

## Measuring active latitudes of Sun-like stars using a-coefficients decomposition on asteroseismic pulsations

OTHMAN BENOMAR,<sup>1,2,3</sup> MASAO TAKATA,<sup>4</sup> MICHAEL BAZOT,<sup>5</sup> TAKASHI SEKII,<sup>1,2</sup> LAURENT GIZON,<sup>6,7,3</sup> AND YUTING LU<sup>8</sup>

<sup>1</sup>*Department of Astronomical Science, School of Physical Sciences, SOKENDAI, 2-21-1 Osawa, Mitaka, Tokyo 181-8588*

<sup>2</sup>*Solar Science Observatory, National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*

<sup>3</sup>*New York University Abu Dhabi, Center for Space Science, PO Box 129188, UAE*

<sup>4</sup>*Department of Astronomy, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan*

<sup>5</sup>*Heidelberg Institut für Theoretische Studien, Schloss-Wolfsbrunnenweg 35, 69118 Heidelberg, Germany*

<sup>6</sup>*Max-Planck-Institut für Sonnensystemforschung, 37077, Göttingen, Germany*

<sup>7</sup>*Institut für Astrophysik, Georg-August-Universität Göttingen, 37077, Göttingen, Germany*

<sup>8</sup>*Department of Physics, The University of Tokyo, Tokyo 113-0033, Japan*

(Received August 28, 2019; Revised August 28, 2019; Accepted March 28, 2023)

### ABSTRACT

We introduce a framework to measure the asphericity of Sun-like stars using  $a_1$ ,  $a_2$  and  $a_4$  coefficients in order to determine their active latitudes. Systematic errors on the coefficients are evaluated in function of key seismic parameters (inclination, average rotation, height-to-noise ratio). The measured a-coefficients are used in pair with a geometrical model assuming a single active band to determine the intensity and location of the stellar active regions. Using the solar SoHo/VIRGO/SPM data, the method is shown to detect the mean active latitude and its intensity changes between 1999-2002 (maximum of activity) and 2006-2009 (minimum of activity). The two solar-analog 16 Cyg A and B are also studied using Kepler observation. An equatorial band of activity of intensity commensurable with the Sun is detected. However, 16 Cyg B shows a difficult-to-explain bi-modality on  $a_4$ . It is suggested that this could be the manifestation of transition between a quiet and an active phase of activity. Invalidating this hypothesis is difficult and might require new observations.

**Keywords:** stars:activity, stars:rotation, stars:solar-type, asteroseismology

### 1. INTRODUCTION

The observations from the space-borne instruments MOST (Walker et al. 2003), CoRoT (Baglin et al. 2006), *Kepler* (Borucki et al. 2010) and TESS (Ricker et al. 2014) have been important in advancing our knowledge of **stellar interiors**. This is particularly true for *Kepler*, which could observe continuously tens of thousand of stars for nearly **four** continuous years, enabling asteroseismic measurements that **almost rival disk-integrated** helioseismic measurements from a decade ago. Those precise measurements have for example enabled us to better understand stellar rotation and its impact on **stellar evolution**. The asteroseismology of Sun-like stars is **based on the study of the pressure modes that are excited by turbulent convection in the outer layers of these stars**. The **acoustic modes may** reach deep into the core or may **be localised** close to the stellar surface, giving **access to the internal structure and dynamics of the stars**. Rotation has a critical impact on the stellar structure and evolution as it induces a material mixing process. **Rotation is also** an essential ingredient of the dynamo-effect (Thompson et al. 2003) and can lead to a distortion of the star's shape due to the centrifugal force (Chandrasekhar 1969). For fast rotators, the **rotational flattening must be taken into account**. These stars show a variety of pulsation modes that are not **present** in slower rotators (e.g. Lignières et al. 2006).

Despite its slow rotation rate, **the solar asphericity can be measured by helioseismology**. While the Sun is seen as **oblate by acoustic modes during the quiet phases of its activity cycle, the modes feel a more complicated shape during activity maxima**. During solar maxima, **the frequencies of acoustic modes increase slightly due to the presence of active regions**.

The perturbation occurs near the surface and is stronger for modes that sense the active latitudes (below  $\simeq 30^\circ$  for the Sun). Physically the magnetic perturbations consist of several components that are not easy to disentangle (incl. stratification and wave speed perturbations (Libbrecht & Woodard 1990; Antia et al. 2000; Dziembowski et al. 2000). The magnetic perturbations affect the acoustic modes near the surface over only a few hundred kilometers, that is of order  $10^{-5}R_\odot$ , but is significant enough to be measured.

