm conclusions and accurate photometry from space, as well as Doppler imaging, is needed for settling this point.

For mid-G to M dwarfs, weaker values of the latitudinal shears are generally found. In particular, [15] analyzed with the Doppler imaging technique a small sample of active stars in the spectral range G2–M2 finding a trend towards decreasing surface differential rotation with decreasing temperature. This suggests that the stellar mass must also play a significant role in this respect. The largest values for stars as cool as about 5000 K are $d\Omega = 0.27\,\mathrm{rad\,d^{-1}}$ found by us in Paper I for KIC 8429280 (K2 V, $P = 1.16\,\mathrm{days}$) and $d\Omega = 0.20\,\mathrm{rad\,d^{-1}}$ found by [100] for LQ Hya (K2 V, $P = 1.60\,\mathrm{days}$).

The slowly rotating ($P_{\mathrm{eq}} = 11.2\,\mathrm{d}$) and mildly active K2 V star ϵ Eri exhibits only little differential surface rotation ($0.017 \leq d\Omega \leq 0.056\,\mathrm{rad\,d^{-1}}$) as a Bayesian reanalysis of the MOST light curve [93, 94] revealed [116]. Thus, there is an indication that a high differential rotation goes along with a high rotation rate.

The high differential rotation that we found for KIC 7985370 and KIC 7765135 disagrees with the hydrodynamical model of [196], which instead predicts a rather low value of $d\Omega \approx 0.08\,\mathrm{rad\,d^{-1}}$ for an (evolved) solar-mass star rotating with a period as short as 1.3 days.

Surface differential rotation may even vary along the activity cycle. Indeed, certain mean-

field dynamo models for rapidly rotating cool stars with deep convection zones predict torsional oscillations with variations of several percent in differential rotation [92]. Of course, this cannot explain such extreme cases as LQ Hya where at times the surface rotation is solid body. According to [200], to maintain the strong shear (~ 0.2 rad d^{-1}) observed for LQ Hya in the year 2000 would imply a dissipated power exceeding the star's luminosity.

The differential rotation of rapidly-rotating solar-like stars has been recently investigated on theoretical grounds by [157]. They found that differential rotation approaches the Taylor-Proudman state, i.e. the iso-rotation surfaces tend to become cylinders parallel to the rotation axis, when stellar rotation is faster than the solar one. In this case, the differential rotation is concentrated at relatively low latitudes with large stellar angular velocity. They show that the latitudinal shear (between the equator and latitude $\beta = 45°$) increases with the angular velocity, in line with our results and the recent literature.

A detailed study of the two Sun-like stars, KIC 7985370 and KIC 7765135, by means of high-resolution spectroscopy and high-precision *Kepler* photometry has been described by [118] and its description goes beyond the scope of this thesis. However, it is important to mention that the high-resolution spectra used by [118] allowed to derive, for the first time, their spectral type, astrophysical parameters ($T_{\rm eff}$, $\log g$, [Fe/H]), rotational and heliocentric radial velocities, and lithium abundance. All this information, combined with the analysis of the SED (spectral energy distribution) and proper motions, allowed to infer their distance and kinematics, and to estimate the age of both stars, found to be in the range 100–200 Myr, although we cannot exclude that they could be as young as 50 Myr. Thus, these two sources should be already in the post-T Tauri phase.

As expected from their young age, both stars were found to be chromospherically active displaying filled-in Hα, Hβ, and Ca II IRT lines, as well as He I D_3 absorption [118]. The surface chromospheric fluxes and the X-ray luminosity (for KIC 7985370), within the ranges found for stars with similar $T_{\rm eff}$ and $v \sin i$ in the Pleiades cluster, are just below the saturation level [280]. The flux ratio of two Ca II IRT lines and the Balmer decrement (for KIC 7765135 only) suggest that the chromospheric emission is mainly due to optically-thick surface regions analogous to solar plages.

In this chapter, we have described a robust spot model (performed by [118] for the study of the two young Suns presented here) based on a Bayesian approach and a MCMC method, to the *Kepler* light curves that span nearly 229 days and have an unprecedent precision ($\approx 10^{-5}$ mag). While seven long-lived spots were needed to perform a reasonable fit (at a 2-mmag level) of the light curve of KIC 7985370, up to nine spots were needed in the case of KIC 7765135 due to a shorter lifetime of its spots. Because of the exceptional precision of the *Kepler* photometry it is impossible to reach the Bayesian noise floor defined by, e.g., the BIC [273] (see also Chapter 5) without increasing significantly the degrees of freedom and, consequently, the non-uniqueness of the solution. Provided spots are indeed long-lived, the equator-to-pole value of the shear amounts for both stars to 0.18 rad d^{-1}. This is in contrast with the theoretical models of [196] that predict a moderate solar-type differential rotation even for fast-rotating main-sequence stars, unless the convection zone is shallower than predicted by the stellar models. The results presented are instead in line with the scenario proposed by other modelers of a differential rotation that increases with the angular velocity [157] and that can be also subject to changes along the activity cycle [92, 200].

Part IV

Appendices

A The AARG Code

In this Appendix we provide a detailed description of the AARG (Asymptotic Analysis of Red Giant stars) code, an IDL based code developed with the purpose of deriving the mean asteroseismic quantities (here denoted as asymptotic parameters) for both p modes and g modes — see Eqs. (1.10), and (1.12) — in large samples of red giant stars that show solar-like oscillations (see Chapter 4 for a brief introduction).

A.1 Getting started

At the beginning, AARG requires a list of stars to be analyzed, together with corresponding values of both $\nu_{\rm max}$ and $\Delta\nu$, and with datasets containing their PSDs (power density spectra), which for the analysis presented in Chapter 4 were derived by means of the SYD pipeline [164]. The input values of $\nu_{\rm max}$ and $\Delta\nu$ do not need to be neither very accurate nor precise, as AARG is supposed to derive the mean asymptotic parameters on its own. The entry display of the code visualizes three different plots for each star, namely the échelle PSD, the collapsed échelle diagram (CED from now on), and the PSD. This helps the user to check for the correctness of the input data (both $\nu_{\rm max}$, $\Delta\nu$ and the PSD) and to decide whether to continue the analysis or discard the star for checking it at a later time. Moreover, the stars are displayed in a decreasing order of $\Delta\nu$, which helps the overall analysis because the asymptotic parameters are expected to change smoothly as the stars evolve.

For the sake of clarity we shall consider the example of a single target, the RGB star KIC 2436818, having $\nu_{\rm max} = 96.1 \pm 2.8\,\mu$Hz and $\Delta\nu = 8.77 \pm 0.09\,\mu$Hz, and belonging to the open cluster NGC 6791 (see Appendix B), which represents a good reference star for explaining all the steps of the analysis done by AARG. Figure A.1 shows the échelle PSD (top panel) and the corresponding CED (bottom panel) as derived by means of the input values of $\nu_{\rm max}$ and $\Delta\nu$. In order to center the region of the power excess — the ordinate direction in Figure A.1(a) — AARG considers an interval in frequency centered at $\nu_{\rm max}$ and having width $\frac{4}{5}\nu_{\rm max}$. This choice resulted to provide a range of frequencies large enough to contain the entire power excess, but sufficiently small to let the regular structures of the oscillation ridges to appear prominent when collapsing the PSD into the CED. The échelle PSD is obtained in a similar manner to that derived by [63, 216].

A.1.1 The CED

The computation of the CED is done by a separate procedure within the AARG code, termed **PFOLD**, which does the following steps:

1. get as input the values for $\Delta\nu$ and the left and right margin values of the frequency range to be collapsed, i.e. $\nu_{\rm left} = \nu_{\rm max} - \frac{2}{5}\nu_{\rm max}$ and $\nu_{\rm right} = \nu_{\rm max} + \frac{2}{5}\nu_{\rm max}$

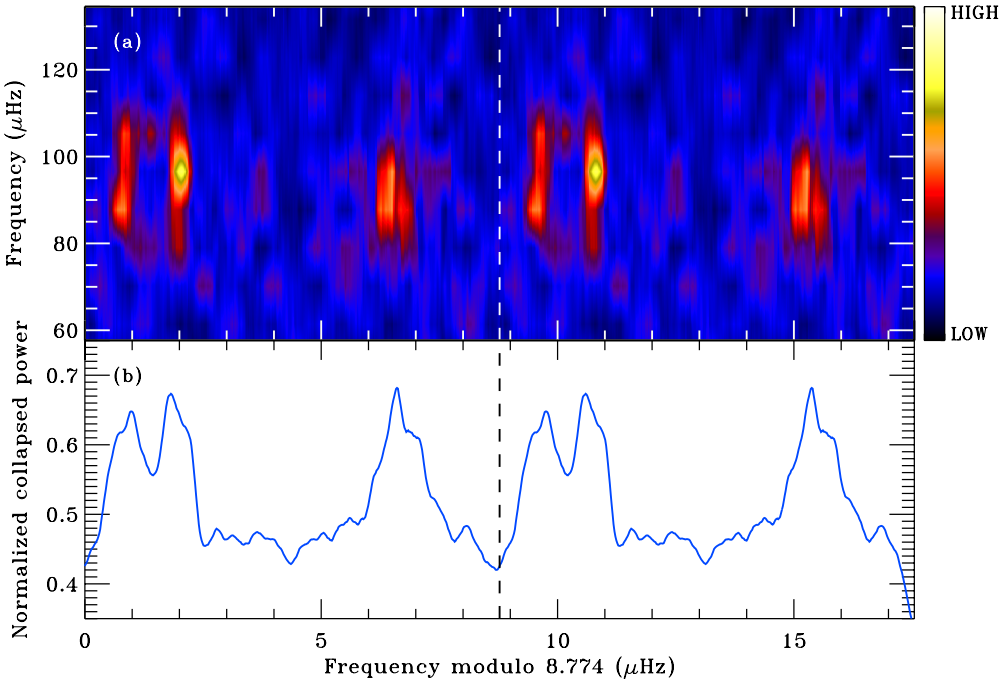

Figure A.1: *Panel (a): doubled échelle PSD of the RGB star KIC 2436818 used by* AARG. *The PSD is shown in a color-coded background. The dashed line overlaid shows the $\Delta\nu$ value of the star, which divides the échelle PSD into two identical parts. Panel (b): corresponding collapsed échelle diagram in normalized units, smoothed by means of a boxcar having width $4\Delta\nu$.*

2. find the correct starting point to collapse the PSD. This is done by computing the shift $c_{\rm ref} = \nu_{\rm left} \bmod \Delta\nu$ so that $\nu_{\rm initial} \bmod \Delta\nu = 0$, where $\nu_{\rm initial} = \nu_{\rm left} - c_{\rm ref}$ is the initial frequency to be considered as the starting point. This allows us to align correctly the structures in the CED (see points below).

3. find the largest integer number of slices having length $\Delta\nu$ contained in the range $[\nu_{\rm left}, \nu_{\rm right}]$. Thus we have $m = \lfloor(\nu_{\rm right} - \nu_{\rm left})/\Delta\nu\rfloor$, where $\lfloor\cdot\rfloor$ denotes the *floor* operator.

4. the i-th slice having length $\Delta\nu$ will be given as

$$\mathbf{s}_i = [\nu_{\rm initial} + i\Delta\nu, \nu_{\rm initial} + (i+1)\Delta\nu] \quad \text{for } i = 0, 1, \ldots, m-1 \tag{A.1}$$

5. the CED (or equivalently the collapsed échelle PSD or simply the collapsed PSD) in normalized units is then build as

$$\mathrm{CED} = \max_{[0,\Delta\nu]}{}^{-1} \left\{ \sum_{i=0}^{m-1} \mathrm{PSD}_i \right\} \times \sum_{i=0}^{m-1} \mathrm{PSD}_i \tag{A.2}$$

where $\mathrm{PSD}_i = \mathrm{PSD}(\mathbf{s}_i)$ is the power density spectrum at each slice $\mathbf{s}_i$. The resulting CED has to be plotted against a modulo abscissa, $\nu \bmod \Delta\nu$, in the range $[0, \Delta\nu]$ which is the interval within which all the structures of the CED can be found.

A.2 Fine-tuning the mean large separation $\Delta\nu$

The correlation between the mean large separation $\Delta\nu$[1] and the shift in position of the centroid of a given mode peak in a collapsed échelle diagram (or alternatively in a simple échelle diagram) is one of the known features of such an asteroseismic tool of analysis. In order to find the best $\Delta\nu$ value for a given star, a first guest for $\Delta\nu$ is our starting point[2]. The corresponding CED to be considered is built by adopting the first guess for $\Delta\nu$ and has a double width in abscissa in order to ensure us to observe the $\ell = 0$ peak in its entire shape[3].

In this section we intend to quantify the aforementioned correlation by assuming $\Delta\nu$ to be fine-tuned around its initial value — few % of $\Delta\nu$, see also [292]. We start by considering the asympotitc relation for the $\ell = 0$ modes [299]:

$$\nu_{n,0} = \Delta\nu(n + \epsilon). \tag{A.3}$$

If we take into account the frequency of maximum power, $\nu_{\max}$, we have

$$\nu_{\max} = \Delta\nu(n_{\max} + \epsilon_0)\,, \tag{A.4}$$

where

$$n_{\max} = \frac{\nu_{\max}}{\Delta\nu} - \epsilon_0 \tag{A.5}$$

is the radial order of the mode with frequency $\nu = \nu_{\max}$ and ϵ_0 is the constant that defines it's position in the échelle diagram. We note that ϵ_0 has to be considered as our first measure of the average ϵ value derived by the position of the $\ell = 0$ peak in the CED[4], obtained with our first guess of $\Delta\nu$. By perturbing Eq. (A.3), we obtain

$$\begin{aligned} \delta\nu &\simeq \delta(n\,\Delta\nu) + \delta(\Delta\nu\,\epsilon) \\ &= n\,\delta(\Delta\nu) + \delta(\Delta\nu\,\epsilon) \end{aligned} \tag{A.6}$$

where we put n out of the variation since it is a constant for a fixed mode of oscillation. If we now turn our attention to the $\ell = 0$ mode that corresponds to the maximum power in the PSD, by exploiting Eq. (A.5) we obtain

$$\begin{aligned} \delta(\nu_{\max}) &\simeq n_{\max}\,\delta(\Delta\nu) + \delta(\Delta\nu\,\epsilon) \\ &= \nu_{\max}\frac{\delta(\Delta\nu)}{\Delta\nu} - \epsilon_0\,\delta(\Delta\nu) + \delta S = 0\,, \end{aligned} \tag{A.7}$$

where $\delta(\nu_{\max}) = 0$ because $\nu_{\max}$ is an unique and observed value that characterizes each star and it does not change within the time spanned by an observation. We also defined $\delta S \equiv \delta(\epsilon\,\Delta\nu)$

[1] We refer to the term *mean large separation*, or *mean large frequency separation* because we are dealing with a global measurement of $\Delta\nu$, hence not a measure of the large separation as a function of frequency, which would arise from a detailed peak bagging of each star. Such an analysis goes beyond the scope of the ensemble analysis presented in Chapter 4.

[2] In general, $\Delta\nu$ can be easily estimated by means of either a Comb-Response function — e.g. see Chapter 2 or [48, 86] — or an Auto-Correlation function — e.g. see [164] — applied to the Power Spectrum of the star.

[3] For instance, if we consider a normalized modulo in abscissa, we have the range $(\nu/\Delta\nu \bmod 1) \in [0, 2]$, or alternatively if we consider a standard modulo (not normalized) we have $(\nu \bmod \Delta\nu) \in [0, 2\Delta\nu]$. In fact, it may happens that the $\ell = 0$ peak in the CED falls in a position close to the border of the diagram, which makes the fine-tuning of its position not possible. By adopting a doubled range we ensure that the peak is repeated without any cuts at least once.

[4] We postpone the discussion about the measurement of ϵ to Section A.3. For this section we assume that ϵ_0 is already given.

as the small shift of the centroid of the $\ell = 0$ peak in the CED caused by a change in $\Delta\nu$ of $\delta(\Delta\nu)$.

Hence, considering the infinitesimal quantities dS and $d(\Delta\nu)$, and integrating between an initial and a final $\Delta\nu$ value, ($\Delta\nu^{(i)}$ and $\Delta\nu^{(f)}$, respectively), the total shift ΔS occurring for a change in $\Delta\nu$, which we define as $\Delta(\Delta\nu) \equiv \Delta\nu^{(f)} - \Delta\nu^{(i)}$, is given by

$$\begin{aligned} \Delta S_{\rm if} &\simeq -\nu_{\rm max} \int_{\Delta\nu^{(i)}}^{\Delta\nu^{(f)}} \frac{d(\Delta\nu)}{\Delta\nu} + \epsilon_0 \int_{\Delta\nu^{(i)}}^{\Delta\nu^{(f)}} d(\Delta\nu) \\ &= -\nu_{\rm max} \ln \frac{\Delta\nu^{(f)}}{\Delta\nu^{(i)}} + \epsilon_0 \left(\Delta\nu^{(f)} - \Delta\nu^{(i)}\right) . \end{aligned} \tag{A.8}$$

The main properties of Eq. (A.8) are written below:

1. For any $\Delta\nu^{(f)}$ such that $\Delta\nu^{(f)} > \Delta\nu^{(i)}$ the shift is always negative, $\Delta S_{\rm if} < 0$, and vice versa if $\Delta\nu^{(f)} < \Delta\nu^{(i)}$, $\Delta S_{\rm if} > 0$.
2. $\Delta S_{\rm if} = 0$ for $\Delta\nu^{(f)} = \Delta\nu^{(i)}$.
3. When $\Delta(\Delta\nu) \sim 10^{-2}\Delta\nu$ (i.e. few % of variation in $\Delta\nu$), as it is the case when fine-tuning $\Delta\nu$ for Red Giants, Eq. (A.8) reduces to the linear equation[5]

$$\Delta S_{\rm if} \simeq -\nu_{\rm max} \left(\frac{\Delta\nu^{(f)}}{\Delta\nu^{(i)}} - 1 \right) + \epsilon_0 \left(\Delta\nu^{(f)} - \Delta\nu^{(i)}\right) \tag{A.9}$$

The property #1 follows from the fact that

$$\frac{\Delta\nu^{(f)}/\Delta\nu^{(i)} - 1}{\Delta\nu^{(f)} - \Delta\nu^{(i)}} \sim \frac{\epsilon_0 \Delta\nu}{\nu_{\rm max}} \sim 0.1 \tag{A.10}$$

being ϵ_0 a number of the order of unity. Thus, the first term appearing on the right-hand side of Eq. (A.8) is always dominant with respect to the second one.

AARG requires the position of the "clear" $\ell = 0$ peak, as it appears from the first doubled CED, to be manually marked by the user (we denote such position as $p_{0,i}$). At this point, it is worth to say that in an updated version of the AARG code, the position of the $\ell = 0$ peak could be estimated automatically by exploiting the ϵ–$\Delta\nu$ relation found by [87], which however holds for RGB stars only. In fact, in the case of clump RGs the observed ϵ is in average lower than that of their RGB counterparts showing similar $\Delta\nu$, with a difference quantified by [87] to be up to $\Delta\epsilon \simeq 0.15$ for the cluster RGs presented in Chapter 4 — see also the discussion in [87, 179, 238]. Hence, when analyzing heterogeneous samples of RGs, it may happen that the predicted value for $p_{0,i}$ provides a peak position that is shifted by some amount with respect to its true position. According to the findings by [87], on one hand this difference can be of the order of $0.15\Delta\nu$ — namely of the order of $\delta\nu_{02}$ — if a clump star is being analyzed, while on the other hand it is expected to be not greater than $0.05\,\mu$Hz in the case of RGB stars, as derived from the error propagation of their ϵ–$\Delta\nu$ relation — see Eq. (A.14) below. Such an amount of shifting becomes very important especially for more evolved stars, whose CED

[5] To derive the approximation of Eq. (A.8) for small $\Delta(\Delta\nu)$ we used the series expansion of the natural logarithm, namely

$$\ln(x+1) = \sum_{i=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} ,$$

where we considered $x = \Delta\nu^{(f)}/\Delta\nu^{(i)} - 1$, being $x \ll 1$.

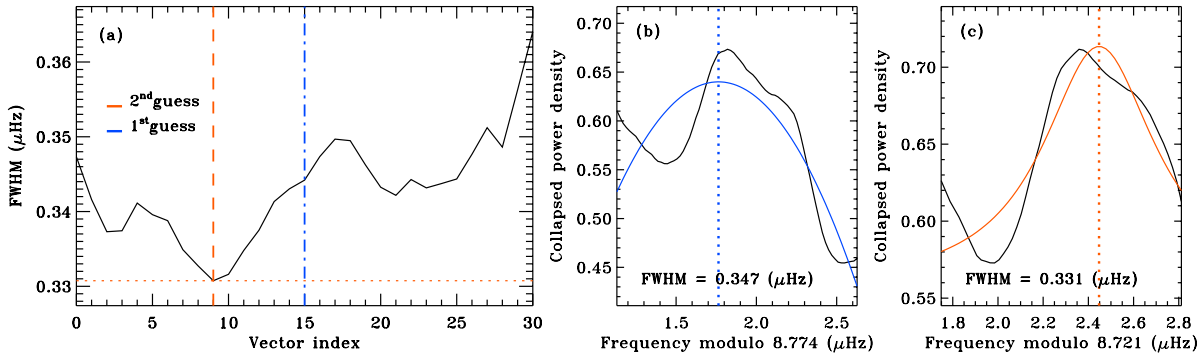

Figure A.2: *Results of the fine-tuning for the RGB star KIC 2436818. Panel (a) shows the FWHM profile as a function of the step in $\Delta\nu$. Index #15 represents our initial value, while the indices #0 and #30 correspond to +1.5% and −1.5% variation of $\Delta\nu$, respectively, which set the extremes of our fine-tuning interval. Panel (b) shows a zoom of the $\ell = 0$ peak in the CED with the initial guess of $\Delta\nu$, where a Lorentzian fit is overlaid for providing a rough estimate of its FWHM. Panel (c) is equivalent to panel (b) but in the case of the CED with the fine-tuned $\Delta\nu$. As visible from the result, the fine-tuned $\Delta\nu$ provides a narrower FWHM than that of the initial value.*

are more sensible to small variations in $\Delta\nu$. For partially getting rid of this issue, AARG performs a Lorentzian fit[6] to the $\ell = 0$ peak in the range $[p_{0,i} - d_{02}/2, p_{0,i} + d_{02}/2]$, where $d_{02} \simeq 0.123\Delta\nu$ (see Section A.3), which allows to recenter the predicted centroid by replacing it with the centroid arising from the fit[7].

Once that $p_{0,i}$ is found, the position of the $\ell = 0$ peak at each new shift in the fine-tuning phase will be given by $p_{0,f} = p_{0,i} + \Delta S_{\rm if}$. Therefore, Eq. (A.8) — or alternatively Eq. (A.9) — can be of great help especially for stars that show very low values of $\nu_{\rm max}$ ($\lesssim 4\,\mu$Hz) since the tracking of the $\ell = 0$ peak in the CED requires particular care due to the limit of the resolution and to the low number of observed modes. Eq. (A.8) has proven to provide a quite accurate and reliable guess for the position of the centroid of the $\ell = 0$ peak when fine-tuning $\Delta\nu$ within 2–3 % of the initial value provided by the SYD pipeline [164]. This method has been tested successfully in over a hundred of cluster RGs observed by *Kepler*.

A.2.1 The FWHM Method

At each step of the fine-tuning, i.e. for each step-value of $\Delta\nu$[8] a Lorentzian fit to the new shifted $\ell = 0$ peak is performed. Such fit has a double purpose: first, the fit centroid can be used for improving the position of the peak centroid in the follow-up; second, and more important, it allows to obtain a rough estimate of the FWHM of the peak. In fact, as discussed by [292], following from the properties of the asymptotic relation for p modes, the correct value of $\Delta\nu$ provides the best vertical alignment of the oscillation ridges in the échelle power spectrum of the star. According to this explanation, the final estimate of $\Delta\nu$ arising from the fine-tuning is reached by taking the value that minimizes the FWHM of the $\ell = 0$ peak. Figure A.2 shows a clarifying example of the method for the star KIC 2436818.

Anyhow, regardless of the sophistication of the method described above, the resulting fine-tuned $\Delta\nu$ of the RGB stars of the clusters NGC 6791 and NGC 6819, resulted to have a spread

[6] All the Lorentzian fits performed by AARG are done by means of the IDL procedure MPFITPEAK distributed within the MPFIT IDL Library.

[7] Such a particular care is only required for this fine-tuning phase, because for secondary steps of the analysis the correct position of the $\ell = 0$ peak must be known already, as we shall discuss in Section A.3 below.

[8] The step considered for the fine-tuning process amounts to $10^{-3}\Delta\nu$.

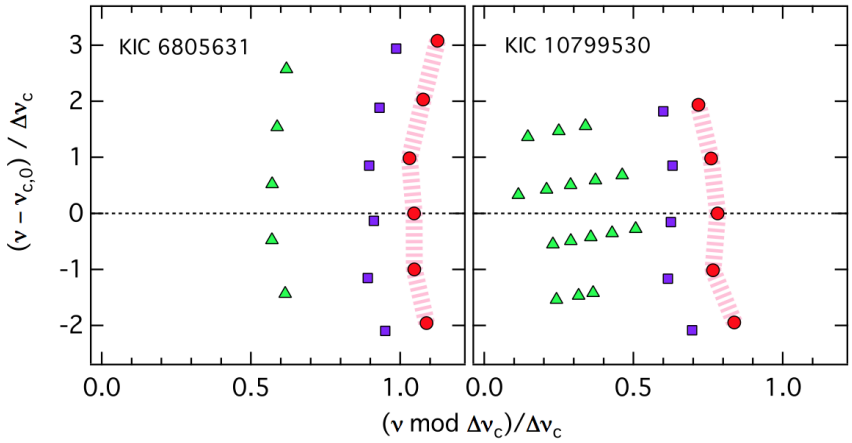

Figure A.3: *Example of échelle diagrams of an RGB (left) and RC (right) star with $\ell = 0, 1$, and 2 modes indicated by red circles, green triangles, and blue squares, respectively. The ordinate is the frequency in units of $\Delta\nu_c$ and centered at the value corresponding to the turning point, here denoted as $\nu_{c,0}$. Image credit by [179].*

larger than that of the values given as input from the SYD pipeline, up to a factor of $\sim 2 - -3$. The reason for this result is very likely to rely on the enhanced curvature of the ridges of oscillation, which is due to the large number of modes detected thanks to the long observing run. In fact, the curvature of the ridges will result in a broader peak width when collapsing the PSD into a CED. A solution to this problem is presented in the forthcoming section.

A.2.2 Overcoming the problem of Ridge Curvature

One of the main effects that hampers the fine-tuning process of the mean large separation by means of the FWHM method is related to the curvature of the ridge corresponding to the radial modes (e.g. see [178, 243, 179] for more details). In fact, as the observations become longer and more modes are so detected, the frequencies of the radial modes start to deviate from their asymptotic values, although in an amount of the order of $\sim 10^{-2}\Delta\nu$, see [243]. In particular, the radial mode sequence of an RGB star becomes "C" shaped, while that of an RC star is more "S" shaped [179] (see Figure A.3). In order to overcome this issue, we follow a similar idea to that introduced by [178] and consider a mean large separation $\Delta\nu_c$ defined for the modes close to the frequency of maximum power $\nu_{\max}$. This choice relies on the fact that, as it appears from all the observations, the ridge curvature has a turning point that is always close to the radial mode corresponding to $\nu_{\max}$. This means that the radial modes close to $\nu_{\max}$ are those less affected by the ridge curvature, hence these modes are those of interest to find the best value for $\Delta\nu_c$ and subsequently $\Delta\nu$.

Nevertheless, for the analysis presented in Chapter 4, we decided not to consider an improved version of the fine-tuning that takes into account the central radial modes because reliable estimates of $\Delta\nu$ as derived by means of the SYD pipeline [164] were already available for all the RGs investigated. For future works, we aim at implementing this method in the AARG code in order to allow the code to perform a completely independent and self-supported analysis of the asymptotic parameters, with the advantage of being automatized so that large numbers of red giant stars can be analyzed in a short time lapse.

A.3 Measuring ϵ and the Small Spacings $\delta\nu_{02}$ and $\delta\nu_{01}$

As already mentioned in Chapter 4, in order to derive the measurements for ϵ and for the small spacings $\delta\nu_{02}$ and $\delta\nu_{01}$, we first collapse the échelle diagram of each star by adopting its corresponding best value of $\Delta\nu$, as derived with, e.g., the method described in Section A.2. The CED to be considered has again a double width, this time with the purpose of allowing us to catch each structure of the diagram at least once without any cuts (each peak corresponding to an oscillation ridge can be repeated in its entire shape twice at most, see also the discussion in Section A.2).

A.3.1 Shifting and clipping the doubled CED

To perform the analysis in a simpler manner, AARG requires the *temporary position* of the $\ell = 0$ peak[9] — referred to as p_0 in the following — which is the position corresponding to the fine-tuned $\Delta\nu$, or alternatively to the position of $\ell = 0$ peak in the doubled CED if no fine-tuning has been used. Thus, we aim at displaying the complete sequence of the $\ell = 0, 2, 1$ peaks without any cuts and respecting the order suggested by the asymptotic relation given by Eq (**??**), i.e. $\ell = 2, 0, 1$ from left to right. This is done by AARG by shifting the range in abscissa of the CED by an amount given by the quantity

$$s_{\rm ref} \equiv p_2 - \frac{3}{4} d_{02} \,, \tag{A.11}$$

where d_{02} is a rough guess of the small spacing $\delta\nu_{02}$ derived by means of the empirical relation $d_{02} \equiv p_2 - p_0 = c\Delta\nu$, with $c \simeq 0.123$ — see also [30] who adopted a similar value — and p_2 a rough guess of the position of the $\ell = 2$ peak in the doubled CED. Consequently AARG considers a slice of length $\Delta\nu$ for applying the clipping to the doubled CED. Thus, in the new CED the positions of the $\ell = 0$ and 2 centroids will be given by $p'_0 = p_0 - s_{\rm ref}$ and by $p'_2 = p_2 - s_{\rm ref}$ (or equivalently by $p'_2 = p'_0 - d_{02}$). These positions will be used in the following Section A.3.4 to derive the small spacings $\delta\nu_{02}$ and $\delta\nu_{01}$.

Lastly, we note that when the fine-tuning process is performed, the shifting and clipping of the CED is done automatically. In this way, the new interval in abscissa of the CED used for the analysis will be always of the form $[s_{\rm ref}, \Delta\nu + s_{\rm ref}]$.

A.3.2 The ϵ Term

In Section A.2 we have seen that AARG uses a very first estimate of ϵ, which we called ϵ_0, in order to exploit the relation given by Eq. (A.8) for fine-tuning the $\Delta\nu$ value. In the following we shall describe in pratice how can ϵ_0 be derived from the doubled CED considered for either the fine-tuning or the shifting and clipping process shown in Section A.3.1.

While measuring ϵ one should keep in mind that, according to the results derived by [165, 317], we expect to find a reliable estimate in the range $[0.5, 1.5]$ for stars having $\Delta\nu < 15\,\mu$Hz. If the $\ell = 0$ peak is repeated twice within the range of the CED, e.g. see Figure A.4, our ϵ_0 can be measured either for the first or the second $\ell = 0$ peak appearing in the doubled CED, no matter which one we choose to consider. In fact, AARG recognizes the true position of the

[9] We term *temporary position* of a peak the position of the peak in the doubled CED, the latter not being necessarily equal to the true position of the centroid. This is because the marked $\ell = 0$ peak could lay in the second half of the CED, i.e. in the range $(\nu \bmod \Delta\nu) \in [\Delta\nu, 2\Delta\nu]$, hence the real and temporary position would differ by a factor $\Delta\nu$. For further details, see also the discussion in Section A.3.2.

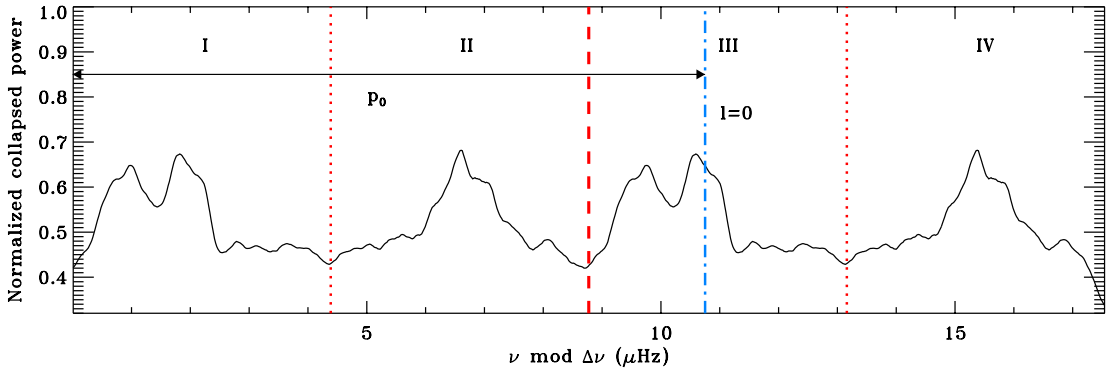

Figure A.4: *Doubled CED of the RGB star KIC 2436818 used by* AARG *for fine-tuning* $\Delta\nu$*. The ordinate shows an arbitrary scale for the collapsed power, while the abscissa is expressed in* μ*Hz. The dashed red line shows the position of the* $\Delta\nu$ *value, which separates the doubled CED into two equal parts containing the same, repeated pattern of the* $\ell = 2, 0, 1$ *peaks (single CEDs). The blue dot-dashed blue line marks our position of the* $\ell = 0$ *peak considered for the analysis, while the dotted red lines mark the values of* $\Delta\nu/2$ *and* $\frac{3}{2}\Delta\nu$*, hence dividing the entire doubled CED into four regions, labeled from I to IV ingoing from left to right. The blue arrow highlights the value of the marked position of the* $\ell = 0$ *peak, which we denoted as* p_0*. The doubled CED has been smoothed by means of a boxcar having width* $4\Delta\nu$ *to simplify the pattern of the peaks.*

$\ell = 0$ peak by simply considering the cases

$$\text{if} \begin{cases} p_0 > \Delta\nu \Rightarrow & p_0^{\text{true}} = p_0 - \Delta\nu \\ p_0 < \Delta\nu \Rightarrow & p_0^{\text{true}} = p_0 \end{cases}, \tag{A.12}$$

where p_0 is once again the temporary position of the $\ell = 0$ peak as derived from the beginning. The estimate of ϵ_0 will depend upon which of the four regions, labeled in Figure A.4 from I to IV, our marked $\ell = 0$ peak falls. Thus, according to the asymptotic relation for p modes — Eq. (1.10) — AARG considers the solutions

$$\begin{cases} \epsilon_0 = (p_0/\Delta\nu) + 1 & \text{(I or III)} \\ \epsilon_0 = p_0/\Delta\nu & \text{(II or IV)} \end{cases}. \tag{A.13}$$

A new version of the code could predict the peak centroid by exploiting the well-calibrated ϵ-$\Delta\nu$ relation derived by [87] for the RGB stars of the open clusters NGC 6791 and NGC 6819 (see Chapter 4, Section 4.3.1), namely the logarithmic relation

$$\epsilon = (0.601 \pm 0.025) + (0.632 \pm 0.032) \log \Delta\nu\,, \tag{A.14}$$

or alternatively the power law relation

$$\epsilon = (0.681 \pm 0.017)\Delta\nu^{0.261\pm0.014}\,. \tag{A.15}$$

In this manner, the position of the $\ell = 0$ peak could be found automatically for any RGB star by simply giving as input the mean large separation $\Delta\nu$. This can be certainly of great advantage when analyzing thousands of Red Giants as those observed by *Kepler*. However, as discussed for the fine-tuning process, such relation does not hold for clump stars, and a manual mark of the peak may still result in being the best solution for a proper analysis of heterogeneous samples of RGs.

Finally, the last estimate of ϵ, which we take as the first asymptotic parameter derived for the analysis presented in Chapter 4, is obtained by fitting a Lorentzian profile to the $\ell = 0$ peak marked by the position p'_0 (or alternatively p_0 if no shift has been applied), within the range $[p'_0 - d_{02}/2, p'_0 + d_{02}/2]$. We also define the parameter $m_0 \equiv d_{02}/2$ as the one giving the margin width for the Lorentzian fit to the $\ell = 0$ peak (see Section A.4). The fit yields us to the final centroid ν_0, and consequently to the final ϵ of the star, as derived by means of Eq. (A.13), in which we use ν_0 instead of p_0 (see also Figure 4.3 in Chapter 4 and the related discussion).

A.3.3 Average Linewidths of $\ell = 0$ Modes

By considering the CED smoothed by means of a boxcar having width $\Delta\nu$, we can use the FWHM of the Lorentzian profile fitted to the $\ell = 0$ peak in the interval $[\nu_0 - m_0, \nu_0 + m_0]$ in order to derive some information about the lifetime of these modes. In fact, the FWHM derived can be used as an upper limit to the real average linewidth of the radial modes [87], and we will term it $\Gamma_0^{\rm up}$ from now on. The resulting values of $\Gamma_0^{\rm up}$ are listed in Appendix B for all the stars analyzed in the sample. These values were adopted for studying the correlation between linewidths and both $\Delta\nu$ and $T_{\rm eff}$, as described in Chapter 4. Nevertheless, we note that our values do not represent accurate measurements for linewidths, because a detailed peak bagging for each star and an accurate estimation of the background level in the power spectrum would be required for obtaining estimates of the real linewdiths of the modes.

For the sake of completeness, we highlight that the empirical relation given by Eq. (4.11), found by [87] for predicting the average linewidths of radial modes Γ in stars from MS to RGs, is also based on the set of measurements $\Gamma_0^{\rm up}$. In particular, the uncertainties on Γ can be derived from the standard error propagation, which yields

$$\sigma_\Gamma\left(T_{\rm eff}\right) = \Gamma\left(T_{\rm eff}\right)\sqrt{\left(\frac{\sigma_{\Gamma_0}}{\Gamma_0}\right)^2 + \left(\frac{T_{\rm eff} - 5777\,{\rm K}}{T_0^2}\right)^2 \sigma_{T_0}^2 + \left(\frac{\sigma_{T_{\rm eff}}}{T_0}\right)^2}, \tag{A.16}$$

where $\Gamma_0 = 1.39\,\mu$Hz, $\sigma_{\Gamma_0} = 0.10\,\mu$Hz, $T_0 = 601\,$K, $\sigma_{T_0} = 3\,$K and $T_{\rm eff}$ is the effective temperature of the star expressed in K.

A.3.4 The Small Spacings $\delta\nu_{02}$ and $\delta\nu_{01}$

A further step in the analysis of p modes performed by AARG consists in measuring the small spacing $\delta\nu_{02}$, namely the spacing between the $\ell = 0$ and $\ell = 2$ peaks in the CED introduced in Section A.3.1, as defined by Eq. (4.1). This is done by AARG by considering the centroid ν_0 derived in the previous step, and subsequently fitting a Lorentzian profile to the $\ell = 2$ peak, in the range $[p'_2 - d_{02}/2, p'_2 + d_{02}/2]$. Again, we introduce the width parameter m_2, which for $\ell = 2$ is given by $m_2 \equiv d_{02}/2 = m_0$. The Lorentzian fit leads us to the final centroid ν_2 of the peak, thus we have that $\delta\nu_{02} = \nu_0 - \nu_2$. An example of this procedure is shown in Figure A.5(a) for the star KIC 2436818.

Conversely to $\delta\nu_{02}$, the estimation of the small spacing $\delta\nu_{01}$ is more subtle because the $\ell = 1$ peak has often a complex structure caused by the presence of mixed modes, which deviate from the regular pattern defined by the asymptotic relation for p modes. In order to improve the capability of AARG to detect and properly fit the $\ell = 1$ ridge, the CED is generally smoothed by means of a boxcar having width $4\Delta\nu$, which makes the pattern of the oscillation ridges clearer and simpler, as shown in Figure A.5(b).

AARG estimates the position of the $\ell = 1$ peak by finding the maximum $c_{\rm max}$ of the CED in the region $[\nu_0 + \delta\nu_{02}, \Delta\nu]$, which is clearly expected to contain the entire central pattern of the dipole modes. Subsequently, it is required to the user to manually mark the left-hand extreme

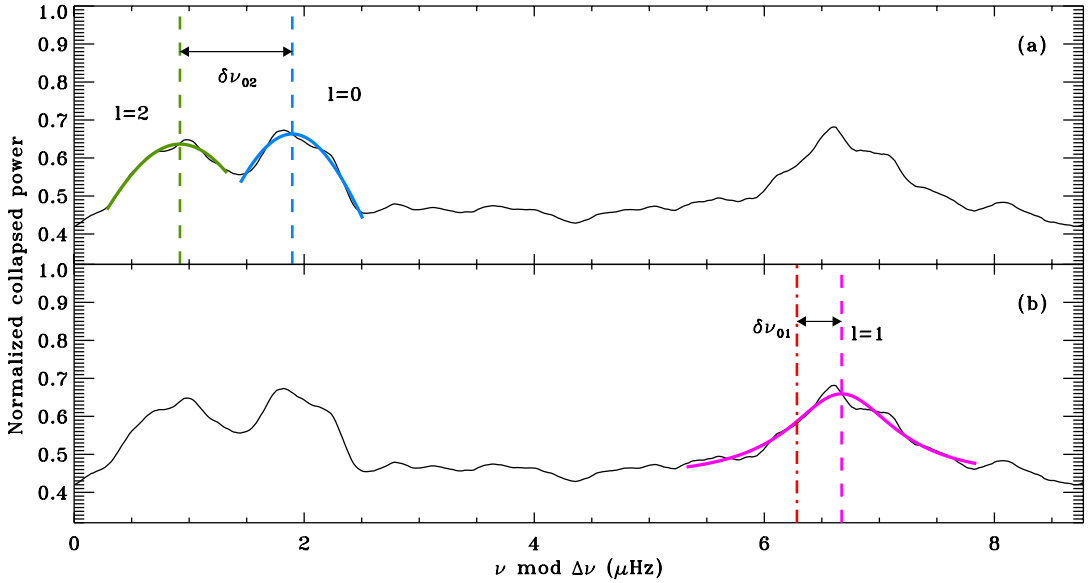

Figure A.5: *CED of the RGB star KIC 2436818 used by* AARG *for the derivation of the asymptotic parameters, smoothed by means of a boxcar having width* $4\Delta\nu$ *to simplify the pattern of the peaks. Panel (a): the dashed lines show the position of the centroids* ν_2 *and* ν_0 *arising from the Lorentzian fits, marked with thick solid lines. Labeled and indicated by an arrow, the final measurement of the small spacing* $\delta\nu_{02}$*. Panel (b): Same case as panel (a) but for the Lorentzian fit to the* $\ell = 1$ *peak (thick purple line), whose centroid* ν_1 *is marked with a dashed purple line. The red dot-dashed line represents the position* $\nu_0 + \Delta\nu/2$*, considered for measuring the small spacing* $\delta\nu_{01}$ *according to the definition given by Eq. (4.2).*

of the interval for the Lorentzian fit, which has to be large enough to contain all the possible prominent structures related to $\ell = 1$. The range between $c_{\max}$ and the left-hand extreme marked by the user is saved by AARG as the width parameter m_1, which will be useful for the derivation of the uncertainties, as we shall discuss in Section A.4. The outcoming centroid ν_1 of the Lorentzian fit applied to the interval $[c_{\max} - m_1, c_{\max} + m_1]$ is used to derive our final estimate of the small spacing $\delta\nu_{01}$. In fact, according to the definition given by Eq. (4.2), we have that $\delta\nu_{01} = \nu_0 - \nu_1 + \Delta\nu/2$, namely $\delta\nu_{01}$ assumes negative or positive values whether the $\ell = 1$ peak falls above or below the semi-distance $\nu_0 + \Delta\nu/2$ occurring between two consecutives $\ell = 0$ peaks (see the example given in Figure A.5).

A.4 Derivation of the Uncertainties

The analysis of the asymptotic parameters of p modes is finally completed with the derivation of the uncertainties for ϵ, $\delta\nu_{02}$, $\delta\nu_{01}$ (see also Chapter 4 for a brief introduction). AARG follows a similar approach to that used by [167], which consists in generating 500 different realizations of each star's PSD by perturbing the observed one by means of a synthetic noise level that is χ^2-distributed with 2 d.o.f., according to the statistics of a PSD [319].

In order to "rescale" the noise level at each realization, AARG multiplies a vector containing χ^2-distributed random numbers by the vector containing the values of the observed PSD. This allows each bin in the PSD to be rescaled, either increased or decreased, by an amount given by the corresponding synthetic noise at that bin. In addition, each synthetic noise vector has been normalized to the χ^2 distribution of the observed PSD, that is it has been rescaled by some factor in order to make the two χ^2 distributions coincide. Such factor depends on the

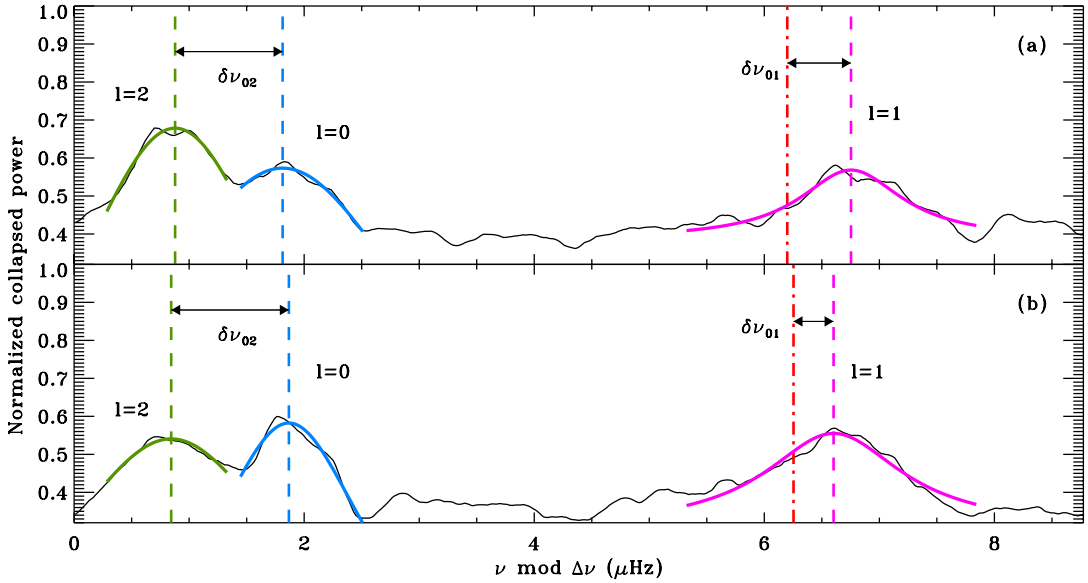

Figure A.6: *Example of two synthetic CEDs of the RGB star KIC 2436818 used by* AARG *for the derivation of uncertainties. The notation is the same as for Figure A.5. As it appears clear by comparing top and bottom panels, different realizations result in different centroids* ν_0, ν_1, ν_2*, hence in a new set of asymptotic values* $\delta\nu'_{02}, \delta\nu'_{01}$*. Both CEDs have been smoothed by means of a boxcar having width* $4\Delta\nu$*.*

star considered and was found to be between 1.5 and 2.5 for most of the stars analyzed.

After a synthetic PSD is derived, it is collapsed into a synthetic CED by using the correct $\Delta\nu$ for the star, in a similar manner to that described in Section A.1.1. The CED is afterwards smoothed by convolving it for a boxcar having width $4\Delta\nu$. AARG exploits the parameters ν_0, ν_1, ν_2 that mark the position of the $\ell = 0, 1, 2$ peaks in the original CED, together with the corresponding margins used for the Lorentzian fits, which we denoted as m_0, m_1, m_2. In fact, AARG applies three Lorentzian fits to the synthetic CED by using the parameters mentioned beforehand. This allows us to derive a new set of centroids ν'_0, ν'_1, ν'_2, used to obtain the synthetic asymptotic parameters $\delta\nu'_{02}, \delta\nu'_{01}$ with the methods described in the sections above. This step is fully automated and it is repeated 500 times for each star. A plot of the CED with the actual Lorentzian fits overlaid (see Figure A.6) is displayed for each simulation with the aim at checking that the fits are performing well during the analysis. For improving the reliability of the fit for the $\ell = 1$ peak, an additional smoothing up to $8\Delta\nu$ to the region $[\nu_1 - m_1, \nu_1 + m_1]$ of the CED can be applied. The additional smoothing can be extremely useful for stars having particularly confusing $\ell = 1$ regions. Figure A.6 shows two different synthetic CED for the star KIC 2436818. When comparing the position of the centroids in the two CEDs, it is apparent that the asymptotic parameters $\delta\nu'_{02}, \delta\nu'_{01}$ derived in each simulation are different.