The origin of stellar activity is not well understood as it depends the complex interplay between rotation, convection and the magnetic field (Brun & Browning 2017). Stellar magnetic cycles are observed in most cool stars (Simon et al. 2002) over a large range of the electromagnetic spectrum and evidence of activity cycles is observed in X-ray (eg. Catura et al. 1975), radio waves (White 1999; White et al. 2017), chromospheric emission lines (Vaughan 1983; Oláh et al. 2009, 2016) or also through luminosity variation due to surface magnetic activity in the visible (Hartmann & Rosner 1979; Silva-Valio et al. 2010; Ceillier et al. 2017). In Sun-like stars, the level of activity is often observed to be cyclic, with activity periods ranging from a few years to decades. Although there may be a relationship between the stellar age, the rotation period and the level of activity (van Saders et al. 2016) the underlying mechanisms at play are not understood.

Since the advent of space-borne photometry and the observation of Sun-like pulsations by CoRoT, it became evident that helioseismic methods used for the Sun may be applied to asteroseismic observations. Regarding rotation, that resulted in robust estimates for the average rotational splitting bringing evidences of nearly-uniform internal rotation in main sequence stars (Benomar et al. 2015; Nielsen et al. 2017) and confirming that a near-constant rotational splitting is expected to be a good approximation (Lund et al. 2014a,b).. For the best Kepler observations, asteroseismology showed evidence of latitudinal differential rotation for main sequence stars (Benomar et al. 2018a) and of the radial differential rotation in RGB (e.g. Deheuvels et al. 2012, 2014).

This paper aims at providing a framework to study the stellar activity and its latitude by analysing its effect on pulsation frequencies. The proposed method involves the use of the *a*-coefficient decomposition (Schou et al. 1994) on the stellar power spectrum, conveniently separating the perturbations caused by rotation and asphericity. The method is tested on the Sun-as-a-star and on the solar-analogs 16 Cyg A and B. These are the two brightest stars in the initial Kepler observation field. Successful attempts to measure the asphericity using *a*-coefficients on other stars than the Sun were made. Using the  $a_2$  coefficient, Gizon et al. (2016) found that KIC 11145123, a  $\gamma$  Doradus– $\delta$  Scuti star has an oblate shape, with an asphericity of  $\Delta R/R = (1.8 \pm 0.6) \times 10^{-6}$  (or  $\Delta R \simeq 3 \pm 1$  km), smaller than expected from rotational oblateness alone. Bazot et al. (2019) using a parametric model, determined the asphericity of 16 Cyg A and B and found that 16 Cyg A is more likely prolate. They suggest that these stars may have a toroidal equatorial magnetic field.

Using simulated power spectra, Gizon (2002) proposed to use the  $a_2$  coefficient in order to constrain the latitude of activity. Here it is proposed to use both  $a_2$  and  $a_4$  coefficients. We start in Section 2 by introducing the effects of rotation on pulsation frequencies and discuss the effect of the centrifugal force and of the activity on the mode cavities. Section 3 presents the assumptions required for the asteroseismic measurement of stellar activity. Section 4 discusses the achievable accuracy of the seismic observables, while sections 5 and 6 present the results on solar data and for 16 CygA and B. This is followed by a discussion and conclusion in section 7.

## 2. THE EFFECT OF ROTATION AND OF MAGNETIC ACTIVITY

This section presents the effect of rotation and of magnetic activity on pulsation modes, and introduces the frequency model used for asteroseismic data analysis.

### 2.1. Frequency splittings

Slowly rotating stars without significant magnetic activity are approximately spherical and it is common to describe the family of modes traveling inside it using spherical harmonics (see e.g. Unno et al. 1989). If the departure from sphericity remains small enough, it is convenient to keep the spherical representation for the equilibrium model and account for distortions through a perturbation analysis. All pulsations can then be described using a set of integers  $(n,l,m)$ , namely the radial order, the mode degree and the azimuthal order, respectively. Acoustic pressure modes observed in Sun-like stars can be identified using their frequencies  $\nu_{nlm}$ .