The last step concerns the computation of the uncertainties. This can be done straightforwardly by evaluating the standard deviation of each sample of synthetic parameters

$$\left\{\delta\nu'_{02,i}\right\}, \quad \left\{\delta\nu'_{01,i}\right\} \qquad \text{with } i = 1, 2, \ldots, 500\,. \tag{A.17}$$

However, the distributions of synthetic parameters may contain some outliers, which in general can arise from a wrong fit especially for stars where $\ell = 0$ and $\ell = 2$ are very close to each other and $\ell = 1$ is splitted into several peaks. In order to get rid of these "bad" points responsible of increasing the effective standard deviation of the distribution, AARG computes a resistant estimate of its dispersion by means of the Tukey's method [152], implemented in the

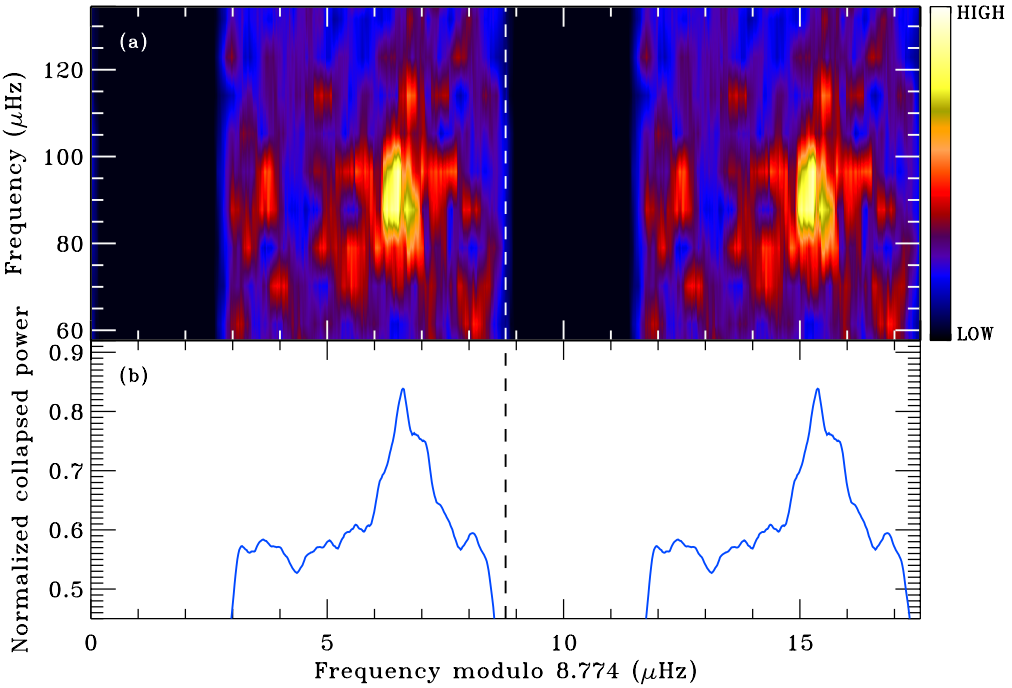

Figure A.7: *Similar description as the one adopted for Figure A.1 but in the case of an "erased" doubled échelle PSD of the RGB star KIC 2436818 — panel (a). The PSD is shown in a color-coded background, where black regions represents those set to zero value. Panel (b) shows the corresponding erased CED in normalized units, smoothed by means of a boxcar having width $4\Delta\nu$. As it appears clear when comparing to Figure A.1, the regions containing radial and quadrupole modes are not appearing in this erased CED.*

IDL procedure **ROBUST_SIGMA**. The robust sigma has proved to give narrower constraints than the Gaussian standard deviation for most of the stars analyzed. This last step is also performed in a completely automated manner. In addition, AARG plots histograms of the distributions of synthetic parameters for each star. Since the distributions are expected to be Gaussian-like, the histograms are an helpful tool to the user for checking wrong results.

Lastly, the uncertainty on ϵ is derived by the error propagation from $\Delta\nu$, according to Eq. (1.10). Thus we have that

$$\sigma_\epsilon = \frac{\nu_{\max}}{\Delta\nu^2}\sigma_{\Delta\nu}\,, \tag{A.18}$$

where $\sigma_{\Delta\nu}$ is the 1σ uncertainty on $\Delta\nu$. All the uncertainties so derived are saved and used for the plots shown in Chapter 4.

A.4.1 Uncertainty on $\Delta\nu$

In the case $\Delta\nu$ is derived by means of a fine-tuning process that uses an improved FWHM method as the one discussed in Section A.2.2, the uncertainty $\sigma_{\Delta\nu}$ can be estimated as the FWHM of the Lorentzian fit to the $\ell = 0$ peak in the CED, which must be collapsed for the fine-tuned value of $\Delta\nu$. Otherwise, the uncertainty on $\Delta\nu$ can be taken as the FWHM of the peak arising from either a CR function (see, e.g., the analysis described in Chapter 2) or an ACF method [164], the latter being the case for the input values adopted for the results presented in Chapter 4.

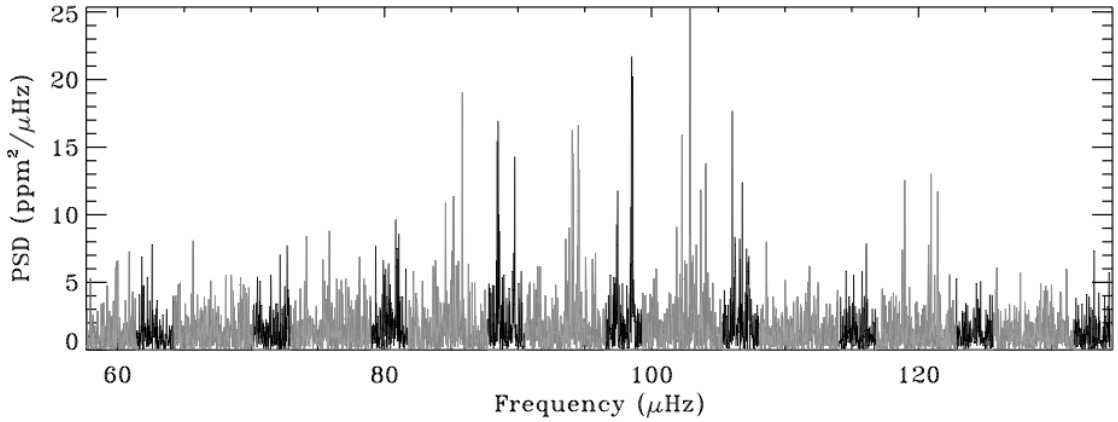

Figure A.8: *"Erased" PSD (gray) of the RGB star KIC 2436818 overlaid on the original PSD (black). Gray PSD has only dipole mixed modes regions.*

A.5 Estimation of $\Delta P_{\rm obs}$ for Mixed Dipole Modes

As already discussed in Chapter 4, the observation of mixed dipole modes — that is oscillation modes that show a mixed behavior between that of pure g and dipole p modes, see [21] for the first discovery with *Kepler* — can be of great importance for our understanding about the evolutionary state of a red giant star. In fact, as noted by [32], and later on by [242], the measurement of the so-called observed period spacing $\Delta P_{\rm obs}$ is an extremely useful tool for distinguishing between H-shell (RGB) and He-core burning (RC) red giant stars. This analysis has been done also for cluster RGs by [87], where $\Delta P_{\rm obs}$ could be measured in almost half of the sample considered, despite of a lower signal-to-noise ratio in the oscillation signal of these stars (see Appendix B for a list of all the results).

As already mentioned in Chapter 4, AARG exploits the method introduced by [32] for measuring the mean $\Delta P_{\rm obs}$ in RGs, which allows us to readily measure the parameter from the PSDs. The analysis of mixed dipole modes done by AARG involves the following steps:

1. by using the centroids ν_0, ν_2 and the corresponding margins m_0, m_2 for the Lorentzian fits derived the analysis of p modes, erase (that is set to zero value) all the regions in the star's CED belonging to the interval $[\nu_2 - 1.3m_2, \nu_0 + 1.5m_0]$, chosen empirically to remove any excess of power related to radial and quadrupole modes. This is done by AARG through a separate procedure, named **ERASEPOWER**, which generates a CED having $\ell = 0, 2$ peaks completely removed. An example of this result is shown in Figure A.7(b) for the star KIC 2436818, together with the corresponding échelle PSD in a color-coded background — panel (a)

2. reconstruct the new "erased" PSD (hereafter EPSD) with an inverse procedure to that used to derive the CED, which simply merges consecutive slices of the erased échelle PSD in increasing order of frequency. This is again done by the procedure **ERASEPOWER**. The result for KIC 2436818 is shown in Figure A.8 in gray color, overlaid on the original PSD in black

3. convert the abscissa of the EPSD into periods (to be expressed in seconds) and apply a Fourier transform to the result. This is done by a separate function named **PPS**, which uses the **LNP_TEST** IDL library function to compute the Lomb Periodogram of the EPSD converted into periods (see also [259], Section 13.8)

4. the power spectrum of the EPSD expressed in periods, is then converted into a power spectrum in periods as well by simply reverting the frequencies into periods and plotting in the range 30–500 s. The result, whenever any mixed modes would be present, will show a power excess peaked around the value $\Delta P_{\rm obs}$, which we aim at measuring for inferring the evolutionary state of the RGs

5. mark a rough position of such power excess and the left (or right) margin of the interval to be used for fitting the power excess with a Gaussian profile. The resulting centroid of the Gaussian fit gives our estimate of $\Delta P_{\rm obs}$.

As auxiliary tools, AARG displays a zoom of the EPSD converted into periods for each slice containing a region of mixed dipole modes. This is done to help the user to search for any regular pattern, as suggested by the asymptotic relation for g modes given by Eq. (1.12). In addition, it is also possible to compute the Lomb Periodogram of each slice in order to visualize a possible power excess arising from the presence of a regular spacing in period. The reason for which checking a single slice separately is suggested derives from the fact that mixed dipole modes may be apparent in a slice rather than in another.

B Asymptotic Parameters

In this Appendix we list the asymptotic values derived by means of the AARG code (see Appendix A) for the Red Giants belonging to the clusters NGC 6791, NGC 6811, and NGC 6819. For each star analyzed, the p-mode asymptotic values $\nu_{\rm max}$, $\Delta\nu$, ϵ, $\delta\nu_{02}$, $\delta\nu_{01}$, together with their 1σ uncertainties, the upper limits to the linewidths of radial modes $\Gamma_0^{\rm up}$, and the period spacing of mixed dipole modes $\Delta P_{\rm obs}$, are reported. Stars for which $\Gamma_0^{\rm up}$ could not be derived properly and stars where no evidence of period spacing has been found, show a dash symbol for the corresponding parameter. We also note that the values of $\nu_{\rm max}$ and $\Delta\nu$, as well as their uncertainties, are those derived by means of the SYD pipeline [164]. We remind that the dataset used to derive these quantities is the one described in Chapter 4.

Table B.1: *Asymptotic parameters for the Cluster NGC 6791.*

KIC ID	$\nu_{\rm max}$ (μHz)	$\Delta\nu$ (μHz)	ϵ	$\delta\nu_{02}$ (μHz)	$\delta\nu_{01}$ (μHz)	$\Gamma_0^{\rm up}$ (μHz)	$\Delta P_{\rm obs}$ (s)
2297384	31.8 ± 1.8	3.80 ± 0.06	0.810 ± 0.128	0.416 ± 0.087	-0.271 ± 0.063	0.136	256.
2297825	29.6 ± 1.6	3.83 ± 0.09	0.806 ± 0.172	0.513 ± 0.070	-0.299 ± 0.086	0.145	254.
2435987	38.1 ± 1.2	4.22 ± 0.01	0.954 ± 0.032	0.553 ± 0.025	-0.175 ± 0.028	0.176	–
2436097	42.4 ± 1.3	4.53 ± 0.03	1.000 ± 0.053	0.568 ± 0.038	-0.166 ± 0.030	0.199	–
2436209	57.8 ± 0.7	5.73 ± 0.03	1.073 ± 0.051	0.732 ± 0.037	-0.184 ± 0.037	0.183	–
2436332	28.1 ± 1.4	3.40 ± 0.02	0.916 ± 0.039	0.455 ± 0.031	-0.126 ± 0.021	0.076	–
2436417	26.7 ± 0.8	3.41 ± 0.06	0.874 ± 0.134	0.342 ± 0.090	-0.237 ± 0.074	0.155	268.
2436458	36.0 ± 1.3	4.15 ± 0.02	0.980 ± 0.038	0.561 ± 0.029	-0.156 ± 0.030	0.147	48.
2436540	57.4 ± 1.3	5.79 ± 0.03	1.088 ± 0.043	0.724 ± 0.033	-0.199 ± 0.039	0.214	–
2436593	111.6 ± 1.7	9.64 ± 0.09	1.238 ± 0.107	1.178 ± 0.094	0.078 ± 0.155	0.248	55.
2436676	132.7 ± 5.8	11.31 ± 0.06	1.263 ± 0.067	1.521 ± 0.228	-0.037 ± 0.204	0.297	55.
2436688	76.9 ± 1.4	7.22 ± 0.03	1.174 ± 0.043	0.872 ± 0.071	-0.240 ± 0.049	0.233	–
2436732	30.9 ± 1.2	3.72 ± 0.06	0.781 ± 0.144	0.483 ± 0.094	-0.249 ± 0.079	0.313	211.
2436759	30.1 ± 3.0	3.75 ± 0.18	0.834 ± 0.391	0.500 ± 0.122	-0.178 ± 0.187	0.231	–
2436814	24.9 ± 0.9	3.13 ± 0.02	0.903 ± 0.053	0.417 ± 0.033	-0.141 ± 0.021	0.086	–
2436818	96.1 ± 2.8	8.77 ± 0.09	1.225 ± 0.113	1.171 ± 0.150	-0.220 ± 0.170	0.222	54.
2436824	33.2 ± 0.7	3.87 ± 0.02	0.934 ± 0.033	0.519 ± 0.025	-0.129 ± 0.022	0.126	–
2436900	35.4 ± 1.5	4.05 ± 0.02	0.968 ± 0.034	0.499 ± 0.035	-0.186 ± 0.031	0.121	–
2436912	30.4 ± 1.4	3.78 ± 0.05	0.820 ± 0.112	0.480 ± 0.090	-0.319 ± 0.085	0.236	–
2437040	26.2 ± 0.6	3.07 ± 0.02	0.901 ± 0.051	0.394 ± 0.034	-0.170 ± 0.029	0.104	–
2437103	29.7 ± 1.7	3.79 ± 0.06	0.770 ± 0.133	0.325 ± 0.153	-0.242 ± 0.114	0.130	306.
2437240	44.3 ± 1.0	4.84 ± 0.02	1.035 ± 0.036	0.614 ± 0.032	-0.196 ± 0.034	0.179	–
2437270	69.8 ± 0.9	6.51 ± 0.02	1.107 ± 0.041	0.780 ± 0.055	-0.232 ± 0.042	0.216	43.
2437325	93.2 ± 1.2	8.50 ± 0.03	1.186 ± 0.042	1.080 ± 0.059	0.054 ± 0.153	0.217	61.
2437340	8.5 ± 0.2	1.36 ± 0.03	0.698 ± 0.126	0.191 ± 0.015	-0.078 ± 0.015	0.019	–
2437353	31.0 ± 1.1	3.85 ± 0.04	0.818 ± 0.080	0.555 ± 0.053	-0.272 ± 0.059	0.174	248.
2437394	164.3 ± 4.2	12.86 ± 0.81	1.378 ± 0.802	1.707 ± 0.354	-0.349 ± 0.288	0.339	–
2437402	45.6 ± 1.2	4.82 ± 0.02	1.019 ± 0.048	0.623 ± 0.029	-0.151 ± 0.028	0.185	–
2437444	18.9 ± 1.2	2.46 ± 0.02	0.877 ± 0.067	0.350 ± 0.027	-0.100 ± 0.031	0.097	–
2437488	64.1 ± 1.2	6.27 ± 0.03	1.096 ± 0.042	0.821 ± 0.039	-0.189 ± 0.041	0.229	–
2437496	4.3 ± 0.5	0.85 ± 0.19	0.701 ± 1.124	0.147 ± 0.049	-0.042 ± 0.025	–	–
2437507	20.4 ± 0.5	2.60 ± 0.02	0.888 ± 0.055	0.351 ± 0.020	-0.128 ± 0.018	0.079	–
2437564	30.6 ± 1.3	3.83 ± 0.06	0.899 ± 0.127	0.479 ± 0.111	-0.239 ± 0.087	0.130	241.
2437589	46.5 ± 1.5	4.60 ± 0.03	1.026 ± 0.057	0.526 ± 0.042	-0.184 ± 0.038	0.186	39.
2437653	74.0 ± 1.4	7.00 ± 0.03	1.195 ± 0.044	0.896 ± 0.050	-0.188 ± 0.037	0.211	–

Table B.1

KIC ID	$\nu_{\rm max}$ (μHz)	$\Delta\nu$ (μHz)	ϵ	$\delta\nu_{02}$ (μHz)	$\delta\nu_{01}$ (μHz)	$\Gamma_0^{\rm up}$ (μHz)	$\Delta P_{\rm obs}$ (s)
2437698	30.1 ± 1.5	3.75 ± 0.03	0.804 ± 0.068	0.507 ± 0.141	-0.250 ± 0.049	0.466	–
2437781	85.1 ± 1.6	7.83 ± 0.03	1.182 ± 0.045	1.029 ± 0.063	-0.046 ± 0.072	0.172	–
2437804	26.5 ± 1.6	3.35 ± 0.07	0.870 ± 0.165	0.478 ± 0.054	-0.266 ± 0.529	0.278	212.
2437805	30.0 ± 1.5	3.80 ± 0.05	0.837 ± 0.094	0.578 ± 0.045	-0.176 ± 0.054	0.174	–
2437816	17.9 ± 0.4	2.33 ± 0.02	0.929 ± 0.051	0.309 ± 0.037	-0.087 ± 0.016	0.083	–
2437933	107.7 ± 1.5	9.38 ± 0.05	1.252 ± 0.066	1.071 ± 0.106	-0.281 ± 0.274	0.230	53.
2437957	90.5 ± 1.8	8.52 ± 0.03	1.178 ± 0.033	1.036 ± 0.073	-0.156 ± 0.114	0.239	48.
2437965	8.3 ± 0.7	1.31 ± 0.04	0.525 ± 0.182	0.199 ± 0.028	-0.120 ± 0.039	–	–
2437972	84.6 ± 1.3	7.83 ± 0.05	1.196 ± 0.064	0.976 ± 0.074	-0.089 ± 0.093	0.204	50.
2437976	89.1 ± 1.6	8.14 ± 0.05	1.221 ± 0.067	1.015 ± 0.106	-0.108 ± 0.073	0.242	51.
2437987	29.3 ± 1.3	3.76 ± 0.08	0.830 ± 0.159	0.444 ± 0.124	-0.239 ± 0.084	0.585	224.
2438038	62.1 ± 1.3	6.15 ± 0.02	1.085 ± 0.036	0.795 ± 0.041	-0.206 ± 0.034	0.187	46.
2438051	30.7 ± 2.1	3.70 ± 0.10	0.832 ± 0.229	0.339 ± 0.106	-0.409 ± 0.102	0.116	207.
2438140	69.3 ± 1.1	6.74 ± 0.02	1.129 ± 0.031	0.794 ± 0.038	-0.234 ± 0.053	0.169	–
2438333	61.2 ± 0.9	6.07 ± 0.02	1.107 ± 0.038	0.737 ± 0.039	-0.259 ± 0.035	0.251	46.
2569055	31.4 ± 2.4	3.76 ± 0.08	0.791 ± 0.175	0.517 ± 0.094	-0.229 ± 0.088	0.208	–
2569360	22.0 ± 0.7	2.73 ± 0.03	0.933 ± 0.092	0.360 ± 0.051	-0.143 ± 0.031	0.089	–
2569618	55.0 ± 1.2	5.65 ± 0.02	1.119 ± 0.037	0.773 ± 0.040	-0.179 ± 0.035	0.180	–
2569935	5.5 ± 0.8	1.01 ± 0.12	0.613 ± 0.629	0.140 ± 0.036	-0.099 ± 0.042	–	–
2569945	29.1 ± 1.0	3.75 ± 0.07	0.984 ± 0.140	0.536 ± 0.158	-0.262 ± 0.087	–	253.
2570094	67.2 ± 0.9	6.50 ± 0.02	1.146 ± 0.034	0.817 ± 0.040	-0.148 ± 0.036	0.174	56.
2570172	73.6 ± 2.3	7.03 ± 0.03	1.173 ± 0.046	0.863 ± 0.086	-0.235 ± 0.051	0.246	–
2570244	103.0 ± 2.8	9.29 ± 0.23	1.142 ± 0.273	1.124 ± 0.117	-0.143 ± 0.157	0.241	56.
2570384	64.3 ± 1.4	6.36 ± 0.03	1.161 ± 0.048	0.841 ± 0.088	-0.232 ± 0.063	0.183	44.
2570518	45.9 ± 1.2	4.94 ± 0.02	1.053 ± 0.031	0.632 ± 0.034	-0.143 ± 0.029	0.146	–

Table B.2: *Asymptotic parameters for the Cluster NGC 6811.*

KIC ID	$\nu_{\rm max}$ (μHz)	$\Delta\nu$ (μHz)	ϵ	$\delta\nu_{02}$ (μHz)	$\delta\nu_{01}$ (μHz)	$\Gamma_0^{\rm up}$ (μHz)	$\Delta P_{\rm obs}$ (s)
9534041	107.4 ± 2.2	8.40 ± 0.04	1.022 ± 0.067	0.837 ± 0.061	-0.247 ± 0.063	0.409	142.
9655101	98.2 ± 2.4	7.86 ± 0.04	1.060 ± 0.061	0.889 ± 0.053	-0.034 ± 0.085	1.072	98.
9655167	100.3 ± 8.7	8.07 ± 0.04	1.034 ± 0.065	0.803 ± 0.064	-0.511 ± 0.115	0.342	–
9716090	101.4 ± 5.9	8.56 ± 0.06	1.001 ± 0.078	0.852 ± 0.167	0.059 ± 0.123	–	172.
9716522	54.9 ± 1.0	4.85 ± 0.04	0.973 ± 0.083	0.592 ± 0.102	-0.116 ± 0.099	0.435	154.