In a non-rotating, non-active star,  $m$ -components are degenerate and cannot be resolved. When rotation or magnetic activity sets in, this degeneracy is lifted. The resulting frequency is treated as a perturbation to the degenerate frequency without rotation and activity,

$$\nu_{nlm} = \nu_{nl} + \delta\nu_{nlm}, \quad (1)$$

with  $\nu_{nl}$ , the equilibrium eigenfrequency **without rotation and activity** and  $\delta\nu_{nlm}$  the frequency splitting. These splitting may depend on multiple physical effects perceived by the modes (Libbrecht & Woodard 1990) within their cavity of propagation. These can be terms of order  $O(\Omega)$ , with  $\Omega$  the rotation rate estimated at the equator,  $\Omega = \Omega(r, \theta = \pi/2)$ . These depend directly on the stellar rotation profile  $\Omega(r, \theta)$ . Higher-order perturbations pertaining to the shape of the mode cavity can also exist. **Note that  $\nu_{nl}$  differs from  $\nu_{nl,m=0}$  as the  $m = 0$  components may have their frequency shifted by perturbations such as the magnetic activity (see e.g. Figure 1).**

Splitting can be described using the Clebsch-Gordan a-coefficient decomposition (Ritzwoller & Lavelle 1991), that corresponds to a representation of the splittings on a base of polynomials  $\mathcal{P}_j^{(l)}(m)$  of degree  $j$  in  $m$ ,

$$\delta\nu_{nlm} = \sum_{j=1}^{j_{\max}} a_j(n, l) \mathcal{P}_j^{(l)}(m), \quad (2)$$

with the polynomials defined orthogonal at discrete values of  $m$ ,

$$\sum_{m=-l}^l \mathcal{P}_i^{(l)}(m) \mathcal{P}_j^{(l)}(m) = 0 \text{ when } i \neq j. \quad (3)$$

Here,  $a_j(n, l)$  is the a-coefficient of order  $j$  and  $j_{\max} = 2l$  is the maximum order to which the decomposition must be carried for a given degree. The standard set of polynomials used in this expansion are those normalized as per described by Schou et al. (1994).

This decomposition is extensively used in helioseismology and was used on Sun-like stars by Benomar et al. (2018a). Asteroseismology has been so far unable to observe modes of degree higher than  $l = 3$  so that in the following the discussion is restricted to  $l \leq 3$  and  $j_{\max} = 2l = 6$ , due to the selection rule of the  $\mathcal{P}_j^{(l)}(m)$ . This limitation is the consequence of full-disk integrated photometric observations. An example of splitting including odd and even  $a_j$  coefficients is given in Figure 1 for  $l = 1, 2$ .

This theoretical model leads to a natural interpretation of the observed splittings. One may decomposes these latter into their symmetrical,  $S_{nlm}$ , and anti-symmetrical,  $T_{nlm}$ , parts. These components can then be described as sums over, respectively, the odd and even a-coefficients,

$$S_{nlm} = \frac{\nu_{n,l,m} - \nu_{n,l,-m}}{2m} = \frac{1}{m} \sum_{j=1}^{j_{\max}/2} a_{2j-1}(n, l) \mathcal{P}_{2j-1}^{(l)}(m), \quad (4)$$

$$T_{nlm} = \frac{\nu_{n,l,m} + \nu_{n,l,-m}}{2} - \nu_{n,l,0} = \sum_{j=1}^{j_{\max}/2} a_{2j}(n, l) (\mathcal{P}_{2j}^{(l)}(m) - \mathcal{P}_{2j}^{(l)}(0)). \quad (5)$$

This arises from the parity relation  $\mathcal{P}_j^{(l)}(-m) = (-1)^j \mathcal{P}_j^{(l)}(m)$ . These equations can be used to express the a-coefficients with  $S_{nlm}$  and  $T_{nlm}$ . They also provide relations between the  $\nu_{nlm}$  with the  $a_j(n, l)$  (Appendix A.2).

It can be seen from Appendix A.1 that the sums in Eqs. (4) and (5) involve, respectively, odd and even functions of  $m$ . Physically, this means that the symmetrical components of the splittings (Gough & Thompson 1990) result from large-scale perturbations sensitive to the prograde or retrograde nature of the waves, such as advection or the Coriolis force. On the other hand, the anti-symmetric splittings are caused by processes that are not affected by the propagation direction of waves. This may include the centrifugal force, that scales as  $O(\Omega^2)$ , and whose effect on the oscillation frequencies varies as  $m^2$ . Magnetic fields or non-spherical deformations of the equilibrium structure will also contribute to the anti-symmetric splittings.