Table B.3: *Asymptotic parameters for the Cluster NGC 6819.*

KIC ID	$\nu_{\rm max}$ (μHz)	$\Delta\nu$ (μHz)	ϵ	$\delta\nu_{02}$ (μHz)	$\delta\nu_{01}$ (μHz)	$\Gamma_0^{\rm up}$ (μHz)	$\Delta P_{\rm obs}$ (s)
4937056	43.4 ± 1.0	4.76 ± 0.06	0.974 ± 0.118	0.581 ± 0.172	-0.213 ± 0.091	0.236	211.
4937576	31.1 ± 1.6	3.56 ± 0.02	0.913 ± 0.047	0.415 ± 0.032	-0.109 ± 0.027	0.129	–
4937770	93.8 ± 2.4	7.82 ± 0.08	1.119 ± 0.116	0.808 ± 0.117	-0.096 ± 0.071	0.345	109.
5023732	26.9 ± 1.1	3.11 ± 0.02	0.892 ± 0.061	0.353 ± 0.045	-0.122 ± 0.031	0.145	–
5023845	108.6 ± 1.1	8.90 ± 0.02	1.223 ± 0.031	1.050 ± 0.048	-0.143 ± 0.099	0.242	–
5023931	50.1 ± 2.4	4.93 ± 0.04	0.968 ± 0.082	0.585 ± 0.058	-0.194 ± 0.048	0.322	–
5023953	50.0 ± 1.1	4.76 ± 0.02	0.936 ± 0.043	0.700 ± 0.057	-0.009 ± 0.361	0.208	214.
5024043	55.9 ± 0.5	5.59 ± 0.02	1.108 ± 0.039	0.700 ± 0.032	-0.132 ± 0.029	0.187	–
5024143	111.3 ± 8.9	9.66 ± 0.03	1.201 ± 0.038	1.139 ± 0.082	-0.049 ± 0.227	0.255	–
5024240	153.7 ± 2.9	11.97 ± 0.05	1.262 ± 0.050	1.485 ± 0.150	-0.152 ± 0.293	0.364	–
5024297	46.1 ± 0.5	4.57 ± 0.02	1.047 ± 0.051	0.517 ± 0.037	-0.091 ± 0.030	0.142	–
5024312	93.6 ± 3.3	8.12 ± 0.02	1.136 ± 0.031	0.892 ± 0.049	-0.116 ± 0.103	0.274	–
5024327	43.9 ± 1.8	4.73 ± 0.08	0.977 ± 0.155	0.687 ± 0.053	-0.049 ± 0.100	0.372	211.
5024404	48.9 ± 0.7	4.86 ± 0.13	0.835 ± 0.261	0.689 ± 0.095	-0.122 ± 0.072	0.144	181.
5024405	96.4 ± 1.6	8.28 ± 0.02	1.141 ± 0.030	0.933 ± 0.044	-0.161 ± 0.076	0.399	–
5024414	77.1 ± 1.5	6.49 ± 0.06	1.013 ± 0.102	0.720 ± 0.072	-0.220 ± 0.143	0.297	178.
5024476	67.0 ± 1.7	5.69 ± 0.10	1.138 ± 0.201	0.656 ± 0.152	-0.203 ± 0.234	0.999	199.
5024512	73.9 ± 2.6	6.68 ± 0.03	1.096 ± 0.044	0.784 ± 0.039	-0.072 ± 0.064	0.318	–
5024517	50.0 ± 1.8	4.91 ± 0.03	1.043 ± 0.059	0.569 ± 0.055	-0.044 ± 0.061	0.163	–
5024582	46.3 ± 2.0	4.78 ± 0.14	0.945 ± 0.287	0.692 ± 0.129	-0.173 ± 0.114	0.192	248.
5024583	38.1 ± 0.8	3.91 ± 0.02	0.931 ± 0.047	0.515 ± 0.025	-0.080 ± 0.035	0.132	–
5024601	31.8 ± 1.7	3.70 ± 0.03	0.862 ± 0.065	0.498 ± 0.061	-0.140 ± 0.107	0.162	–
5024750	11.7 ± 0.6	1.76 ± 0.04	0.946 ± 0.162	0.198 ± 0.047	-0.109 ± 0.034	0.131	–
5024851	4.6 ± 0.2	0.75 ± 0.03	0.580 ± 0.241	0.087 ± 0.016	-0.026 ± 0.016	0.006	–

Table B.3

KIC ID	$\nu_{\max}$ (μHz)	$\Delta\nu$ (μHz)	ϵ	$\delta\nu_{02}$ (μHz)	$\delta\nu_{01}$ (μHz)	$\Gamma_0^{\rm up}$ (μHz)	$\Delta P_{\rm obs}$ (s)
5024967	46.9 ± 1.7	4.72 ± 0.06	0.964 ± 0.131	0.595 ± 0.073	-0.181 ± 0.136	0.414	194.
5111718	135.2 ± 1.4	10.52 ± 0.04	1.264 ± 0.050	1.144 ± 0.066	-0.112 ± 0.078	0.320	–
5111940	51.9 ± 1.1	5.18 ± 0.02	1.041 ± 0.039	0.597 ± 0.033	-0.044 ± 0.039	0.178	–
5111949	48.7 ± 1.1	4.96 ± 0.06	0.672 ± 0.125	0.746 ± 0.068	-0.001 ± 0.111	0.639	235.
5112072	125.0 ± 1.3	10.01 ± 0.03	1.262 ± 0.040	1.154 ± 0.059	-0.102 ± 0.082	0.288	–
5112288	47.8 ± 1.8	4.81 ± 0.03	0.892 ± 0.062	0.618 ± 0.051	-0.146 ± 0.082	0.191	–
5112361	67.4 ± 1.4	6.18 ± 0.03	1.066 ± 0.044	0.712 ± 0.066	-0.102 ± 0.048	0.259	75.
5112373	43.7 ± 1.1	4.67 ± 0.06	0.849 ± 0.123	0.623 ± 0.047	-0.118 ± 0.113	0.205	187.
5112387	45.7 ± 1.3	4.75 ± 0.05	0.868 ± 0.105	0.585 ± 0.059	-0.287 ± 0.102	0.489	208.
5112401	38.2 ± 0.7	4.05 ± 0.06	0.892 ± 0.148	0.476 ± 0.082	-0.169 ± 0.069	0.142	209.
5112403	138.8 ± 2.0	11.09 ± 0.05	1.313 ± 0.061	1.264 ± 0.153	-0.314 ± 0.259	0.331	–
5112467	46.3 ± 1.1	4.71 ± 0.05	1.061 ± 0.108	0.760 ± 0.072	-0.222 ± 0.088	0.241	220.
5112481	4.9 ± 0.6	0.88 ± 0.06	0.894 ± 0.352	0.133 ± 0.025	-0.020 ± 0.017	0.247	–
5112491	45.0 ± 1.9	4.68 ± 0.04	0.984 ± 0.079	0.732 ± 0.042	-0.031 ± 0.134	0.197	240.
5112730	45.7 ± 0.9	4.55 ± 0.04	0.997 ± 0.079	0.631 ± 0.052	-0.028 ± 0.160	0.317	232.
5112734	40.7 ± 0.9	4.13 ± 0.03	1.008 ± 0.076	0.443 ± 0.065	-0.141 ± 0.046	0.152	–
5112744	45.4 ± 1.0	4.42 ± 0.02	1.022 ± 0.050	0.569 ± 0.046	-0.061 ± 0.037	0.278	–
5112786	8.3 ± 0.8	1.14 ± 0.04	0.781 ± 0.274	0.146 ± 0.025	-0.072 ± 0.035	0.030	–
5112880	26.3 ± 0.7	2.81 ± 0.02	0.859 ± 0.059	0.320 ± 0.043	-0.057 ± 0.027	0.072	–
5112938	44.1 ± 1.2	4.75 ± 0.07	0.940 ± 0.145	0.706 ± 0.175	-0.089 ± 0.089	0.338	257.
5112948	43.3 ± 0.8	4.28 ± 0.02	1.024 ± 0.046	0.546 ± 0.029	-0.116 ± 0.033	0.209	–
5112950	42.8 ± 1.3	4.30 ± 0.04	1.082 ± 0.083	0.584 ± 0.104	0.010 ± 0.181	0.166	249.
5112974	41.3 ± 0.7	4.36 ± 0.04	0.874 ± 0.097	0.655 ± 0.115	0.064 ± 0.073	0.369	239.
5113041	37.7 ± 0.9	3.97 ± 0.02	1.020 ± 0.052	0.485 ± 0.035	-0.160 ± 0.035	0.130	–
5113441	154.8 ± 2.2	11.71 ± 0.03	1.271 ± 0.032	1.360 ± 0.068	-0.087 ± 0.128	0.323	56.
5200152	46.3 ± 0.9	4.79 ± 0.06	0.866 ± 0.113	0.685 ± 0.050	0.032 ± 0.133	0.372	236.

C Numerical Methods in Statistical Analysis

In this appendix we discuss some interesting numerical methods used in frequentist and Bayesian statistical analysis, the latter in particular referred to Chapter 6. In the first part of this appendix we highlight the differences between the frequentist definition of confidence intervals and the counterpart Bayesian credible intervals, the latter representing one of the fundamental information arising from a Bayesian parameter estimation. In Section C.1, we show how confidence intervals are directly related to the χ^2 distribution and how they can be derived in the case of a maximum likelihood estimator (MLE). In Section C we focus on the meaning of credibility intervals and on the numerical method adopted to derive them for the analysis presented in Chapter 6. Lastly, in Section C.3 we present an useful numerical method that has been adopted for computing the Bayesian evidence in the low dimensional problems ($k \leq 4$) discussed in Chapter 6.

C.1 Confidence intervals

In the present section we shall describe some details of parameter estimation in the frequentist approach. Referring to the definitions given in Section 5.3, let us consider a sampling distribution of a statistic S, with mean μ_s and standard deviation σ_s. Assuming S to be an unbiased estimator of the population mean, then its mean value is equal to the population mean. Thus, if the sampling distribution of S is approximately a normal distribution, then we can expect to find S lying in the intervals $\mu_s - \sigma_s$ to $\mu_s + \sigma_s$, $\mu_s - 2\sigma_s$ to $\mu_s + 2\sigma_s$, or $\mu_s - 3\sigma_s$ to $\mu_s + 3\sigma_s$ about 68.27 %, 95.45 %, and 99.73 % of the time, respectively. Equivalently, we can expect to find, or be *confident* of finding, μ_s in the intervals $S - \sigma_s$ to $S + \sigma_s$, $S - 2\sigma_s$ to $S + 2\sigma_s$, or $S - 3\sigma_s$ to $S + 3\sigma_s$ about 68.27 %, 95.45 %, and 99.73 % of the time, respectively. For this reason, the intervals considered are called the 68.27 %, 95.45 %, and 99.73 % *confidence intervals* (CI) for estimating μ_s, respectively [281]. The end points of these intervals ($S \pm \sigma_s$, $S \pm 2\sigma_s$, $S \pm 3\sigma_s$) are known as *confidence limits* or *fiducial limits*, while the percentage confidence is often named *confidence level* (CL).

Therefore, given some quantity to be normal distributed, it is possible to find a correlation between the percentage of the area under the distribution (e.g. 100 % $\rightarrow p = 1$ if one considers the entire area) and the product $z_c\sigma$ between the standard deviation σ of the normal distribution and the so called *confidence coefficient* z_c. Some of the most used values of the confidence coefficient, expressed in units of standard deviation, and the corresponding confidence levels are listed in Table C.1. We should note that the definition of confidence intervals, limits, levels and coefficients only holds in the frequentist approach to statistics, although their introduction is mandatory for properly defining and understanding the case of Bayesian inference. The

Table C.1: *Most used confidence levels and their corresponding confidence coefficients* z_c.

CL	50 %	68.27 %	90 %	94.45 %	99 %	99.73 %
z_c (σ)	0.6745	1.00	1.645	2.00	2.58	3.00

Bayesian point of view will be discussed in Section C.2.

C.1.1 Confidence Limits for a χ^2 Distribution

In cases where the sampling distribution is not a normal distribution, appropriate modifications to confidence intervals introduced above are easily made. In particular, we shall discuss the case of a *chi-square distribution*, which for $x \geq 0$ is given by

$$P(\chi^2 \leq x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^x u^{(\nu/2)-1} e^{-u/2} du \tag{C.1}$$

and $P(\chi^2 \leq x) = 0$ for $x < 0$, where ν is the *number of degrees of freedom* (d.o.f) and Γ is the *gamma function*[1]. The corresponding density function reads

$$f_\nu(x) = \begin{cases} \frac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x > 0 \\ 0 & x \leq 0 \end{cases} \tag{C.2}$$

and is shown in Figure C.1 for some degrees of freedom.

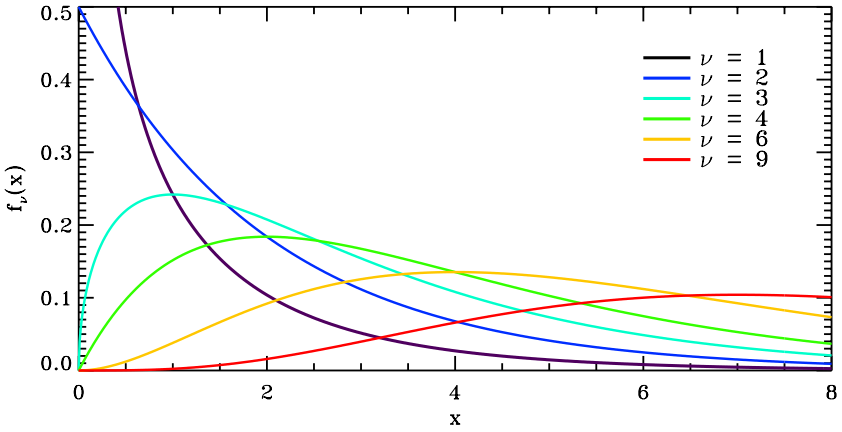

Figure C.1: *Probability density function* $f_\nu(x)$*, where* ν *is the number of degrees of freedom.*

The chi-square distribution is of particular interest when dealing with quantities that are Gaussian distributed. In fact, the minimization of a chi-square function is an useful and practical method for the derivation of the best model parameters. Consider a model $\mathcal{M}$ with ν free parameters ξ_j in order to fit a set of N independent and identically distributed (i.i.d.) random variables, namely the observations $\{d_i{}^{\text{obs}}\}$, having uncertainties $\{\sigma_i\}$ corresponding to the standard deviation of their normal distributions. Then, the χ^2 function to be built is represented by

$$\chi^2\left(\xi_1, \xi_2, \ldots, \xi_\nu\right) = \sum_{i=1}^{N} \left[\frac{d_i{}^{\text{obs}} - d_i{}^{\text{th}}\left(\xi_1, \xi_2, \ldots, \xi_\nu\right)}{\sigma_i} \right]^2, \tag{C.3}$$

[1]The gamma function is defined as $\Gamma(n) = \int_0^\infty t^{n-1} e^{-t} dt$, con $n > 0$.

Table C.2: *Most useful values of $\Delta\chi^2$ for the first three degrees of freedom. Corresponding confidence levels and confidence coefficients are also reported.*

CL	z_c (σ)	$\Delta\chi^2_{\nu=1}$	$\Delta\chi^2_{\nu=2}$	$\Delta\chi^2_{\nu=3}$
50 %	0.6745	0.46	1.39	2.37
68.27 %	1.00	1.00	2.30	3.53
90 %	1.645	2.71	4.60	6.25
94.45 %	2.00	4.00	6.18	8.02
99 %	2.58	6.64	9.21	11.34
99.73 %	3.00	9.00	11.83	14.16

where ${d_i}^{\text{th}}$ is the value derived according to the model adopted, which clearly depends on the ν free parameters ξ_j. Hence, the chi-square function here considered has ν d.o.f..

In analogy to the values presented in Table C.1, it is possible to correlate the probability given by the area under the chi-square distribution (namely the confidence level if we refer to an area expressed in percentage), and the increasing amount in χ^2 from its minimum, represented by the x parameter of Eq. (C.1) and usually known as the percentile $\Delta\chi^2$ (see Figure C.2). Intuitively, the more we increase $\Delta\chi^2$, hence a larger confidence level, the higher becomes the probability that by repeating our measurements infinite times ($N \to \infty$), the limit result will have a value of χ^2 laying in the interval χ^2_{min} to $\chi^2_{\text{min}} + \Delta\chi^2$. Computations to derive confidence levels for the chi-square distribution can be easily made for different d.o.f. and are shown in Table C.2 for some of the most useful cases, as derived with a small IDL code developed by the author for this purpose[2]. Tables reporting values of $\Delta\chi^2$ as a function of the number of d.o.f. and the confidence level can be found in [281, 252]. It is clear that by increasing the number of d.o.f. the chi-square distribution flattens and therefore the percentile required for attaining the same confidence level becomes larger.

C.1.2 Confidence Limits for a Gaussian Likelihood Distribution

Now that the definition of confidence level has been introduced, and the computation in the case of chi-square distribution given, we focus on an application of particular interest for the statistical inference that will be of help for a proper understanding of the Bayesian point of view, i.e. the calculation of confidence levels in the case of a likelihood function. Let us assume we have a model $\mathcal{M}$, whose hypotheses are formalized with the ν free parameters ξ_j (briefly the parameter vector $\boldsymbol{\xi}$), and a set of N i.i.d. measurements $\{d_i\}$, whose distributions we denote as $p(d_i \mid \boldsymbol{\xi})$, for $i = 1, 2, \ldots, N$. The likelihood function $\mathcal{L}(\boldsymbol{\xi} \mid d, \mathcal{M})$ of the ν free parameters ξ_j, given the N measurements $d_1, d_2, \ldots, d_N$, is then built according to the definition of the joint probability that N independent events occur simultaneously, namely the product of the probabilities that each single event occurs

$$\mathcal{L}(\boldsymbol{\xi} \mid d, \mathcal{M}) = \prod_{i=1}^{N} p(d_i \mid \boldsymbol{\xi})\,. \tag{C.4}$$

Assuming that each measurement d_i is normally distributed with variance σ_i, as it may happen in most of the cases when dealing with measurements of some physical quantity, then the

[2]The results were derived by integrating Eq. (C.2) with a direct Romberg integration over a fine grid of $2 \cdot 10^3$ values of $\Delta\chi^2$, in the interval 0 to 15. The choice of this range of values allows to cover approximately the entire interval of the distribution for a number of d.o.f. up to $\nu = 3$. The integration produces a cumulative distribution function for each d.o.f.. The desired $\Delta\chi^2$ value arising from a given CL is then derived by means of a cubic interpolation of the cumulative distribution function considered.

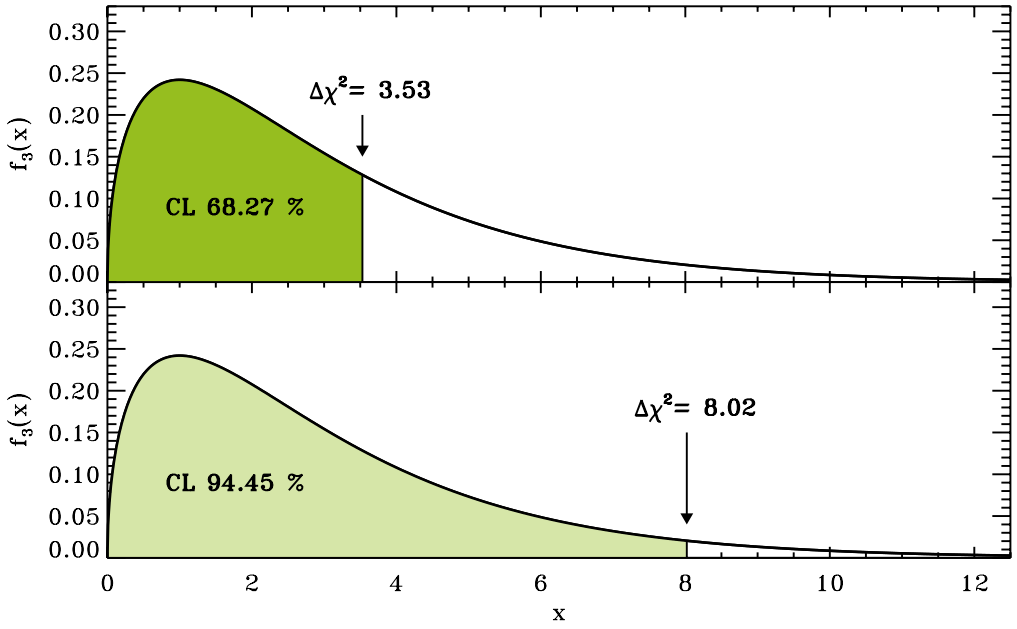

Figure C.2: *Probability density function $f_3(x)$ for a chi-square with three degrees of freedom. Confidence levels of 68.27 % (1σ) and 94.45 % (2σ) are drawn. Percentile values of χ^2 are also reported with an arrow at the end of the interval considered.*

likelihood function is built as a product of N Gaussian distributions and is expressed as[3]

$$\begin{aligned}\mathcal{L}(\boldsymbol{\xi} \mid \boldsymbol{d}, \mathcal{M}) &= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{d_i}{\sigma_i}\right)^2\right] \\ &= \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-\frac{1}{2}\left(\frac{{d_i}^{\mathrm{obs}} - {d_i}^{\mathrm{th}}(\boldsymbol{\xi})}{\sigma_i}\right)^2\right],\end{aligned} \tag{C.5}$$

where we put $d_i = {d_i}^{\mathrm{obs}} - {d_i}^{\mathrm{th}}(\boldsymbol{\xi})$, namely the difference between the observations and the predictions for a given physical quantity[4]. The main property of the likelihood function is that the position of its maximum in the parameter space built according to the free parameters ξ_j, is an estimate of the most likely value of $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_\nu)$, hereafter denoted as $\bar{\boldsymbol{\xi}} = (\bar{\xi_1}, \bar{\xi_2}, \ldots, \bar{\xi_\nu})$. This is done by solving the maximization problem of the likelihood function,

[3] As already mentioned in Section 5.3, the likelihood function ought to reflect the nature of the data, i.e. the way data are collected by the observer. In most of the cases, the residuals arising from the difference between observations and predictions appear to be normal distributed, leading us to the choice of a Gaussian likelihood as the best representative one. However, Gaussian likelihoods also rely on the assumption that the uncertainties involved in the problem are correct, that is we consider them to be derived properly, which might not be always the case especially when dealing with quantities that suffer from some model-dependent method to be measured. Any systematic contained in the data would be responsible of distorting the Gaussianity of each of the i.i.d. random variables considered. In order to overcome the problem of the uncertainties, more robust functional forms of the likelihood function ought to be chosen, as for instance the Median likelihood introduced by [132] and also adopted in [38] for cosmological studies.

[4] Referring to the notation used in the text, when measuring some physical quantity d_i, the least assumption we can do is to consider our set of observations ${d_i}^{\mathrm{obs}}$ to be log-normal distributed, which ensures us to deal correctly with the uncertainties of the problem (see also the discussion in Chapter 6).

which results into the set of ν equations

$$\frac{\partial \mathcal{L}}{\partial \xi_j} = 0 \quad j = 1, 2, \ldots, \nu \,. \tag{C.6}$$

A more useful method, and of much easier numerical handling, is represented by the minimization of the so-called *log-likelihood* function[5] $\Lambda \equiv \ln \mathcal{L}$ (e.g. see [43, 302, 259, 303] for more details). In fact, by considering the natural logarithm of $\mathcal{L}$ we arrive at the Gaussian log-likelihood

$$\Lambda\left(\boldsymbol{\xi}\right) = -\sum_{i=1}^{N} \ln \sqrt{2\pi}\sigma_i - \frac{1}{2}\sum_{i=1}^{N}\left[\frac{d_i{}^{\mathrm{obs}} - d_i{}^{\mathrm{th}}\left(\boldsymbol{\xi}\right)}{\sigma_i}\right]^2 . \tag{C.7}$$

The first term on the right-hand side of Eq. (C.7) is a constant, hereafter Λ_0, while the second term, in case the variances do not depend upon the free parameters ξ_j, represents the χ^2 function with ν degrees of freedom given by Eq. (C.3). Hence, briefly $\Lambda\left(\boldsymbol{\xi}\right) = \Lambda_0 - \frac{1}{2}\chi^2\left(\boldsymbol{\xi}\right)$. The inference problem becomes

$$\frac{\partial \chi^2}{\partial \xi_j} = -2\frac{\partial \Lambda}{\partial \xi_j} = -2\frac{\partial \ln \mathcal{L}}{\partial \xi_j} = -\frac{2}{\mathcal{L}_{\max}}\frac{\partial \mathcal{L}}{\partial \xi_j} \quad j = 1, 2, \ldots, \nu \,, \tag{C.8}$$

where $\mathcal{L}_{\max} = \mathcal{L}(\bar{\boldsymbol{\xi}})$. Therefore, when $\sigma_i \neq \sigma_i(\xi_j)$ for any j and i, the χ^2 function, and the Gaussian likelihood and log-likelihood functions are directly correlated, hence their choice becomes equivalent for estimating the most likely set of values $\bar{\boldsymbol{\xi}}$. The properties of the chi-square distribution (including the derivation of the confidence limits) can be either applied to the case of a likelihood or log-likelihood function. In this manner confidence intervals for parameter estimation are derived by simply converting the likelihood function into a chi-square function with the relation mentioned above. A given confidence level will correspond to a region in parameter space where the likelihood (or log-likelihood) drops (or increases) by a specified amount (see Table C.2). The correlation discussed is of great utility for methods based on likelihood functions, e.g. the Maximum Likelihood Estimation (MLE) method [43, 302, 259].

For the sake of completeness, we consider a one-dimensional case of a parameter ϕ that represents the hypotheses of a given model. Let us assume to take into account a function of the parameter, $f(\phi)$. As the confidence intervals are random variables themselves in the context of frequentist parameter estimation, given they are known for ϕ, one can derive the corresponding confidence intervals for $f(\phi)$ by means of the error propagation law, which yields

$$\sigma_f = \left(\frac{df}{d\phi}\right)\sigma_\phi \,, \tag{C.9}$$

where we denoted as σ_ϕ the confidence intervals for ϕ and with σ_f the corresponding confidence intervals for $f(\phi)$.

C.2 Credible intervals

As mentioned at the end of the section above, an important feature of the frequentist approach is that confidence intervals are considered random variables themselves, i.e. they provide a

[5] In statistical inference the log-likelihood function is greatly more favored than the likelihood one because it is computationally easier to handle with logarithmic quantities since very high values (or conversely very low values) are then flattened. This allows to overcome underflow and overflow errors. A practical example of this case is described in Section C.3 below.

range of values within which we expect our parameter to fall in a known percentage of the times by repeating our measurements $N \to \infty$ times. Conversely, in the Bayesian context we refer to so-called *credible intervals*, which are evaluated only from the marginal PDF, obtained as described in Section 5.3. In this case, if a credible interval contains for example the 94.45 % of the posterior probability mass (i.e. the area under the distribution), then it represents our degree of belief about the value of the parameter itself. It often happens that credible intervals are termed confidence intervals because when adopting Gaussian likelihoods, as in most of the cases, the two results are formally identical, though their interpretation is profoundly different (see Section C.2.1 for a practical example).

But how can we compute credible intervals from a marginal PDF of some physical quantity? We shall turn into that in a moment. Before proceeding with the discussion, it is worth to highlight an important difference between the frequentist and the Bayesian minded way of considering the proper estimator for a given parameter. Assuming to have some marginal PDF of a quantity to be estimated, according to the Bayesian's principle of the degree of belief, the outcoming value of interest for a Bayesian is not the one corresponding to the modal value of the distribution (i.e. the most likely value of the parameter, as conversely frequentists do), but the mean value (or first momentum) of the distribution, also called the *expectation value* or simply the *expectation* of the parameter [44]. Nonetheless, some authors consider the median value of the marginal PDF (e.g. see [303]) — found as the value that divides the area under the distribution into two equal parts — as the most credible one despite of its more complicated derivation. The choice of the latter one usually depends upon the skewness of the distribution, which is responsible for making median and mean values differ from one another. Although, in general, the median estimator is able to provide a result that is less affected than the mean by the presence of any outliers in the distribution, the latter represents in any case the best parameter estimate that a Bayesian may claim for, as discussed in [44]. The most common estimators are defined in the following for the one-dimensional case of a quantity ϕ:

$$\phi_{\text{modal}} : p(\phi_{\text{modal}}) = \max_{\phi} p(\phi) \tag{C.10}$$

$$\phi_{\text{median}} : \int_{\phi_{\text{min}}}^{\phi_{\text{median}}} p(\phi) d\phi = \int_{\phi_{\text{median}}}^{\phi_{\text{max}}} p(\phi) d\phi = \frac{1}{2} \tag{C.11}$$

$$\phi_{\text{mean}} = \int_{\phi_{\text{min}}}^{\phi_{\text{max}}} p(\phi) \phi d\phi \,, \tag{C.12}$$

where $p(\phi)$ is the marginal PDF of the parameter.

Once this preamble has been done, we are now able to describe how to derive numerically the credible intervals of some marginal PDF. Related to this derivation, it important to mention that different definitions of Bayesian credible intervals are available in literature, as noted by [44]. For the computation presented in this dissertation we adopt the definition of the *shortest* credible intervals, which consist in estimating the intervals by starting from the maximum of the marginal PDF. In this manner, we ensure to include the largest values of the distribution for first, hence the resulting credible intervals will be the shortest possible[6]. For simplicity, we consider the example of a dimensionless quantity ϕ, defined in the interval $[2, 12]$.

[6] Other authors define the credible intervals by starting from the tail points of the distribution, which makes them the largest intervals possible because the smallest values of the marginal PDF are considered for first. However, the shortest credible intervals represent the most common choice for most of the Bayesian minded statisticians.

C.2.1 Gaussian Marginal PDFs

As a first case, let us assume that the marginal PDF $p(\phi)$ is a simple Gaussian function. This situation is often encountered when adopting Gaussian likelihood functions and flat priors, as for the Bayesian parameter estimation described in Chapter 6. For our purpose here, we consider the Gaussian marginal PDF

$$p(\phi) = \frac{1}{\sqrt{2\pi}\sigma_\phi} \exp\left[-\frac{1}{2}\left(\frac{\phi - \phi_0}{\sigma_\phi}\right)^2\right] \tag{C.13}$$

where we put $\phi_0 = 7$ and $\sigma_\phi = 1$. The result is shown in the left panel of Figure C.3, where the expectation value of ϕ is marked with a dashed line. For this symmetric case, we have that $\phi_{\text{mean}} = \phi_{\text{modal}} = \phi_{\text{median}}$. The shortest Bayesian credible region marked with a shaded band is identified by the credible limits (red points) found with our numerical method. The quantity h denotes the dropping amount in height of the marginal PDF from its maximum to the value delimited by the credible limits (red line). The ordinates of the upper and lower credible limits must be the same because they identify equal values of the PDF by definition.

The cumulative histogram of the marginal PDF given by Eq. (C.13), namely the area under the distribution, denoted as P_1, is shown as a function of the height h in the right panel of Figure C.3. The function $P_1(h)$ can be expressed as

$$P_1(h) = \int_{\phi_{c_1}(h)}^{\phi_{\text{modal}}} p(\phi)d\phi + \int_{\phi_{\text{modal}}}^{\phi_{c_2}(h)} p(\phi)d\phi \tag{C.14}$$

whose values are probabilities, hence normalized to unity. The terms $\phi_{c_1}(h)$ and $\phi_{c_2}(h)$ represent the ordinates of lower and upper credible limits, respectively, which clearly depend upon the height h. The integral is split into two parts for highlighting the inclusion of the modal value of the distribution within the range of the credible intervals. The integral can be computed numerically by means of a simple trapezoidal summation over the range $[c_1(h), c_2(h)]$. The parameter h is chosen in the range $[0, \max_\phi p(\phi)]$, which allows us to compute all the possible values of $P_1(h)$, varying from 0 to 1.

Therefore, we first compute $P_1(h)$ for each value of h, assuming a step $\Delta h \sim 10^{-3} \max_\phi p(\phi)$. As a second step, supposed we are interested in 68.27 % credible intervals, we identify the value h^* corresponding to $P_1(h^*) = 0.6827$ in the cumulative histogram by means of either a cubic or linear interpolation. Afterwards, we divide the marginal PDF into two parts, the first ranging from ϕ_{min} to ϕ_{modal} (left distribution) and the second from ϕ_{modal} to ϕ_{max} (right distribution). The 68.27 % credible limits $\phi^*_{c_1}$ and $\phi^*_{c_2}$ are found by means of either a cubic or linear interpolation to the left and right distributions, respectively, in order to satisfy the condition

$$p(\phi^*_{c_1}) = p(\phi^*_{c_2}) = \max_\phi p(\phi) - h^* \,. \tag{C.15}$$

The expectation value ϕ_{mean} is then reported together with the credible intervals $\phi^{\pm}_{\text{CI}} = \left|\phi_{\text{mean}} - \phi^*_{c_1}\right| = \left|\phi_{\text{mean}} - \phi^*_{c_2}\right|$, which are symmetric with respect to the mean (or equivalently the mode). For the marginal PDF given by Eq. (C.13), the 68.27 % Bayesian credible intervals are formally identical to the frequentist 1σ confidence interval computed for the same distribution, although their meanings remain completely different from one another. In particular, we describe a substantial difference between the two definitions (frequentist and Bayesian ones) by considering once again the case of a function $f(\phi)$ for which we want to estimate the credible intervals. In fact, by adopting Eq. (5.9) introduced in Chapter 5, we have that

$$p\,(f) = p(\phi)\left|\frac{d\phi}{df}\right| , \tag{C.16}$$

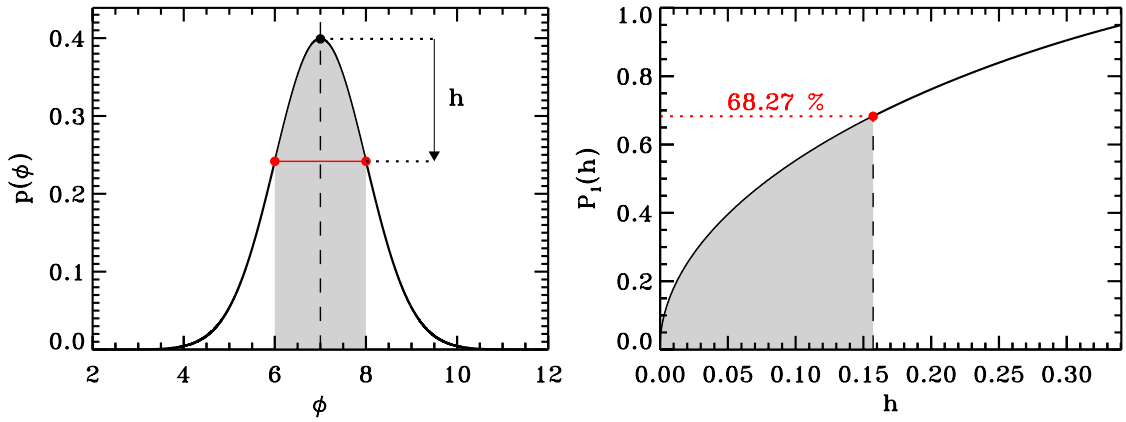

Figure C.3: *Left panel: Gaussian (symmetric) marginal PDF of a dimensionless parameter ϕ. 68.27 % Bayesian credible region is marked with a shaded band. The mean value of ϕ is reported with a dashed vertical line. Red points represent the credible limits to the shaded region, while the arrow indicates the dropping amount in height of the distribution from its maximum (black point) to the value corresponding to the credible limits (red line). Right panel: Cumulative distribution of the area under the marginal PDF as computed from its maximum ($h = 0$). The shaded region corresponds to the amount in h that gives the 68.27 % of the total area (red point and red dashed line).*

namely, since we are dealing with PDFs, first we must evaluate the corresponding marginal PDF $p\,[f(\phi)]$, then we apply to it the method described above. As it appears clear, this procedure is completely different from the one described for the frequentist case in Section C.1.2. In fact, Bayesian credible intervals are not random variables: they represent our degree of belief upon the model investigated or, in simpler words, the "elbow" of the model.

C.2.2 Generalization to Asymmetric Marginal PDFs

We have seen the case of a symmetric marginal PDF, which is clearly the simplest situation we can deal with. In fact, it may happen that the Bayesian inference problem involves either non-uniform priors or a likelihood function that differs from the Gaussian one. This may lead us to a final marginal PDF having an irregular shape. For treating this case, we consider the example of an asymmetric marginal PDF for the parameter ϕ, expressed as

$$p(\phi) = \frac{1}{2\sqrt{2\pi}} \left\{ \frac{1}{\sigma_{\phi,1}} \exp\left[-\frac{1}{2} \left(\frac{\phi - \phi_1}{\sigma_{\phi,1}} \right)^2 \right] + \frac{1}{\sigma_{\phi,2}} \exp\left[-\frac{1}{2} \left(\frac{\phi - \phi_2}{\sigma_{\phi,2}} \right)^2 \right] \right\} \tag{C.17}$$

where we put $\phi_1 = 7$, $\sigma_{\phi,1} = 1$, $\phi_2 = 6$, $\sigma_{\phi,2} = 0.4$. The result is shown in the left panel of Figure C.4, which adopts the same notation to that used for Figure C.3. As one can see immediately, the expectation value of ϕ (dashed vertical line) is now differing from $\phi_{\rm modal}$ (black dot) because the distribution is not symmetric anymore.

Altogether, we are still able to compute the confidence limits with the same method used for the symmetric case, because we are able to evaluate the cumulative distribution $P_1(h)$ straightforwardly (right panel of Figure C.4). The only difference to the previous case is that now the values $\phi^*_{c_1}$ and $\phi^*_{c_2}$ are not symmetric with respect to $\phi_{\rm modal}$. Therefore, the expectation value $\phi_{\rm mean}$ has to be reported with two different credible intervals, computed as $\phi^-_{\rm CI} = \phi_{\rm mean} - \phi^*_{c_1}$ and $\phi^+_{\rm CI} = \phi^*_{c_2} - \phi_{\rm mean}$. It is obvious that, for this asymmetric case the analogy with the Gaussian confidence limits does not hold anymore.

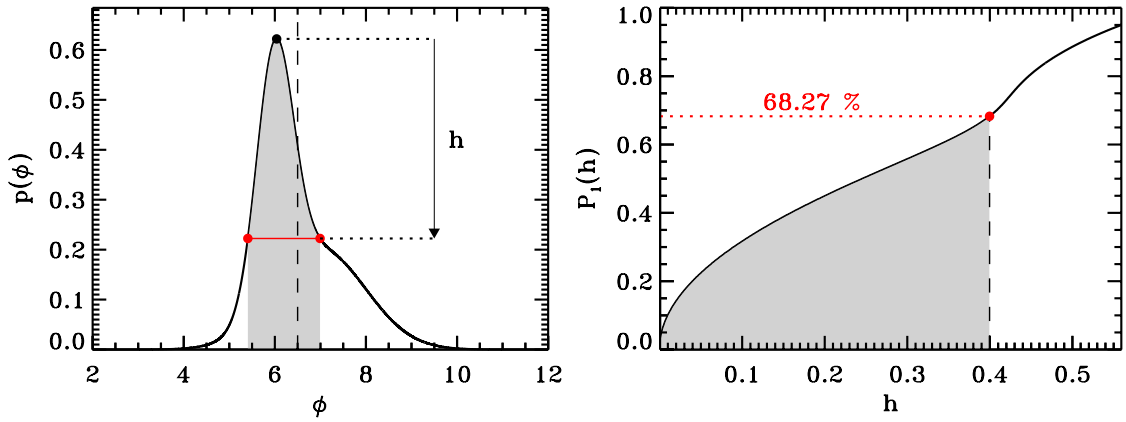

Figure C.4: *Left panel: asymmetric marginal PDF of a dimensionless parameter ϕ. 68.27 % Bayesian credible region is marked with a shaded band. The mean value of ϕ is reported with a dashed vertical line. Red points represent the credible limits to the shaded region, while the arrow indicates the dropping amount in height of the distribution from its maximum (black point) to the value corresponding to the credible limits (red line). Right panel: Cumulative distribution of the area under the marginal PDF as computed from its maximum ($h = 0$). The shaded region corresponds to the amount in h that gives the 68.27 % of the total area (red point and red dashed line).*

C.3 Computing the Bayesian evidence

In most of the problems encountered in Bayesian statistics, we are often unable to compute analytically he Bayesian evidence introduced by Eq. (5.29), hereafter $\mathcal{E}_{\mathcal{M}}$. This happens because the integrand function of Eq. (5.29) has usually a complicated dependence upon the free parameters of the problem, and because a large number of free parameters, i.e. of dimensions, is taken into account. A nice example concerning this aspect is given can be found in the case of a likelihood function given by Eq. (6.27), which was used for the set of amplitude measurements investigated in Chapter 6 . In fact, a particular numerical method has been adopted for computing $\mathcal{E}_{\mathcal{M}}$ in the k-dimensional cases up to $k = 4$, as we shall describe in the following. From now on, we shall adopt the same notation used in Chapter 6, referring to the parameter vector $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_k\}$ and to the models there discussed, which we will simply denote as $\mathcal{M}$ in general.

First of all, in order to avoid underflow errors in our computations, we consider a modified version of the log-likelihood $\Lambda(\boldsymbol{\xi})$ given by Eq. (C.7), which is

$$\widehat{\Lambda}(\boldsymbol{\xi}) \equiv \Lambda(\boldsymbol{\xi}) - \Lambda_{\max}\,, \tag{C.18}$$

where $\Lambda_{\max}$ is simply the maximum value of Λ found in the inference problem (see Section 6.3). Since Λ shows very negative values ($\sim -10^3$ dex), as reported in Tables 6.2, 6.3, and 6.4, by subtracting its maximum we can deal with larger values, numerically easier to be handled. Thus, the likelihood function $\mathcal{L}(\boldsymbol{\xi})$ can be expressed as

$$\mathcal{L}(\boldsymbol{\xi}) = A \exp\left[\widehat{\Lambda}(\boldsymbol{\xi})\right] = A\,\widehat{\mathcal{L}}(\boldsymbol{\xi})\,, \tag{C.19}$$

where $A = \exp(\Lambda_{\max})$ is a constant and $\widehat{\mathcal{L}}(\boldsymbol{\xi})$ is the modified likelihood function.

After this preliminary manipulation, we can comeback to the problem of computing $\mathcal{E}_{\mathcal{M}}$. Let us assume to include flat priors (i.e. uniform priors made proper, see the discussion in Chapter 5) in our problem, as it happens in many cases studied indeed, especially for the one

Table C.3: *Natural logarithms of Bayesian evidence* $\ln \mathcal{E}_{\mathcal{M}}$ *for each model as derived by means of Eqs. (C.22) and for the three cases of sample considered in Chapter 6. The shaded row shows the model having the largest evidences among the others.*

Model	$\ln \mathcal{E}_{\mathcal{M}}$	$\ln \mathcal{E}_{\mathcal{M}}^{(\mathrm{SC})}$	$\ln \mathcal{E}_{\mathcal{M}}^{(\mathrm{LC})}$
$\mathcal{M}_1$	−3544.7	−331.1	−808.8
$\mathcal{M}_{1,\beta}$	−2507.2	−213.0	−750.3
$\mathcal{M}_2$	−4171.4	−453.5	−778.6
$\mathcal{M}_4$	−962.4	−105.9	194.4
$\mathcal{M}_{4,\beta}$	145.2	2.9	284.7
$\mathcal{M}_5$	−930.8	−87.5	−363.1
$\mathcal{M}_{5,\beta}$	−796.9	−89.1	−318.0
$\mathcal{M}_6$	−735.6	−86.8	−363.9
$\mathcal{M}_{6,\beta}$	−735.4	−70.4	−318.6

dealing with the amplitude scaling relations discussed in Chapter 6. Since flat priors have to satisfy the normalization condition

$$\int_{\Omega_{\mathcal{M}}} \pi\left(\boldsymbol{\xi} \mid \mathcal{M}\right) d\boldsymbol{\xi} = 1 \tag{C.20}$$

we obtain that $\pi(\boldsymbol{\xi} \mid \mathcal{M}) = \mathcal{V}_{\mathcal{M}}^{-1}$, where

$$\mathcal{V}_{\mathcal{M}} = (\xi_1^{\max} - \xi_1^{\min}) \cdot (\xi_2^{\max} - \xi_2^{\min}) \ldots (\xi_k^{\max} - \xi_k^{\min}) \,. \tag{C.21}$$

is the hyper-volume of the parameter space $\Omega_{\mathcal{M}}$, namely the simple product of the intervals considered for the free parameters. As a note of caution, one should keep in mind that when computing the evidence for a set of models, the intervals of the free parameters must be the same for all the models considered. In addition, the hyper-volume $\mathcal{V}_{\mathcal{M}}$ depends upon the dimensionality of the model because the corresponding parameter space $\Omega_{\mathcal{M}}$ may differ from one model to another according to the free parameters involved.

Therefore, taking into account the advantage of having flat priors, and by means of Eq. (C.19), Eq. (5.29) can be rearranged to provide a logarithmic expression of the evidence, i.e.

$$\ln \mathcal{E}_{\mathcal{M}} = \ln \int_{\Omega_{\mathcal{M}}} \widehat{\mathcal{L}}(\boldsymbol{\xi}) d\boldsymbol{\xi} + \Lambda_{\max} - \ln \mathcal{V}_{\mathcal{M}} \,. \tag{C.22}$$

which is extremely useful for our computations. In fact, the integral appearing in the right-hand side of Eq. (C.22), namely the multidimensional integral of the modified likelihood function over the entire parameter space, can be evaluated with a direct numerical integration for up to $k = 4$ dimensions at least, as we shall discuss in the section below.

Referring to the study described in Chapter 6, the natural logarithms of the evidence, $\ln \mathcal{E}_i$, for each model M_i, have been computed according to the method described in this section and are listed in Table C.3 for the three samples described in Chapter 6, Section 6.1.

C.3.1 Multidimensional Integration

As discussed in [259], multidimensional integrals are often a very difficult task by a numerical point of view. However, the problem can be addressed in most of the cases, depending on the complexity of the boundary of the region of integration, the smoothness of the integrand function and the accuracy level one wants to reach [259].

For the analysis described in Chapter 6, we deal with a very simple boundary, namely the one of the parameter space $\Omega_{\mathcal{M}}$ given by Eq. (5.12), a smoothed function (i.e. a multidimensional Gaussian profile), and we aim at reaching a good accuracy level in our computations as it is fundamental for a reliable model comparison. According to the numerical methods described in [259], we split the multidimensional integration into a series of one-dimensional integrals, which represent one of the most effective and fast methods to adopt for solving our problem. For explaining the method in more detail, let consider the example of a function $f(\xi_1, \xi_2, \xi_3)$, depending on the three free parameters ξ_i. Suppose we are interested in estimating the three-dimensional integral

$$\begin{aligned} I \equiv \iiint_{\Omega_{\mathcal{M}}} & f(\xi_1, \xi_2, \xi_3) d\xi_1 d\xi_2 d\xi_3 \\ & = \int_{\xi_1^{\min}}^{\xi_1^{\max}} d\xi_1 \int_{\xi_2^{\min}}^{\xi_2^{\max}} d\xi_2 \int_{\xi_3^{\min}}^{\xi_3^{\max}} f(\xi_1, \xi_2, \xi_3) d\xi_3 \end{aligned} \tag{C.23}$$

where we considered the simple parameter space

$$\Omega_{\mathcal{M}} = \left[\xi_1^{\min}, \xi_1^{\max}\right] \times \left[\xi_2^{\min}, \xi_2^{\max}\right] \times \left[\xi_3^{\min}, \xi_3^{\max}\right] . \tag{C.24}$$

We can define a function $g(\xi_1, \xi_2)$ that does the innermost integral, namely

$$g(\xi_1, \xi_2) = \int_{\xi_3^{\min}}^{\xi_3^{\max}} f(\xi_1, \xi_2, \xi_3) d\xi_3 \,, \tag{C.25}$$

and subsequently another function $h(\xi_1)$ that does the integral of $g(\xi_1, \xi_2)$

$$h(\xi_1) = \int_{\xi_2^{\min}}^{\xi_2^{\max}} g(\xi_1, \xi_2) d\xi_2 \,. \tag{C.26}$$

The integral I is then simply computed as the one-dimensional integral

$$I = \int_{\xi_1^{\min}}^{\xi_1^{\max}} h(\xi_1) d\xi_1 \,. \tag{C.27}$$

This method can be, in principle, extended to any dimension by defining iteratively new functions that depend upon a reduced number of free parameters. For each one-dimensional computation, we have adopted the IDL routine **INT_TABULATED**, which evaluates one-dimensional integrals for tabulated functions, i.e. a given vector containing the values of the function to be integrated. In particular, for the evidence given by Eq. (C.22), the integrand function $\widehat{\mathcal{L}}(\boldsymbol{\xi})$ is basically that of Eq. (6.27) — apart from the constant term $\exp(-\Lambda_{\max})$ — which we tabulated for a grid of values for all the models described in Chapter 6. Thus, the method described above can be applied to the case of $\widehat{\mathcal{L}}(\boldsymbol{\xi})$ — given in the form of a multidimensional matrix — by simply passing to the routine a one-dimensional matrix depending on the free parameter we want to integrate for, and repeating this procedure iteratively within a loop that fixes at each step different values of the remaining variables upon which the function depends.

The method has proved to be very efficient and fast for up to 10^6 tabulated values of the integrand function. The resulting integrals have been compared to those derived from integrations made by means of MCMC simulations to the posterior PDF, showing a very good agreement.

C.3.2 The Newton-Cotes Formulas

For gaining more feeling about the integration technique implemented in the IDL routine **INT_TABULATED** we will explain it in the following. The routine exploits the so-called *five point Newton-Cotes* formula over a closed intervals of values. Such formula belongs to the family of Newton-Cotes formulas[7], or quadrature formulas, of numerical integration techniques, and it is also known as the *Boole's rule*.