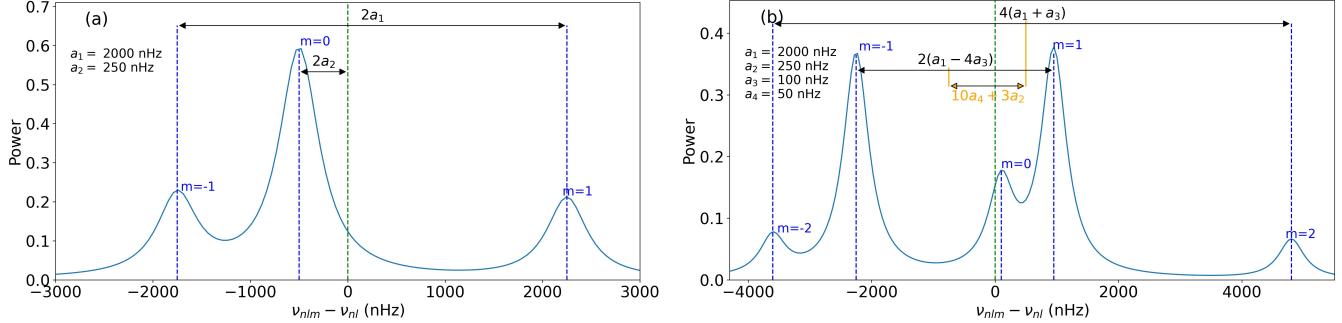
We further decompose the anti-symmetric splittings into a term depending on centrifugal-force-induced distortions and another one accounting for activity-related distortions,

$$T_{nlm} = \delta\nu_{nlm}^{(CF)} + \delta\nu_{nlm}^{(AR)}. \quad (6)$$

The symmetric part of the splitting corresponds to a term  $\delta\nu_{nlm}^{(rot)}$  that stems from perturbations of order  $O(\Omega)$  and the total observed splitting is,

$$\delta\nu_{nlm} = \delta\nu_{nlm}^{(rot)} + \delta\nu_{nlm}^{(CF)} + \delta\nu_{nlm}^{(AR)}. \quad (7)$$

Using the Sun as an archetype of Sun-like star, it is possible to measure these contributions to  $\delta\nu_{nlm}$  as shown in Sections 2.2 and 2.3.



**Figure 1.** Example of Lorentzian mode profiles showing the a-coefficients and their relationship with frequencies spacings for  $l = 1$  (a) and  $l = 2$  modes (b). The orange spacing is  $T_{n22} - T_{n21}$ .  $\nu_{nl}$  is the m-averaged frequency of each multiplet. The stellar inclination is 40°.

## 2.2. Expressions for $\delta\nu_{nlm}^{(rot)}$ , $\delta\nu_{nlm}^{(CF)}$ and $\delta\nu_{nlm}^{(AR)}$

To the first order, the perturbation on the frequency due to rotation is,

$$\delta\nu_{nlm}^{(rot)} = \frac{m}{2\pi} \int_0^R \int_0^\pi K_{nlm}(r, \theta) \mathcal{Q}(r, \theta) r dr d\theta. \quad (8)$$

R is the radius of the star and the kernel  $K_{nlm}(r, \theta)$  (Hansen et al. 1977) expresses the sensitivity of a mode to the rotation at the radial point  $r$  and co-latitude  $\theta$ .  $\mathcal{Q}(r, \theta)$  is the rotation profile of the star. It can be shown that  $\delta\nu_{nlm}^{(rot)}$  actually only depends on symmetrical splittings (Ritzwoller & Lavelle 1991) which in turns depend only on odd coefficients. For example and for  $l \leq 3$ , it is expressed as,

$$\delta\nu_{nlm}^{(rot)} \approx \mathcal{P}_1^{(l)}(m) a_1 + \mathcal{P}_3^{(l)}(m) a_3 + \mathcal{P}_5^{(l)}(m) a_5. \quad (9)$$

Centrifugal forces typically distorts a spherical rotating sphere of gas into an oblate ellipsoid, elongated at the equator. Functional analysis shows that the contribution of centrifugal forces to the mode splitting scales with  $\mathcal{Q}^2 R^3 / 2\pi GM$ . Properly integrating higher-order terms of the perturbation expansion over the aspherical volume of the star and using asymptotic expressions for the equilibrium mode eigenfunctions (assuming  $n$  is large enough), leads to the following expression for the centrifugal-force component of the frequency splitting (Gough & Taylor 1984; Gough & Thompson 1990),

$$\delta\nu_{nlm}^{(CF)} = \frac{\mathcal{Q}^2 R^3}{GM} \nu_{nl} Q_{lm}, \quad (10)$$