We consider the general example of a function $f(x)$ to be integrated over the closed interval $[a, b]$. For a n-point Newton-Cotes formula, we divide the integral into $n-1$ parts having same length, $h = (b-a)/(n-1)$, which determines the points $x_1 = a, x_2 = a+h, x_3 = a+2h, \ldots, x_{n-1} = a+(n-2)h, x_n = a+(n-1)h = b$. We then define the function in the regularly spaced points x_i so that $f_i = f(x_i)$. Thus, we have the analytical expression

$$\int_a^b f(x)dx = \int_{x_1}^{x_n} P_n(x)dx = h\sum_{i=1}^{n} H_{n,i} f_i \,, \tag{C.28}$$

where

$$H_{n,r+1} = \frac{(-1)^{n-r}}{r!(n-r)!}\int_0^n t(t-1)\ldots(t-r+1)(t-r-1)\ldots(t-n)dt \,. \tag{C.29}$$

In particular, $P_n(x)$ is the so-called *Lagrange Interpolating Polynomial*[8] and it is a polynomial of degree $\leq (n-1)$ that passes through the n points defined, in order to reproduce the function $f(x)$ in the given range. We have that

$$P_n(x) = \sum_{j=1}^{n} P_j(x) \,, \tag{C.30}$$

where

$$P_j(x) = \prod_{k=1, k\neq j}^{n} \frac{x - x_k}{x_j - x_k} f_j \tag{C.31}$$

In the case of $n = 5$, which is used by **INT_TABULATE**, the Boole's rule is then expressed as

$$\int_a^b f(x)dx = \int_{x_1}^{x_1+4h} P_5(x)dx = \frac{2}{45}h\left[7(f_1+f_5) + 32(f_2+f_4) + 12f_3\right] \,, \tag{C.32}$$

where the function is replaced by the fourth-order Lagrange interpolating polynomial, $P_5(x)$. When adopting grids that present a large number of values (e.g. $n > 100$), **INT_TABULATED** computes a new grid of values having $n_{\rm new} > n$ so that $(n_{\rm new} \bmod 4) = 0$, by means of a cubic-spline interpolation to the values of the tabulated function. Subsequently, the Boole's rule is applied consecutively to each group of five points of the new grid of values, from x_1 to $x_{n_{\rm new}}$. Since the method is repeated $m = n_{\rm new}/4$ times, the final integral is thus given as

$$\int_a^b f(x)dx = \sum_{i=0}^{m-1} \int_{x_1+i4h}^{x_1+(i+1)4h} P_5(x)dx \,. \tag{C.33}$$

[7] Weisstein, Eric W. "Newton-Cotes Formulas." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/Newton-CotesFormulas.html.

[8] Archer, Branden and Weisstein, Eric W. "Lagrange Interpolating Polynomial." From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/LagrangeInterpolatingPolynomial.html.

In cases the integrand function is not very smooth within each interval $[x_1+i4h, x_1+(i+1)4h]$, the adoption of a larger number of grid points for the input tabulated function may be required. Lastly, the error to a single application of the Boole's rule is given by

$$\frac{8}{945}h^7 f^{(6)}(\xi)\,, \tag{C.34}$$

ξ being a value within the closed interval $[a, b]$ and $f^{(6)}$ the sixth-order derivative of $f(x)$. For the case described above, having a grid of $n_{\rm new}$ points, the corresponding error in the total integral is

$$\frac{8}{945}mh^7 f^{(6)}(\xi)\,, \tag{C.35}$$

namely m-times the error of a single application, with m the number of groups containing 5-points each.

D Principal Component Analysis

Although numerical techniques for the exploration and sampling of high-dimensionality parameter spaces, such as the MCMC method described in this dissertation, have nowadays reached a high level of performance for attaining the outcomes, it is sometimes useful to optimize their efficiency by introducing analysis techniques that deal with *multivariate statistics* and whose notion is presented in this appendix. The method described here is known as *Principal Component Analysis (PCA)*. Our choice to dedicate an appendix to this method relies on the fact that it is employed within the Bayesian star spot modeling described in Chapter 7. However, the reader should note that this argument can be treated separately from the discussion about Bayesian statistics and that PCA represents a numerical technique aimed at improving the computational speed of the calculations (thus not essential to bring off an entire analysis).

PCA is in origin a mathematical procedure that uses an *orthogonal transformation* to convert a set of observations of possibly correlated variables (e.g. the parameters of a model) into a set of values of uncorrelated variables that are called *principal components (PCs)*. Clearly, the number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has as high a variance as possible (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it be orthogonal to (uncorrelated with) the preceding components.

It is generally accepted that the earliest descriptions of PCA were given in 1901 by Karl Pearson [253] and in 1933 by Harold Hotelling [156]. PCA is now mostly used as a tool in exploratory data analysis and for the developing of predictive models. It is analytically performed through an *eigenvalue decomposition (ED)* of a data covariance matrix, although it can be numerically computed by a *singular value decomposition (SVD)* of a data matrix. Moreover it is the simplest of the true eigenvector-based multivariate analyses. Often, its operation can be thought of as a revealing of the internal structure of the data in a way which best explains the variance in the data. If a multivariate dataset is visualized as a set of coordinates in a high dimensional data space (one axis per variable, as in the case of parameter spaces used within Bayesian inference), PCA can supply the user with a lower-dimensional picture, a "shadow" say, of this space as viewed from its most informative viewpoint. This is done by using only the first few principal components so that the dimensionality of the transformed data is reduced.

However, for an exhaustive explanation of PCA the introduction of some new statistical notions is required, which we briefly illustrate below [143]. In probability theory and statistics, the *multivariate normal distribution* or *multivariate Gaussian distribution*, is a generalization of the one-dimensional (univariate) normal distribution to higher dimensions. A *random vector* $\mathbf{x}$ of elements $x_1, x_2, \ldots, x_p$ is said to be multivariate normally distributed if every linear combination of its components $y = a_1x_1 + a_2x_2 + \cdots + a_px_p$ has a univariate normal distribution. That

is, for any constant vector $\mathbf{a} \in \mathbb{R}^p$, such that $\mathbf{a} = (a_1, a_2, \ldots, a_p)$, the random variable $y = \mathbf{a}'\mathbf{x}$, where $\mathbf{a}'$ denotes the transpose, is normally distributed. The multivariate normal distribution is often used to describe, at least approximately, any set of (possibly) correlated real-valued *random variables*, each of which clusters around a mean value. The reason for introducing these definitions relies on the fact that PCs are guaranteed to be independent only if the data set is *jointly normally distributed*. This means that, considering two random vectors $\mathbf{x}$ and $\mathbf{y}$ to be jointly normally distributed, the pair $(\mathbf{x}, \mathbf{y})$ must have a bivariate (two-dimensional) normal distribution. Lastly, it is a common notation to refer to a multivariate normal distribution of a p-dimensional random vector $\mathbf{x} = (x_1, x_2, \ldots, x_p)$ as $x \sim \mathcal{N}(\mu, \Sigma)$, where μ is the p-dimensional vector of the mean values and Σ is the $(p \times p)$ covariance matrix. In a two-dimensional case the bivariate density function of two vectors $\mathbf{x}$ and $\mathbf{y}$, having mean values μ_x, μ_y and standard deviations σ_x and σ_y respectively, is given by

$$f(x, y) = A \exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu-x)(y-\mu_y)}{\sigma_x\sigma_y}\right)\right], \quad \text{(D.1)}$$

where

$$A = \frac{1}{2\pi\sqrt{(1-\rho^2)}\sigma_x\sigma_y},$$

and ρ is the correlation term of the associated covariance matrix

$$\Sigma = \begin{pmatrix} \sigma_x^2 & \rho\sigma_x\sigma_y \\ \rho\sigma_y\sigma_x & \sigma_y^2 \end{pmatrix}.$$

D.1 Principal Components

As already announced, the central idea of PCA is to reduce the dimensionality of a dataset consisting of a large number of interrelated variables, while retaining as much as possible of the variation contained in the data. This is achieved by transforming the old set of parameters into a new set of variables (PCs), which are completely uncorrelated and ordered so that the first *few* retain most of the variation present in *all* of the original variables.

For the purpose of providing an exhaustive and pivotal explanation of the argument, we will follow the mathematical approach used in [176]. Suppose that $\mathbf{x}$ is a vector of p random variables (e.g. the parameters of a spot modeling) and the structures of covariances or correlations between the p variables are of interest. Therefore, let be $\mathbf{x} \in \mathbb{M}^p$, where $\mathbb{M}^p \subset \mathbb{R}^p$ is a p-dimensional vector space of multivariate normally distributed random vectors. Unless p is small, or the structure is very simple, it will often not be very helpful to simply look at the p variances and all of the $\frac{1}{2}p(p-1)$ correlations or covariances. Alternatively one can look for a few ($\ll p$) derived variables that preserve most of the information given by these variances and correlations or covariances. PCA represents an alternative that concentrates on variances, although covariances and correlations are not ignored.

The first step consists in finding a linear function of the elements of $\mathbf{x}$ (i.e. a linear combination of its elements) having maximum variance, which we will denote as $z_1(\mathbf{x})$, or briefly z_1, where $z_1 \colon \mathbb{M}^p \mapsto \mathbb{R}$. By adopting a vector $\mathbf{a_1} \in \mathbb{R}^p$, of p constants $a_{11}, a_{12}, \ldots, a_{1p}$, we can write the function $z_1(\mathbf{x})$ as the product $\mathbf{a_1'}\mathbf{x}$, where $\mathbf{a_1'}$ is the transpose, so that

$$z_1(\mathbf{x}) = \mathbf{a_1'}\mathbf{x} = a_1x_1 + a_2x_2 + \cdots + a_px_p = \sum_{j=1}^{p} a_{1j}x_j\,. \quad \text{(D.2)}$$

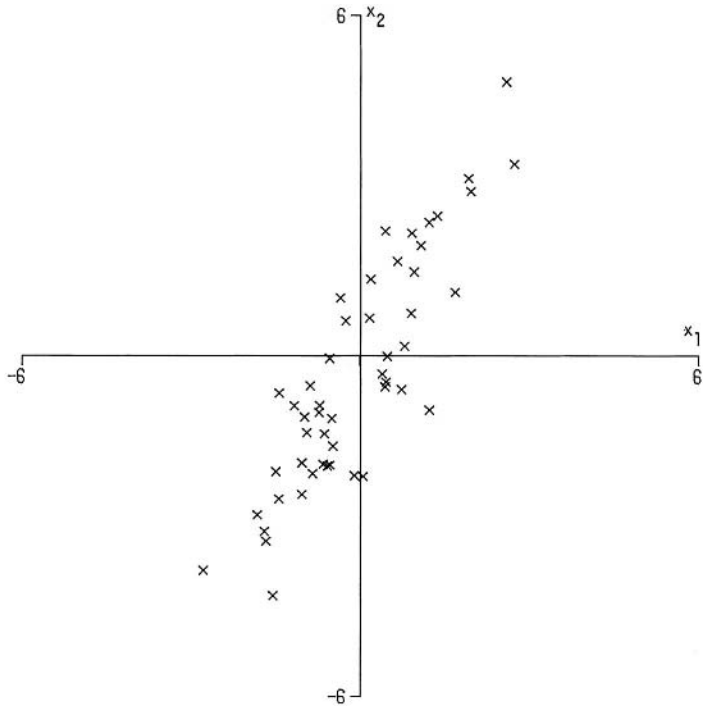

Figure D.1: *Plot of the 50 observations on two variables x_1, x_2.*

The next step is to look for a linear function $z_2(\mathbf{x}) = \mathbf{a_2'x}$, uncorrelated with $z_1(\mathbf{x})$, having maximum variance, and so on, so that at the kth stage a linear function $z_k(\mathbf{x}) = \mathbf{a_k'x}$ is found to have maximum variance subject to being uncorrelated with $z_1(\mathbf{x}), z_2(\mathbf{x}), \ldots, z_{k-1}(\mathbf{x})$. Therefore, the kth derived variable, $\mathbf{a_k'x}$ is the kth PC. Of course, up to p PCs could be found, but it is hoped in general that most of the variation in $\mathbf{x}$ will be accounted for by m PCs, where $m \ll p$. The reduction in complexity attained by transforming the original variables to PCs can be demonstrated by adopting an unrealistic, but simple, case where $p = 2$. The advantage of $p = 2$ is, of course, that the data can be plotted exactly in two dimensions. Assume we have a series of 50 observations[1] on two highly correlated variables x_1, x_2 and that there is a considerable variation in both variables, though rather more in the direction of x_2 than x_1, as shown in Figure D.1. If we transform the variables to the PCs z_1, z_2 we obtain the plot shown in Figure D.2. It is clear that most of the variation occurs in the direction of z_1 than in either of the original variables, but very little variation in the direction of z_2. More generally, if p (> 2) random variables have significant correlation among them, then the first few PCs will account for most of the variation in the original variables. Conversely, the last few PCs identify directions in which there is a very little variation; that is, they identify near-constant linear relationships among the original variables.

D.2 PCs as ED of a Covariance Matrix

Having defined PCs, it is useful to know how to derive them analytically. Let us consider the case where the vector of random variables $\mathbf{x}$ has a known covariance matrix $\mathbf{\Sigma}$. The element Σ_{ij} is the (known) covariance between the ith and jth elements of $\mathbf{x}$ when $i \neq j$ ($\Sigma_{ij} = \sigma_i\sigma_j$),

[1] We generally use the name *observations* for indicating any type of input information. In the case of star spot modeling presented in this thesis, the observations correspond to the sampling points of the posterior pdf.

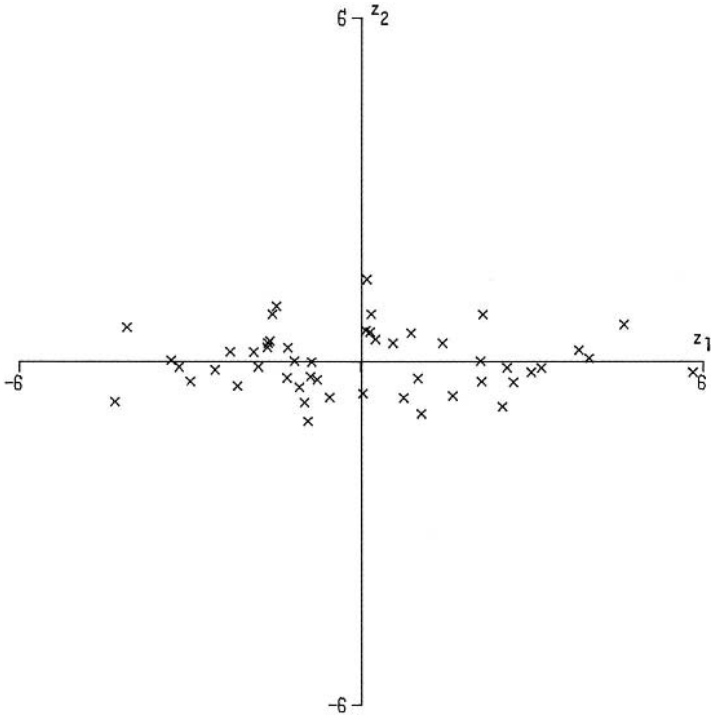

Figure D.2: *Plot of the 50 observations with respect to their PCs z_1, z_2.*

and variance of the jth element of $\mathbf{x}$ when $i = j$ ($\Sigma_{jj} = \sigma_j^2$). In the more realistic case, where $\mathbf{\Sigma}$ is unknown, such matrix is replaced by a sample covariance matrix[2] $\mathbf{S}$. However, it turns out that for $k = 1, 2, \ldots, p$, the kth PC is given by $z_k = \mathbf{a_k'x}$ where $\mathbf{a_k}$ is an eigenvector of $\mathbf{\Sigma}$ corresponding to its kth largest eigenvalue λ_k. Furthermore, if $\mathbf{a_k}$ is chosen to have unit length, i.e. $\mathbf{a_k'a_k} = 1$, then the variance of z_k gives $\text{var}\,[z_k] = \lambda_k$.

The following derivation of PCs is the standard one discussed in many multivariate textbooks. To derive the form of PCs, consider first the product $\mathbf{a_1'x}$, as for the definition; the vector $\mathbf{a_1}$ must maximize the variance $\text{var}\,[\mathbf{a_1'x}] = \mathbf{a_1'\Sigma a_1}$. It is clear that, as it stands, the maximum will not be achieved for a finite $\mathbf{a_1}$ so a normalization constraint must be imposed. The constraint used in this derivation[3] is $\mathbf{a_1'a_1} = 1$, that is, the sum of squares of elements of $\mathbf{a_1}$ equals 1. To maximize $\mathbf{a_1'\Sigma a_1}$ subject to $\mathbf{a_1'a_1} = 1$, the standard approach is to use the technique of Lagrange multipliers. Maximize

$$\mathbf{a_1'\Sigma a_1} - \lambda(\mathbf{a_1'a_1} - 1)\,, \tag{D.3}$$

where λ is a Lagrange multiplier. Differentiation with respect to $\mathbf{a_1}$ gives

$$\mathbf{\Sigma a_1} - \lambda\mathbf{a_1} = \mathbf{0}\,, \tag{D.4}$$

or

$$(\mathbf{\Sigma} - \lambda\mathbf{I}_p)\mathbf{a_1} = \mathbf{0}\,, \tag{D.5}$$

[2] A thorough discussion about this case goes beyond the arguments we intend to fulfill. Further details can be found in [176].

[3] Other constraints can be used, as for example $\max_j \mid a_{1j} \mid$ may be more useful in other circumstances. However, the use of constraints other than $\mathbf{a_1'a_1} = 1$ in the derivation leads to a more difficult optimization problem, and it will produce a set of derived variables different from the PCs.

where $\mathbf{I}_p$ is the $(p \times p)$ identity matrix. Thus, λ is an eigenvalue of $\mathbf{\Sigma}$ and $\mathbf{a_1}$ its corresponding eigenvector. Since we considered a p-dimensional problem, p different eigenvectors will be then derived. To decide which of them gives $\mathbf{a_1'x}$ with maximum variance, note that the quantity to be maximized is

$$\mathbf{a_1'\Sigma a_1} = \mathbf{a_1'}\lambda\mathbf{a_1} = \lambda\mathbf{a_1'a_1} = \lambda\,, \tag{D.6}$$

so λ must be as large as possible. Thus, $\mathbf{a_1}$ is the eigenvector corresponding to the largest eigenvalue of $\mathbf{\Sigma}$, and $\text{var}\,[\mathbf{a_1'x}] = \mathbf{a_1'\Sigma a_1} = \lambda_1$, the largest eigenvalue.

In general, the kth PC of $\mathbf{x}$ is $\mathbf{a_k'x}$ and $\text{var}\,[\mathbf{a_k'x}] = \lambda_k$, where λ_k is the kth largest eigenvalue of $\mathbf{\Sigma}$, and $\mathbf{a_k}$ its corresponding eigenvector. This will now be proved for $k = 2$; the proof for $k \geq 3$ is slightly more complicated, but very similar. Obviously, the second PC, $\mathbf{a_2'x}$, maximizes $\mathbf{a_2'\Sigma a_2}$ subject to being uncorrelated with $\mathbf{a_1'x}$, or equivalently subject to $\text{cov}\,[\mathbf{a_1'x}, \mathbf{a_2'x}] = 0$, where we denoted as $cov(x, y)$ the covariance between the two random variables x and y. Anyhow we should keep in mind that

$$\text{cov}\left[\mathbf{a_1'x}, \mathbf{a_2'x}\right] = \mathbf{a_1'\Sigma a_2} = \mathbf{a_2'\Sigma a_1} = \mathbf{a_2'}\lambda_1\mathbf{a_1} = \lambda_1\mathbf{a_2'a_1} = \lambda_2\mathbf{a_1'a_2}\,. \tag{D.7}$$

Thus, any of the equations $\mathbf{a_1'\Sigma a_2} = 0$, $\mathbf{a_2'\Sigma a_1} = 0$, $\mathbf{a_2'a_1} = 0$, $\lambda_2\mathbf{a_1'a_2} = 0$ could be used to express zero correlation between the two PCs. By choosing the last equation (arbitrary choice) we have to maximize the quantity $\mathbf{a_2'\Sigma a_2}$, but subject to the conditions $\mathbf{a_2'a_2} = 1$ and $\mathbf{a_2'a_1} = 0$. By adopting the technique of the Lagrange multipliers once more, we have to maximize the quantity

$$\mathbf{a_2'\Sigma a_2} - \lambda(\mathbf{a_2'a_2} - 1) - \phi\mathbf{a_2'a_1}\,, \tag{D.8}$$

where λ and ϕ are the two Lagrange multipliers. Differentiation with respect to $\mathbf{a_2}$ gives

$$\mathbf{\Sigma a_2} - \lambda\mathbf{a_2} - \phi\mathbf{a_1} = \mathbf{0}\,, \tag{D.9}$$

and by multiplying on the left by $\mathbf{a_1'}$ we have

$$\mathbf{a_1'\Sigma a_2} - \lambda\mathbf{a_1'a_2} - \phi\mathbf{a_1'a_1} = 0\,, \tag{D.10}$$

which, since the first two terms on the right-hand side are zero and $\mathbf{a_1'a_1} = 1$, reduces to $\phi = 0$. Hence, we obtain again the eigenvalues problem $\mathbf{\Sigma a_2} - \lambda\mathbf{a_2} = 0$ or equivalently $(\mathbf{\Sigma} - \lambda\mathbf{I}_p)\mathbf{a_2} = \mathbf{0}$, so λ represents once more an eigenvalue of $\mathbf{\Sigma}$, and $\mathbf{a_2}$ its corresponding eigenvector.

Since $\lambda = \mathbf{a_2'\Sigma a_2}$, so λ has to be as large as possible. Assuming that $\mathbf{\Sigma}$ does not have repeated eigenvalues, λ cannot equal λ_1. If it did, it follows that $\mathbf{a_2} = \mathbf{a_1}$, violating the constraint $\mathbf{a_1'a_2} = 0$. Hence λ is the second largest eigenvalue of $\mathbf{\Sigma}$, and $\mathbf{a_2}$ is the corresponding eigenvector.

As stated above, it can be shown that the for third, fourth, ..., pth PCs, the vectors of coefficients $\mathbf{a_3}, \mathbf{a_4}, \ldots, \mathbf{a_p}$ are the eigenvectors of $\mathbf{\Sigma}$ corresponding to $\lambda_3, \lambda_4, \ldots, \lambda_p$, the third and fourth largest, ..., and the smallest eigenvalue, respectively. Furthermore

$$\text{var}\left[\mathbf{a_k'x}\right] = \lambda_k \quad \text{for } k = 1, 2, \ldots, p\,. \tag{D.11}$$

D.3 Singular Value Decomposition

We shall now discuss a powerful method used to perform numerically a dimensionality reduction of a given parameter space (represented by the dataset mentioned above) accordingly to the PCA here described. Thus, we shall discuss about the singular value decomposition. SVD, which is usually the method of choice for solving most *linear least-squares* problems. We will

outline its relevant theory in this section, in order to furnish a general understanding of the problem, since a depth discussion about the argument goes beyond our scope[4]. SVD methods are based on the following theorem of linear algebra, whose proof will not be provided in this work:

Theorem 2. *Any $m \times n$ matrix $\mathbf{A}$ can be written as the product of an $m \times n$ column-orthogonal matrix $\mathbf{U}$, an $n \times n$ diagonal matrix $\mathbf{M}$ with positive or zero elements (the singular values), and the transpose of a $n \times n$ orthogonal matrix $\mathbf{V}$.*

Depending on the dimensionality of the problem, the SVD of a matrix $\mathbf{A}$, expressed as

$$\mathbf{A} = \mathbf{U}\mathbf{W}\mathbf{V}^T \,, \tag{D.12}$$

can produce an *overdetermined* situation of more equations than unknowns, $m > n$, or conversely an *undetermined* situation of fewer equations than unknowns, $m < n$. Moreover, the matrix $\mathbf{V}$ is orthogonal in the sense that its columns are orthonormal,

$$\sum_{j=0}^{n-1} v_{jk} v_{jl} = \delta_{k,l} \,, \qquad \begin{matrix} 0 \leq k \leq n-1 \\ 0 \leq l \leq n-1 \end{matrix} \tag{D.13}$$

that is, $\mathbf{V}^T\mathbf{V} = \mathbf{1}$. Since $\mathbf{V}$ is square, it is also row-orthonormal, $\mathbf{V}\mathbf{V}^T = \mathbf{1}$. The rows of the matrix $\mathbf{V}^T$ contain the elements of the *right singular vectors*, $\{\mathbf{v}_p\}$, for $p = 0, 1, \ldots, n-1$. When $m \geq n$, the matrix $\mathbf{U}$ is also column orthogonal,

$$\sum_{j=0}^{n-1} u_{jk} u_{jl} = \delta_{k,l} \,, \qquad \begin{matrix} 0 \leq k \leq n-1 \\ 0 \leq l \leq n-1 \end{matrix} \tag{D.14}$$

that is, $\mathbf{U}^T\mathbf{U} = \mathbf{1}$. The columns of $\mathbf{U}$ are called the *left singular vectors*, $\{\mathbf{u}_q\}$, for $q = 0, 1 \ldots, m-1$. In the case of an undetermined problem, $m < n$, then two possibilities rise up:

- The singular values w_j, i.e. the elements of the diagonal matrix $\mathbf{W}$, for $j = m, m+1, \ldots, n-1$, are all zero.
- The corresponding columns of $\mathbf{U}$ are also zero. Eq. (D.14) then only holds for $k, l \leq n-1$.

A remarkable note about the strength of the SVD analysis is that the decomposition $\mathbf{A} = \mathbf{U}\,\mathbf{W}\,\mathbf{V}^T$ can always be done, no matter how singular the matrix is. By convention, the ordering of the singular vectors is determined by high-to-low sorting of singular values, with the highest singular value in the upper left index of the $\mathbf{W}$ matrix. Also note that for a square, symmetric matrix $\mathbf{A}$, singular value decomposition is equivalent to diagonalization, or solution of the eigenvalue problem.

D.4 Calculation of SVD

Therefore, we can now proceed in describing how to calculate SVD and associate the singular vectors to the PCs defined above[5] by adopting some practical case.

Taking the advantage of the notation introduced in the section before, let $\mathbf{A}$ denote a $m \times n$ real-valued data matrix of rank r, where, without loss of generality, $m \geq n$ and $r \leq n$. Moreover,

[4] For further details and a computational treatment of SVD we refer to [259].
[5] More details and applications can be found in [314] and references therein.

we denote as $\mathbf{g}_i$, for $i = 0, 1, \ldots, m-1$ the n-dimensional vector formed by the elements of the ith row of $\mathbf{A}$ and in the same way $\mathbf{a}_j$, for $j = 0, 1, \ldots, n-1$, the m-dimensional vector formed by the elements of the jth column of $\mathbf{A}$. Therefore, the singular value decomposition of $\mathbf{A}$ is represented by Eq. (D.12). Thus, $\mathbf{W} = diag\,(w_0, w_1, \ldots, w_{n-1})$ with $w_k > 0$ for $0 \leq k \leq r-1$ and $w_k = 0$ for $r \leq k \leq n-1$.

One way to calculate the SVD is to first calculate $\mathbf{V}^T$ and $\mathbf{W}$ by diagonalizing the $n \times n$ matrix $\mathbf{A}^T\mathbf{A}$, and then to calculate $\mathbf{U}$. In fact, since $(\mathbf{UWV}^T)^T = \mathbf{VW}^T\mathbf{U}^T$, we have that

$$\mathbf{A}^T\mathbf{A} = \mathbf{VW}^T\mathbf{U}^T\mathbf{UWV}^T = \mathbf{VW}^2\mathbf{V}^T\,, \tag{D.15}$$

where clearly $\mathbf{U}^T\mathbf{U} = \mathbf{1}$ and $\mathbf{W}^T\mathbf{W} = \mathbf{W}^2$. Hence $\mathbf{U}$ can be evaluated as follows

$$\mathbf{U} = \mathbf{AVW}^{-1}\,, \tag{D.16}$$

where the $r, \ldots, n-1$ columns of $\mathbf{V}$ for which $w_k = 0$ are ignored in the matrix multiplication of Eq. (D.15). Choices for the remaining $n-r$ singular vectors in $\mathbf{V}$ or $\mathbf{U}$ may be calculated using the Gram-Schmidt orthogonalization process or some other extension method. On the other hand one can decide to calculate $\mathbf{U}$ and $\mathbf{W}$ by diagonalizing the $m \times m$ matrix $\mathbf{AA}^T$ and then calculate $\mathbf{V}^T$ instead. This yields to

$$\mathbf{AA}^T = \mathbf{UWV}^T(\mathbf{UWV}^T)^T = \mathbf{UWV}^T\mathbf{VW}^T\mathbf{U}^T = \mathbf{UW}^2\mathbf{U}^T\,, \tag{D.17}$$

equivalently to Eq. (D.15). Therefore we can derive $\mathbf{V}^T$ as

$$\mathbf{V}^T = \mathbf{W}^{-1}\mathbf{U}^T\mathbf{A}\,. \tag{D.18}$$

However, a direct relation between PCA and SVD occurs in the case where principal components are calculated from the covariance matrix (see Section D.2) and we investigate it here. If one conditions the data matrix $\mathbf{A}$ by centering each column[6], then the $n \times n$ matrix $\mathbf{A}^T\mathbf{A} = \mathbf{g}_i^T\mathbf{g}_j = \sum_{k=0}^{n-1} g_i^k g_j^k$ is proportional to the $n \times n$ covariance matrix of the column vectors $\mathbf{a}_k$. In fact, since the covariance between two m-dimensional vectors $\mathbf{a}_i$ and $\mathbf{a}_j$ is given as

$$C(\mathbf{a}_i, \mathbf{a}_j) = \frac{1}{m-1}\sum_{k=0}^{m-1}(a_i^k - \langle \mathbf{a}_i \rangle)(a_j^k - \langle \mathbf{a}_j \rangle)\,, \tag{D.19}$$

where a_i^k denotes the kth element of the column vector $\mathbf{a}_i$ and $\langle \mathbf{a}_i \rangle$ its mean value (which is chosen to be zero in this case), we have the $n \times n$ covariance matrix $\mathbf{\Sigma}_n$ as

$$\mathbf{\Sigma}_n = \|C(\mathbf{a}_i, \mathbf{a}_j)\| \qquad \text{for} \quad i, j = 0, 1, \ldots, n-1\,. \tag{D.20}$$

By means of Eq. (D.15) diagonalization of $\mathbf{A}^T\mathbf{A}$ yields $\mathbf{V}^T$, which also yields the principal components of the set of n-dimensional row vectors $\{\mathbf{g}_i\}$. So, the right singular vectors $\{\mathbf{v}_k\}$ are the same as the principal components of $\{\mathbf{g}_i\}$, being the eigenvectors of $\mathbf{A}^T\mathbf{A}$. Moreover, its eigenvalues are given by w_k^2, for $k = 0, 1, \ldots, n-1$, which are proportional to the variances of the principal components. From Eq. (D.12) the matrix $\mathbf{UW}$ then contains the *principal component scores*, which are the values of the new set of coordinates $\{\mathbf{v}_k\}$ with respect to the old set $\{\mathbf{g}_k\}$, in the space of principal components.

If instead each row of $\mathbf{A}$ is centered, then the $m \times m$ matrix $\mathbf{AA}^T = \mathbf{a}_i\mathbf{a}_j^T = \sum_{k=0}^{m-1} a_i^k a_j^k$ is proportional to the $m \times m$ covariance matrix of the row vectors $\mathbf{g}_k$, which is given as

$$\mathbf{\Sigma}_m = \|C(\mathbf{g}_i, \mathbf{g}_j)\| \qquad \text{for} \quad i, j = 0, 1, \ldots, m-1\,, \tag{D.21}$$

[6] Centering a vector means to apply a linear transformation that shifts its elements in order to give a zero mean.

where the covariance between two n-dimensional vectors $\mathbf{g}_i$ and $\mathbf{g}_j$ is expressed as

$$C(\mathbf{g}_i, \mathbf{g}_j) = \frac{1}{n-1} \sum_{k=0}^{n-1} (g_i^k - \langle \mathbf{g}_i \rangle)(g_j^k - \langle \mathbf{g}_j \rangle) \,, \tag{D.22}$$

in which again g_i^k denotes the kth element of the row vector $\mathbf{g}_i$ and $\langle \mathbf{g}_i \rangle$ its mean value (zero). In this case the left singular vectors $\{\mathbf{u}_k\}$ are the same as the principal components of the set of m-dimensional column vectors $\{\mathbf{a}_k\}$, since they are the eigenvectors of $\mathbf{A}\mathbf{A}^T$. The eigenvalues w_k^2 for $k = 0, 1, \ldots, m-1$ are once more proportional to the variances of the principal components. Hence, as Eq. (D.12) suggests the matrix $\mathbf{W}\mathbf{V}^T$ contains again the principal component scores, which are the values of the new set of coordinates $\{\mathbf{u}_k\}$ with respect to the old set $\{\mathbf{a}_k\}$, in the space of principal components.

The application of PCA and SVD to Bayesian modelings consists in having a data matrix formed by the values of the posterior probability density function and hence the set of coordinates we are interested in reducing in number refers to the set of free parameters that formalize the hypotheses of our model. Therefore, the purpose becomes to reduce the dimensionality of the parameter space in order to simplify its exploration and sampling by means of MCMC when high-dimensionality problems have to be addressed. In the case of the Bayesian spot modeling presented in this dissertation it was empirically proved that PCA furnishes a very helpful method for reducing the computational effort required for the exploration of parameter space in the case of a large number of free parameters ($N > 100$).

Acknowledgements

During my three years Ph.D I could improve considerably my knowledge and expertise related to the field of asteroseismology, in which I put my efforts by starting from scratch at the beginning of the year 2010.

Foremost, I would like to express my deep and heartfelt gratitude to my first tutor, and also mentor and friend, Dr. Alfio Bonanno, for his support during my Ph.D and for teaching me a lot about science and real life. Thanks to his motivation, enthusiasm, to his wide knowledge and logical way of thinking, and the very interesting and fruitful conversations we had, he has always provided me with new important idea for research projects and problem solution. He has been guiding me through this path for the entire period of my doctoralship, giving me the good chances to create important work connections and meet a lot of people interested in this exciting field of astrophysics. Without him, all the research done during my Ph.D could not be possible and I could not be what I am now.

Besides my first tutor, I would like to thank my second tutor Dr. Dennis Stello for his exceptional supervision, advice and the very fruitful collaboration during my period of stay at the SIfA, University of Sydney. In fact, despite of the shorter time we worked together, thanks to his great availability, his criticism and methodical way of operating, I could produce a very important part of the research described in this thesis and learn a lot about cluster seismology and paper writing.

I wish to express my warm and sincere thanks to Dr. Hans-Erich Fröhlich for his great kindness, availability and assistance, and for sharing his deep knowledge and expertise about the Bayesian analysis, all along the entire period of my Ph.D. Without his help, encouragement and insightful comments, which gave me a clear answer to all of my questions, many of the results discussed here could not be produced.

I am also very grateful to Prof. Frank Grundahl and Dr. Hans Kjeldsen for their great collaboration, advice and supervision during my stay at the Department of Physics and Astronomy of Aarhus Univeristy, which conducted me through a very interesting and challenging research related to the asteroseismology from ground-based observations.

It is a pleasant task to express my thanks to all those who contributed in many ways to the success of this study and made it an unforgettable experience for me. Among them, I would like to mention Prof. Tim Bedding, for many interesting and productive discussions where he shared his great knowledge and experience related to the asterosemology field, and for his support during my stay at the SIfA; Dr. Antonio Frasca for sharing his expertise related to the photometry and spectroscopy fields and for very useful suggestions; Dr. Othman Benomar for interesting discussions as well; Christian Napoli for his sincere friendship and advice and all those real friends who helped me in enjoying this important period of my life.

Last but certainly not least, I am tremendously grateful to my parents for their continuous support, for always believing in me and encouraging me especially at difficult times because without them all this would not even exist; to my girlfriend Loredana Spampinato, who through her love, patience, support and unwavering belief in me, allowed me to grow up, learning on how getting over the difficulties together, and to complete this dissertation journey successfully.

Funding Supports and other acknowledgements

Funding for the *Kepler* Discovery mission is provided by NASA's Science Mission Directorate. The research leading to thee results discussed in this thesis has received financial support from:

- the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 269194
- the PRIN-INAF 2010 *Asteroseismology: looking inside the stars with space- and ground-based observations*
- the FWO-Flanders under project O6260 - G.0728.11
- the MHD group of the Astrophysics Institute of Postdam

I would like to acknowledge the Sydney Institute for Astronomy (SIfA), School of Physics of the University of Sydney; the Department of Physics and Astronomy of the Aarhus University; the Instituut voor Sterrenkunde at K.U. Leuven, for hosting part of this work.

I also would like to thank the entire Kepler team, without whom the investigation of *Kepler* data would not have been possible.

Bibliography

[1] Aerts C., Christensen-Dalsgaard J., Kurtz D. W., *Asteroseismology*, 2010, Springer Verlag

[2] Akaike, H., I.E.E.E. Trans. Auto. Control AC-19 716–723 (1974)

[3] Anderson, E. R., Duvall, T. L., Jr., & Jefferies, S. M. 1990, ApJ, 364, 699

[4] Andreon, S., & Hurn, M. A. 2012, arXiv:1210.6232

[5] Appourchaux, T., Gizon, L., & Rabello-Soares, M.-C. 1998, A&AS, 132, 107

[6] Appourchaux, T., Chaplin, W. J., García, R. A., et al. 2012, A&A, 543, A54

[7] Appourchaux, T., Benomar, O., Gruberbauer, M., et al. 2012, A&A, 537, A134

[8] Arentoft, T., Kjeldsen, H., Nuspl, J., et al. 1998, A&A, 338, 909

[9] Arentoft, T., Kjeldsen, H., Bedding, T. R., et al. 2008, ApJ, 687, 1180

[10] Arentoft, T., Kjeldsen, H., & Bedding, T. R. 2009, Solar-Stellar Dynamos as Revealed by Helio- and Asteroseismology: GONG 2008/SOHO 21, 416, 347

[11] Baglin, A., Michel, E., Auvergne, M., & The COROT Team. 2006, in ESA Special Publication, Vol. 624, Proceedings of SOHO 18/GONG 2006/HELAS I, Beyond the spherical Sun

[12] Ballot, J., Barban, C., & van't Veer-Menneret, C. 2011, A&A, 531, A124

[13] Balmforth, N. J. 1992, MNRAS, 255, 639

[14] Barnes, J. R., Collier Cameron, A., Unruh, Y. C., Donati, J.-F., & Hussain, G. A. J. 1998, MNRAS, 299, 904

[15] Barnes, J. R., Collier Cameron, A., Donati, J.-F., et al. 2005, MNRAS, 357, L1

[16] Barrado y Navascués, D., Stauffer, J. R., & Jayawardhana, R. 2004, ApJ, 614, 386

[17] Basu, S., Grundahl, F., Stello, D., et al. 2011, ApJ, 729, L10

[18] Batalha, N. M., Rowe, J. F., Bryson, S. T., et al. 2012, arXiv:1202.5852

[19] Baudin, F., Barban, C., Belkacem, K., et al. 2011, A&A, 529, A84

[20] Bayes, T., Phil. Trans. Roy. Soc. 53 370-418, 1763

[21] Beck, P. G., Bedding, T. R., Mosser, B., et al. 2011, Science, 332, 205

[22] Beck, P. G., Montalban, J., Kallinger, T., et al. 2012, Nature, 481, 55

[23] Beck, P. G., De Ridder, J., Aerts, C., et al. 2012, arXiv:1211.0153

[24] Bedding, T. R. 2011, arXiv:1107.1723v1 [astro-ph.SR]

[25] Bedding, T. R. & Kjeldsen, H. 2003, PASA, 20, 203

[26] Bedding, T. R. & Kjeldsen, H. 2006, in ESA Special Publication, Vol. 624, Proceedings of SOHO 18/GONG 2006/HELAS I, Beyond the spherical Sun

[27] Bedding, T. R. & Kjeldsen, H. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 384, 14th Cambridge Workshop on Cool Stars, Stellar Systems, and the Sun, ed. G. van Belle, 21

[28] Bedding, T. R., Kjeldsen, H., Butler, R. P., et al. 2004, ApJ, 614, 380

[29] Bedding, T. R., Kjeldsen, H., Arentoft, T., et al. 2007, ApJ, 663, 1315

[30] Bedding, T. R., Huber, D., Stello, D., et al. 2010, ApJ, 713, L176

[31] Bedding, T. R., Kjeldsen, H., Campante, T. L., et al. 2010, ApJ, 713, 935

[32] Bedding, T. R., Mosser, B., Huber, D., et al. 2011, Nature, 471, 608

[33] Belkacem, K., Samadi, R., & Goupil, M. J. 2011, Journal of Physics Conference Series, 271, 012047

[34] Belkacem, K., Goupil, M. J., Dupret, M. A., et al. 2011, A&A, 530, A142

[35] Belkacem, K., Dupret, M. A., Baudin, F., Appourchaux, T., Marques, J. P., & Samadi, R. 2012, A&A, 540, L7

[36] Belkacem, K. 2012, SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics, 173

[37] Benomar, O., Appourchaux, T., & Baudin, F. 2009, A&A, 506, 15

[38] Bentivegna, E., Bonanno, A., & Reuter, M. 2004, JCAP, 1, 1

[39] Berger, J., & Bernardo, J., *Estimating a product of means: Bayesian analysis with reference priors*, 1989, Journal of the American Statistical Association, 89, 200-207

[40] Bernardo, J., *Reference posterior distributions for Bayesian inference*, 1979, J. R. Stat. Soc. B, 41, 113-147

[41] Bernd A. Berg, *Markov Chain Monte Carlo Simulations and Their Statistical Analysis - With Web-Based Fortran Code*, World Scientific Publishing, 2004

[42] Bildsten, L., Paxton, B., Moore, K., & Macias, P. J. 2012, ApJ, 744, L6

[43] Blobel, V., Least Squares Methods and Function Minimization, in *Formulae and Methods in Experimental Data Evaluation*, Vol. 3, European Physical Society, CERN, 1984

[44] Bolstad W. M., *Introduction to Bayesian Statistics*, 2nd Edition, Wiley & Sons, 2007

[45] Bonaca, A., Tanner, J. D., Basu, S., et al. 2012, ApJ, 755, L12

[46] Bonanno, A., Elstner, D., Rüdiger, G., & Belvedere, G. 2002, A&A, 390, 673

[47] Bonanno, A., Küker, M., & Paternò, L. 2007, A&A, 462, 1031

[48] Bonanno, A., Benatti, S., Claudi, R., et al. 2008, ApJ, 676, 1248

[49] Bonanno, A., Brandenburg, A., Del Sordo, F., & Mitra, D. 2012, Phys. Rev. E, 86, 016313

[50] Borucki, W. J., Koch, D., Basri, G., et al. 2010, Science, 327, 977

[51] Bragaglia, A., Carretta, E., Gratton, R. G., et al. 2001, AJ, 121, 327

[52] Breger, M., Stich, J., Garrido, R., et al. 1993, A&A, 271, 482

[53] Brewer, B. J., Bedding, T. R., Kjeldsen, H., & Stello, D. 2007, ApJ, 654, 551

[54] Brewer, B. J., & Stello, D. 2009, MNRAS, 395, 2226

[55] Brogaard, K., Bruntt, H., Grundahl, F., Clausen, J. V., Frandsen, S., Vandenberg, D. A., & Bedin, L. R. 2011, A&A, 525, A2

[56] Brogaard, K., VandenBerg, D. A., Bruntt, H., et al. 2012, A&A, 543, A106

[57] Broomhall, A.-M., Chaplin, W. J., Elsworth, Y., Appourchaux, T., & New, R. 2010, MNRAS, 406, 767

[58] Brown, T. M., Gilliland, R. L., Noyes, R. W., & Ramsey, L. W. 1991, ApJ, 368, 599

[59] Brown, A., Korhonen, H., Berdyugina, S., et al. 2011, AAS Meeting Abstracts #218, p. 205.02

[60] Brown, T. M., Latham, D. W., Everett, M. E., & Esquerdo, G. A. 2011, AJ, 142, 112

[61] Butler, R. P., Marcy, G. W., Williams, E., et al. 1996, PASP, 108, 500

[62] Butler, R. P., Bedding, T. R., Kjeldsen, H., et al. 2004, ApJ, 600, L75

[63] Campante, T. L., Handberg, R., Mathur, S., et al. 2011, A&A, 534, A6

[64] Carrier, F., Kjeldsen, H., Bedding, T. R., et al. 2007, A&A, 470, 1059

[65] Casagrande, L., Portinari, L., & Flynn, C. 2006, MNRAS, 373, 13

[66] Casagrande, L., Ramírez, I., Meléndez, J., Bessell, M., & Asplund, M. 2010, A&A, 512, A54

[67] Chaplin, W. J., *The Music of the Sun: the Story the Heliosismology*, 2006, Oneworld

[68] Chaplin, W. J., Houdek, G., Appourchaux, T., et al. 2008, A&A, 485, 813

[69] Chaplin, W. J., Houdek, G., Karoff, C., Elsworth, Y., & New, R. 2009, A&A, 500, L21

[70] Chaplin, W. J., Appourchaux, T., Elsworth, Y., et al. 2010, ApJ, 713, L169

[71] Chaplin, W. J., Kjeldsen, H., Christensen-Dalsgaard, J., et al. 2011a, Science, 332, 213

[72] Chaplin, W. J., Kjeldsen, H., Bedding, T. R., et al. 2011b, ApJ, 732, 54

[73] Chaplin, W. J., Bedding, T. R., Bonanno, A., et al. 2011c, ApJ, 732, L5

[74] Christensen-Dalsgaard, J., & Frandsen, S. 1983, SoPh, 82, 469

[75] Christensen-Dalsgaard, J. 1984, in Space Research in Stellar Activity and Variability, ed. A. Mangeney & F. Praderie, 11

[76] Christensen-Dalsgaard, J., *Lecture Notes on Stellar Oscillations. Fifth Edition*, 2003, University of Aarhus

[77] Christensen-Dalsgaard, J. 2004, Sol. Phys., 220, 137

[78] Christensen-Dalsgaard, J. 2011, arXiv: 1106.5946v1 [astro-ph.SR]

[79] Christensen-Dalsgaard, J., Gough, D. O. & Toomre, J., 1985, Science, 229, 923 – 931

[80] Christensen-Dalsgaard, J. & Berthomieu, G., 1991, In Solar interior and atmosphere, p. 401 – 478, eds Cox, A. N., Livingston, W. C. & Matthews, M., Space Science Series, University of Arizona Press.

[81] Christensen-Dalsgaard, J., & Perez Hernandez, F. 1992, MNRAS, 257, 62

[82] Claret, A., & Bloemen, S. 2011, A&A, 529, 75

[83] Claudi, R., Benatti, S., Bonanno, A., et al. 2009, Communications in Asteroseismology, 159, 21

[84] Collier Cameron, A. 1995, MNRAS, 275, 534

[85] Corsaro, E. *Analisi dell'abbondanza di Litio in stelle di tipo TTauri*, University of Catania, 2009

[86] Corsaro, E., Grundahl, F., Leccia, S., et al. 2012a, A&A, 537, A9

[87] Corsaro, E., Stello, D., Huber, D., et al. 2012b, ApJ, 757, 190

[88] Corsaro, E., Fröhlich, H.-E., Bonanno, A., et al. 2012c, MNRAS (submitted), arXiv:1212.1156v1 [astro-ph.SR]

[89] Cox, R., J. A. Stat. Assoc. 14 1-13, 1946

[90] Cox, J. P. & Smith, R. C. 1981, The Observatory, 101, 87

[91] Cox, A. N. & Pilachowski, C. A. 2000, Physics Today, 53, 100000

[92] Covas, E., Moss, D., & Tavakol, R. 2005, A&A, 429, 657

[93] Croll, B. 2006, PASP, 118, 1351

[94] Croll, B., Walker, G. A. H., Kuschnig, R., et al. 2006, ApJ, 648, 607

[95] D'Agostini, G. 2005, arXiv:physics/0511182

[96] De Ridder, J., et al. 2009, Nature, 459, 398

[97] Deubner, F.-L. & Gough, D. O., 1984, Ann. Rev. Astron. Astrophys. 22, 593 – 619

[98] Diaconis, P., Holmes, S., and Montgomery, R., SIAM Review 49 211-235, 2007

[99] Dikpati, M., & Gilman, P. A. 2001, ApJ, 559, 428

[100] Donati, J.-F., Collier Cameron, A., & Petit, P. 2003, MNRAS, 345, 1187

[101] Dorren, J. D. 1987, ApJ, 320, 756

[102] Duncan, D. K., Vaughan, A. H., Wilson, O. C., et al. 1991, ApJS, 76, 383

[103] Dupret, M.-A., et al. 2009, A&A, 506, 57

[104] Dziembowski, W. A., & Soszyński, I. 2010, A&A, 524, A88

[105] Eddington A. S., *The internal constitution of the stars*, 1926, Cambridge University Press, Cambridge

[106] Eggenberger, P., Carrier, F., Bouchy, F., & Blecha, A. 2004, A&A, 422, 247

[107] Elsworth, Y. P. & Thompson, M. J. 2004, Astronomy and Geophysics, 45, 050000

[108] Eyer, L., & Bartholdi, P. 1999, A&AS, 135, 1

[109] Eyer, L., & Mowlavi, N. 2008, Journal of Physics Conference Series, 118, 012010

[110] Endl, M., Kürster, M., & Els, S. 2000, A&A, 362, 585

[111] Flower, P. J. 1996, ApJ, 469, 355

[112] Frandsen, S., Jones, A., Kjeldsen, H., et al. 1995, A&A, 301, 123

[113] Frandsen, S., Carrier, F., Aerts, C., et al. 2002, A&A, 394, L5

[114] Frasca, A., Guillout, P., Marilli, E., et al. 2006, A&A, 454, 301

[115] Frasca, A., Fröhlich, H.-E., Bonanno, A., et al. 2011, A&A, 532, A81

[116] Fröhlich, H.-E. 2007, Astron. Nachr., 328, 1037

[117] Fröhlich, H.-E., Küker, M., Hatzes, A. P., & Strassmeier, K. G. 2009, A&A, 506, 263

[118] Fröhlich, H.-E., Frasca, A., Catanzaro, G., et al. 2012, A&A, 543, A146

[119] Fröhlich, H.-E. 2012, Astron. Nachr., in prep.

[120] Fuhrmann, K. 1998, A&A, 338, 161

[121] Fuhrmann, K. 2004, Astronomische Nachrichten, 325, 3

[122] García, R. A., Hekker, S., Stello, D., et al. 2011, MNRAS, 414, L6

[123] Gauss, C. F. 1866, Carl Friedrich Gauss Werke, 3, 265

[124] Ghosh, J., Delampady, M., and Samanta, T., *An Introduction to Bayesian Analysis*, Springer, New York, 2006

[125] Gilliland, R. L., Morris, S. L., Weymann, R. J., Ebbets, D. C., & Lindler, D. J. 1992, PASP, 104, 367

[126] Gilliland, R. L. 2008, AJ, 136, 566

[127] Gilliland, R. L., Jenkins, J. M., Borucki, W. J., et al. 2010a, ApJ, 713, L160

[128] Gilliland, R. L., Brown, T. M., Christensen-Dalsgaard, J., et al. 2010b, PASP, 122, 131

[129] Girardi, L., Bressan, A., Bertelli, G., & Chiosi, C. 2000, A&AS, 141, 371

[130] Girardi, L., Groenewegen, M. A. T., Hatziminaoglou, E., & da Costa, L. 2005, A&A, 436, 895

[131] Glushkova, E. V., Batyrshinova, V. M., Ibragimov, M. A. 1999, AstL, 25, 86

[132] Gott, J. R., III, Vogeley, M. S., Podariu, S., & Ratra, B. 2001, ApJ, 549, 1

[133] Gough, D. O. 1986, in Hydrodynamic and Magnetodynamic Problems in the Sun and Stars, ed. Y. Osaki, 117

[134] Gough, D. 1987, Nature, 326, 257

[135] Gough, D. O. & Toomre, J., 1991, Ann. Rev. Astron. Astrophys., 29, 627 – 685

[136] Gough, D. O., Leibacher, J. W., Scherrer, P. H., & Toomre, J. 1996, Science, 272, 1281

[137] Gratton, R. G., Bonanno, G., Bruno, P., et al. 2001, Experimental Astronomy, 12, 107

[138] Gray F. David, *The Observation and Analysis of Stellar Photospheres, Third Edition*, 2005, Cambridge University Press

[139] Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, AJ, 132, 161

[140] Grec, G., Fossat, E., & Pomerantz, M. A. 1983, Solar Physics, 82, 55

[141] Gruberbauer, M., Guenther, D. B., & Kallinger, T. 2012, ApJ, 749, 109

[142] Grundahl, F., Christensen-Dalsgaard, J., Kjeldsen, H., et al. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 416, Astronomical Society of the Pacific Conference Series, ed. M. Dikpati, T. Arentoft, I. González Hernández, C. Lindsey, & F. Hill, 579

[143] Gut, Allan, An intermediate Course in Probability, 2009, Chapter 5.

[144] Handberg, R., & Campante, T. L. 2011, A&A, 527, A56

[145] Hanson, K., M., *Markov Chain Monte Carlo posterior sampling with the Hamiltonian method*, in M. Sonka and K. M. Hanson (Eds) Medical Imaging: Image Processing Vol. 4322, Proc. SPIE, pp. 456-467

[146] Harvey, J. 1985, Future Missions in Solar, Heliospheric & Space Plasma Physics, 235, 199

[147] Hastings, W. K., 1970, Biometrika, 57, 97

[148] Hatzes, A.P. 1995, ApJ, 451, 784

[149] Hekker, S., Broomhall, A.-M., Chaplin, W. J., et al. 2010, MNRAS, 402, 2049

[150] Hekker, S., Basu, S., Stello, D., et al. 2011a, A&A, 530, A100

[151] Hekker, S., Gilliland, R. L., Elsworth, Y., et al. 2011b, MNRAS, 414, 2594

[152] Hoaglin, D., C., Mosteller, F., Tukey, J., W., *Understanding robust and exploratory data analysis*, 1983, John Wiley & Sons, NY

[153] Högbom, J. A. 1974, A&AS, 15, 417

[154] Hole, K. T., Geller, A. M., Mathieu, R. D., Platais, I., Meibom, S., & Latham, D. W. 2009, AJ, 138, 159

[155] Holman, M. J., & Murray, N. W. 2005, Science, 307, 1288

[156] Hotelling, H., *Analysis of a Complex of Statistical Variables with Principal Components*, 1933, Journal of Educational Psychology

[157] Hotta, H., & Yokoyama, T. 2011, ApJ, 740, 12

[158] Houdek, G., Balmforth, N. J., Christensen-Dalsgaard, J., & Gough, D. O. 1999, A&A, 351, 582

[159] Houdek, G., & Gough, D. O. 2002, MNRAS, 336, L65

[160] Houdek, G. 2006, Proceedings of SOHO 18/GONG 2006/HELAS I, Beyond the spherical Sun, 624

[161] Hrudková, M. 2006, *The Accurate Barycentric Corrections for the Detection of Extrasolar Planets*, WDS'06 Proceedings of Contributed Papers: Part III – Physics (eds. J. Šafránková and J. Pavlů), Prague, Matfyzpress, p. 18–23

[162] Hrudková, M., *Planets by other suns*, PhD Thesis, Charles University in Prague, 2006

[163] Hubble, E., & Humason, M. L. 1931, ApJ, 74, 43

[164] Huber, D., Stello, D., Bedding, T. R., Chaplin, W. J., Arentoft, T., Quirion, P.-O., & Kjeldsen, H. 2009, Communications in Asteroseismology, 160, 74

[165] Huber, D., Bedding, T. R., Stello, D., et al. 2010, ApJ, 723, 1607

[166] Huber, D., Bedding, T. R., Arentoft, T., et al. 2011, ApJ, 731, 94

[167] Huber, D., Bedding, T. R., Stello, D., et al. 2011, ApJ, 743, 143

[168] Huber, K. F., Wolter, U., Czesla, S., et al. 2009, A&A, 501, 715

[169] Järvinen, S. P., Korhonen, H., Berdyugina, S. V., et al. 2008, A&A, 488, 1047

[170] Jaynes, E., T., Probability Theory. The logic of science, Cambridge University Press, Cambridge, UK, 2003

[171] Jeffers, S. D., & Donati, J.-F. 2008, MNRAS, 390, 635

[172] Jeffreys, H., *Theory of probability*, 3rd Ed, Oxford Classics series (reprinted 1998), Oxford University Press, Oxford, UK, 1961

[173] Jenkins, C. R., & Peacock, J. A. 2011, MNRAS, 413, 2895

[174] Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010a, ApJ, 713, L120

[175] Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010b, ApJ, 713, L87

[176] Jolliffe I.T., *Principal Component Analysis, Series: Springer Series in Statistics*, 2nd ed., Springer, NY, 2002, 487 p. 28 ills. ISBN 978-0-387-95442-4

[177] Kallinger, T., Mosser, B., Hekker, S., et al. 2010a, A&A, 522, A1

[178] Kallinger, T., Weiss, W. W., Barban, C., et al. 2010, A&A, 509, A77

[179] Kallinger, T., Hekker, S., Mosser, B., et al. 2012, A&A, 541, A51

[180] Kartunnen H. et al., *Foundamental Astronomy, Fifth Edition*, Springer Berlin Heidelberg New York, 2007

[181] Karoff, C., *Observational asteroseismology*, PhD Thesis, University of Aarhus, 2008

[182] Karoff, C., Metcalfe, T. S., Chaplin, W. J., et al. 2009, MNRAS, 399, 914

[183] Karoff, C., Campante, T. L., & Chaplin, W. J. 2010, arXiv:1003.4167

[184] Kass, R. E., & Wasserman, L. 1996, J. Am. Stat. Association, 91, 1343

[185] Kjeldsen, H. 2003, Ap&SS, 284, 1

[186] Kjeldsen, H. & Frandsen, S. 1992, PASP, 104, 413

[187] Kjeldsen, H. & Bedding, T. R. 1995, A&A, 293, 87

[188] Kjeldsen, H., & Bedding, T. R. 2011, A&A, 529, L8

[189] Kjeldsen, H., Bedding, T. R., Viskum, M., & Frandsen, S. 1995, AJ, 109, 1313

[190] Kjeldsen, H., Bedding, T. R., Butler, R. P., et al. 2005, ApJ, 635, 1281

[191] Kjeldsen, H., Bedding, T. R., Arentoft, T., et al. 2008, ApJ, 682, 1370

[192] Kiss, L. L., Szatmáry, K., Cadmus, R. R., Jr., & Mattei, J. A. 1999, A&A, 346, 542

[193] Koch, D. G., Borucki, W. J., Basri, G., et al. 2010, ApJ, 713, L79

[194] Korhonen, H., & Elstner, D. 2011, A&A, 532, 106

[195] Kővári, Zs., Frasca, A., Biazzo, K., et al. 2011, Proceedings IAU Symposium No. 273 "Physics of Sun and Star Spots", D. P. Choudhary & K. G. Strassmeier, eds., Cambridge Univ. Press, p. 121

[196] Küker, M., Rüdiger, G., & Kitchatinov, L. L. 2011, A&A, 530, A48

[197] Kullback, S. and R. A. Leibler, R., A., Ann. Math. Stat. 22 79–86 (1951)

[198] Kunz, M., Trotta, R., & Parkinson, D. R. 2006, Phys. Rev. D, 74, 023503

[199] Lanza, A. F., Rodonò, M., & Zappala, R. A. 1994, A&A, 290, 861

[200] Lanza, A. F. 2006, MNRAS, 373, 819

[201] Lanza, A. F., Bonomo, A. S., Pagano, I., et al. 2011, A&A, 525, A14

[202] Leccia, S., *Studio sismologico della stella di tipo solare Procione A*, PhD Thesis, University of Catania, 2004

[203] Leccia, S., Kjeldsen, H., Bonanno, A., et al. 2007, A&A, 464, 1059

[204] Leibacher, J. W., Noyes, R. W., Toomre, J. & Ulrich, R. K., 1985, Scientific American, 253, (September) p. 34 – 43 (US p. 48 – 57)

[205] Ledoux P., 1951, ApJ, 114, 373 – 384

[206] Libbrecht, K. G., 1988, Space Sci. Rev., 47, 275 – 301

[207] Lindborg, M., Korpi, M. J., Hackman, T., et al. 2011, A&A, 526, A44

[208] Lomb, N. R. 1976, Ap&SS, 39, 447

[209] Luck, R. E. & Heiter, U. 2005, AJ, 129, 1063

[210] Lüftinger, T., Fröhlich, H.-E., Weiss, W. W., et al. 2010, A&A, 509, A43

[211] Marcy, G. W., & Butler, R. P. 1992, PASP, 104, 270

[212] Marigo, P., Girardi, L., Bressan, A., et al. 2008, A&A, 482, 883

[213] Marsden, S. C., Donati, J.-F., Semel, M., Petit, P., & Carter, B. D. 2006, MNRAS, 370, 468

[214] Marsden, S. C., Jardine, M. M., Ramírez Vélez, J. C., et al. 2011, MNRAS, 413, 1939

[215] Mathur, S., García, R. A., Régulo, C., et al. 2010, A&A, 511, A46

[216] Mathur, S., Handberg, R., Campante, T. L., et al. 2011, ApJ, 733, 95

[217] Mathur, S., Hekker, S.,Trampedach, R., et al. 2011, ApJ, 741, 119

[218] Mathur, S., Campante, T. L., Handberg, R., et al. 2012, Progress in Solar/Stellar Physics with Helio- and Asteroseismology, 462, 180

[219] Mathur, S., Hekker, S., Trampedach, R., et al. 2012, Progress in Solar/Stellar Physics with Helio- and Asteroseismology, 462, 375

[220] Mathur, S. 2013, Advances in Solid State Physics, 31, 237

[221] Matthews, J. M. 2007, Communications in Asteroseismology, 150, 333

[222] Mazumdar, A. 2005, A&A, 441, 1079

[223] Meibom, S., Barnes, S. A., Latham, D. W., et al. 2011, ApJ, 733, L9

[224] Messina, S. & Guinan, E. F. 2003, A&A, 409, 1017

[225] Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H. & Teller E., 1953, J. Chem. Phys., 21, 1087

[226] Metcalfe, T. S., Monteiro, M. J. P. F. G., Thompson, M. J., et al. 2010, Astronomische Nachrichten, 331, 977

[227] Metcalfe, T. S., Chaplin, W. J., Appourchaux, T., et al. 2012, ApJ, 748, L10

[228] Meyn, S. & Richard T., *Markov Chains and Stochastic Stability*, Springer Verlag, 1993

[229] Michel, E., Baglin, A., Auvergne, M., et al. 2008, Science, 322, 558

[230] Michel, E., Samadi, R., Baudin, F., et al. 2009, A&A, 495, 979

[231] Miglio, A., Montalbán, J., Baudin, F., et al. 2009, A&A, 503, L21

[232] Miglio, A., et al. 2012, MNRAS, 419, 2077

[233] Miglio, A., Montalbán, J., & Noels, A. 2012, *Red Giants as Probes of the Structure and Evolution of the Milky Way*, Springer-Verlag

[234] Miglio, A., Chiappini, C., Morel, T., et al. 2012, arXiv:1211.0146

[235] Monet, D. G., Bryson, S. T., Dotson, J. L., et al. 2010, Bulletin of the American Astronomical Society, 42, #305.05

[236] Montalbán, J., Miglio, A., Noels, A., Scuflaire, R., & Ventura, P. 2010, ApJ, 721, L182

[237] Montgomery, M. H. & O'Donoghue, D. 1999, Delta Scuti Star Newsletter, 13, 28

[238] Mosser, B. 2012, arXiv:1210.7301

[239] Mosser, B., Baudin, F., Lanza, A. F., et al. 2009, A&A, 506, 245

[240] Mosser, B., & Appourchaux, T. 2009, A&A, 508, 877

[241] Mosser, B., Belkacem, K., Goupil, M.-J., et al. 2010, A&A, 517, A22

[242] Mosser, B., Barban, C., Montalbán, J., et al. 2011a, A&A, 532, A86

[243] Mosser, B., Belkacem, K., Goupil, M. J., et al. 2011b, A&A, 525, L9

[244] Mosser, B., Elsworth, Y., Hekker, S., et al. 2012a, A&A, 537, A30

[245] Mosser, B., Goupil, M. J., Belkacem, K., et al. 2012b, A&A, 540, A143

[246] Moya, A., García Hernández, A., Suárez, J. -., Rodríguez-López, C., & Garrido, R. 2010, arXiv:1004.0100

[247] Neal, R., M., *Probabilistic Inference Using Markov Chain Monte Carlo Methods*, Department of Computer Science, University of Toronto, 1993

[248] Nelson, N. J., Brown, B. P., Brun, S., Miesch, M. S., & Toomre, J. 2011, 217th AAS Meeting, #155.12, Bulletin of the American Astronomical Society, Vol. 43

[249] Nordgren, T. E., Germain, M. E., Benson, J. A., et al. 1999, AJ, 118, 3032

[250] Oja, T. 1986, A&AS, 65, 405

[251] Parker, E. N. 1955, ApJ, 122, 293

[252] Pearson, E. S. & Hartley, H. O., Biometrika Tables for Statisticians, Vol. 1, Table 8, pages 137 and 138, 1966

[253] Pearson, K., On Lines and Planes of Closest Fit to Systems of Points in Space, Philosophical Magazine 2 (6): 559-572, 1901

[254] Perryman, M. A. C., Lindegren, L., Kovalevsky, J., et al. 1997, A&A, 323, L49

[255] Pigulski, A., Pojmański, G., Pilecki, B., & Szczygieł, D. M. 2009, Acta Astronomica, 59, 33

[256] Pinsonneault, M. H., An, D., Molenda-Żakowicz, J., et al. 2012a, ApJS, 199, 30

[257] Pinsonneault, M. H., An, D., Molenda-Zakowicz, J., et al. 2012b, VizieR Online Data Catalog, 219, 90030

[258] Piskunov, N. E., & Valenti, J. A. 2002, A&A, 385, 1095

[259] Press, W. H. et al, *Numerical Recipes - The Art of Scientific Computing*, 3rd Ed., 2007, Cambridge University Press, Chapters 2 and 15

[260] Quirion, P., Christensen-Dalsgaard, J., & Arentoft, T. 2010, ApJ, 725, 2176

[261] Radford M. Neal, *Probabilist Inference using Markov Chain Monte Carlo methods*, Department of Computer Science, University of Toronto

[262] Ramírez, I., & Meléndez, J. 2005, ApJ, 626, 446

[263] Reiners, A. 2006, A&A, 446, 267

[264] Reiners, A., & Schmitt, J. H. M. M. 2003, A&A, 398, 647

[265] Richichi, A., Percheron, I., & Khristoforova, M. 2005, A&A, 431, 773

[266] Roberts, D. H., Lehar, J., & Dreher, J. W. 1987, AJ, 93, 968

[267] Rodonò, M., Messina, S., Lanza, A. F., Cutispoto, G., & Teriaca, L. 2000, A&A, 358, 624

[268] Rosvick, J. M., & Vandenberg, D. A. 1998, AJ, 115, 1516

[269] Samadi, R., Georgobiani, D., Trampedach, R., et al. 2007, A&A, 463, 297

[270] Samadi, R., Ludwig, H.-G., Belkacem, K., et al. 2010, A&A, 509, A16

[271] Samadi, R., Belkacem, K., Dupret, M.-A., et al. 2012, A&A, 543, A120

[272] Scargle, J. D. 1982, ApJ, 263, 835

[273] Schwarz, G., Ann. Statist. 6 461–464 (1978)

[274] Siess, L., Dufour, E., & Forestini, M. 2000, A&A, 358, 593

[275] Silva Aguirre, V., Chaplin, W. J., Ballot, J., et al. 2011, ApJ, 740, L2

[276] Sing, D. K. 2010, A&A, 510, A21

[277] Skrutskie, M. F., Cutri, R. M., Stiening, R., et al. 2006, AJ, 131, 1163

[278] Smith, A. and Roberts, G., J. R. Statist. Soc D 55 3-23, 1993

[279] Soubiran, C., Le Campion, J.-F., Cayrel de Strobel, G., & Caillo, A. 2010, A&A, 515, A111

[280] Soderblom, D. R., Stauffer, J. R., Hudon, J. D., & Jones, B. F. 1993b, ApJS, 85, 315

[281] Spiegel, M. R., *Theory and Problems of Probability and Statistics*, 23rd Edition, Schaum's Outline Series, McGraw-Hill, USA, 1998

[282] Spiegelhalter, D., Best, N., G., Carlin, B., P., et al., J. Royal. Stat. Soc. B 64 583–639 (2002)

[283] Steenbeck, M., Krause, F., & Rädler, K.-H. 1966, Z. Naturforsch. 21a, 369

[284] Stello, D., *Time-series analysis of oscillating red giant stars*, PhD Thesis, University of Sydney, 2006

[285] Stello, D., Kjeldsen, H., Bedding, T. R., & Buzasi, D. 2006, A&A, 448, 709

[286] Stello, D., Chaplin, W. J., Basu, S., Elsworth, Y., & Bedding, T. R. 2009a, MNRAS, 400, L80

[287] Stello, D., Chaplin, W. J., Bruntt, H., et al. 2009b, ApJ, 700, 1589

[288] Stello, D., Basu, S., Bruntt, H., et al. 2010, ApJ, 713, L182

[289] Stello, D., Basu, S., Bedding, T. R., et al. 2010, Astronomische Nachrichten, 331, 985

[290] Stello, D., Basu, S., Bruntt, H., et al. 2010b, ApJL, 713, L182

[291] Stello, D., Huber, D., Kallinger, T., et al. 2011, ApJ, 737, L10

[292] Stello, D., Meibom, S., Gilliland, R. L., et al. 2011, ApJ, 739, 13

[293] Stello, D. 2012, Progress in Solar/Stellar Physics with Helio- and Asteroseismology, 462, 200

[294] Stetson, P. B., Bruntt, H., & Grundahl, F. 2003, PASP, 115, 413

[295] Still, M. D. 2012, American Astronomical Society Meeting Abstracts, 220, #419.01

[296] Suhov, Y. and Kelbert, M., *Probability and Statistics by Example: II Markov Chain: A Primer in Random Processes and their Applications*, Cambridge University Press, 2008

[297] Syversveen, A. R., *Noninformative Bayesian priors. Interpretation and problems with construction and applications.* Preprint Statistics No. 3, 1998, Department of Mathematical Sciences, NTNU, Trondheim

[298] Szatmary, K., Vinko, J., & Gal, J. 1994, A&AS, 108, 377

[299] Tassoul, M. 1980, ApJS, 43, 469

[300] Tassoul, M. 1990, ApJ, 358, 313

[301] Trampedach, R., Christensen-Dalsgaard, J., Nordlund, A., & Stein, R. F. 1998, The First MONS Workshop: Science with a Small Space Telescope, 59

[302] Toutain, T., & Appourchaux, T. 1994, A&A, 289, 649

[303] Trotta, R. 2008, Contemporary Physics, 49, 71

[304] Ulrich, R. K. 1986, ApJ, 306, L37

[305] Uytterhoeven, K., Moya, A., Grigahcène, A., et al. 2011, A&A, 534, A125

[306] Valenti, J. A., Butler, R. P., & Marcy, G. W. 1995, PASP, 107, 966

[307] Valenti, J. A. & Fischer, D. A. 2005, ApJS, 159, 141

[308] Vandakurov, Y. V. 1968, Soviet Ast., 11, 630

[309] van Leeuwen, F. 2007, A&A, 474, 653

[310] Verner, G. A., Elsworth, Y., Chaplin, W. J., et al. 2011, MNRAS, 415, 3539

[311] Voges, W., Aschenbach, B., Boller, T., et al. 1999, A&A, 349, 389

[312] Voges, W., Aschenbach, B., Boller, T., et al. 2000, IAU Circ., 7432

[313] Waite, I. A., Marsden, S. C., Carter, B. D., et al. 2011, MNRAS, 413, 1949

[314] Wall, M., E., et al. *Singular value decomposition and principal component analysis - A Practical Approach to Microarray Data Analysis*, Kluwer: Norwell, MA, 2003. pp. 91-109. LANL LA-UR-02-4001

[315] White, T. R., Brewer, B. J., Bedding, T. R., Stello, D., & Kjeldsen, H. 2010, Communications in Asteroseismology, 161, 39

[316] White, T. R., Bedding, T. R., Stello, D., Christensen-Dalsgaard, J., Huber, D., & Kjeldsen, H. 2011a, ApJ, 743, 161

[317] White, T. R., Bedding, T. R., Stello, D., et al. 2011, ApJ, 742, L3

[318] White, T. R., Bedding, T. R., Gruberbauer, M., et al. 2012, ApJ, 751, L36

[319] Woodard M., PhD Thesis, University of California, San Diego, 1984

Declaration

I herewith declare that I have produced this work without the prohibited assistance of third parties and without making use of aids other than those specified; notions taken over directly or indirectly from other sources have been identified as such. This work has not previously been presented in identical or similar form to any other italian or foreign examination board.

The PhD Thesis work was conducted from 1st January 2010 to 6th December 2012 under the tutoring of Dr. Alfio Bonanno at INAF OACT (Astrophysical Observatory of Catania, University of Catania, Catania, Italy) and of Dr. Dennis Stello at SIfA (Sydney Institute for Astronomy, University of Sydney, Sydney, Australia).

Catania,
6th December 2012