with  $Q_{lm} \approx \frac{2}{3} \frac{l(l+1)-3m^2}{(2l-1)(2l+3)}$  factor depending on the density. Equation (10) means that

$$\delta\nu_{nlm}^{(CF)} = \mathcal{P}_2^{(l)}(m) a_2^{(CF)}(n, l) \quad (11)$$

with

$$a_2^{(CF)}(n, l) = -\frac{1}{2l+3} \frac{\mathcal{Q}^2 R^3}{GM} \nu_{nl} \quad (12)$$

since  $-(2l+3)Q_{lm} = \mathcal{P}_2^{(l)}(m)$ . It should be noted that the contribution of the centrifugal-force-induced deformation to the frequency splittings can be described by a linear combination of the  $\mathcal{P}_{2j}$  polynomials (Gough & Taylor 1984; Gough & Thompson 1990). Since the stars are assumed to be slowly rotating, one may only retain terms of order  $O(\mathcal{Q}^2)$ , which correspond to the contribution of the  $a_2$  coefficient alone.

One can further approximate equation (12) in order to express  $\delta\nu_{nlm}^{(CF)}$  as a function of quantities that can be obtained directly from the modeling of the acoustic power spectra of Sun-like stars (Benomar et al. 2018a). First, the stellar density of a Sun-like star scales to a good approximation with its large separation, that is the average distance in the frequency space between two modes of identical degree and consecutive orders (Ulrich 1986). In solar units, this reads  $\rho = (\rho_\odot / \Delta\nu_\odot) \Delta\nu$ , with the solar density  $\rho_\odot = (1.4060 \pm 0.0005) \cdot 10^3$  kg.m<sup>-3</sup> and  $\Delta\nu_\odot = 135.20 \pm 0.25$  μHz (García et al. 2011a). With an accuracy estimated to a few percents for Sun-like stars (White et al. 2011), the use of this scaling relation is thought to be a decent approximation.

The second simplification uses the fact that the Clebsch-Gordan coefficient decomposition of the frequency splitting imposes a one-to-one relationship between the a-coefficients and the coefficients of the decomposition of the velocity field into poloidal and

toroidal components (Ritzwoller & Lavelle 1991). Helioseismology suggests that the Sun rotates as a near-solid-body down to at least ( $r/R_\odot \approx 0.2$ ) (e.g. Thompson et al. 2003), which is the maximum depth at which measurements from low-degree  $\mathbf{p}$  modes remain accurate and precise. Furthermore, its outer-convective zone shows a differential rotation of  $\approx 30\%$  from the equator to the pole, which leads to an  $a_{3,\odot}(n, l) \approx 4$  nHz and to even smaller values for higher odd a-coefficients. This is significantly smaller than  $a_{1,\odot}(n, l) \approx 420$  nHz. Therefore, the aforementioned one-to-one relation ensures that we can retain only the leading order in the expansion of the rotation rate and treat it as an average value, given in terms of seismic observables by  $\mathcal{Q} \approx 2\pi a_1$ . This leads to

$$a_2^{(CF)}(n, l) \approx -\frac{\nu_{nl}}{2l+3} \frac{3\pi}{G\rho_\odot} \frac{\Delta\nu_\odot^2}{\Delta\nu^2} a_1^2. \quad (13)$$

Regarding  $\delta\nu_{nlm}^{(AR)}$ , there is no unambiguous theory to describe the effect of the near-surface magnetic activity on the shape of the cavity. Due to this and following Gizon (2002), a geometrical description is preferred to a physical model. **This description assumes that the activity perturbation separates in latitudinal and radial coordinates.** The proposed form of the perturbation in frequency is,

$$\delta\nu_{nlm}^{(AR)} = \nu_{nl}\epsilon_{nl} \int \int F(\theta|\mathbf{x}) |Y_l^m(\theta, \phi)|^2 \sin \theta d\theta d\phi = \nu_{nl}\epsilon_{nl} A_{lm}(\mathbf{x}) = \sum_{j=0}^{j_{max}/2} \mathcal{P}_{2j}^{(l)}(m) a_{2j}^{(AR)}(n, l). \quad (14)$$

The term  $a_{2j}^{(AR)}(n, l)$  of equation (14) refers to the contribution of the activity or a latitudinal temperature gradient to the a-coefficients. The geometrical weight function  $A_{lm}(\mathbf{x})$  describes the effect of an active zone at the co-latitude  $\theta$ , on a mode of degree  $l$  and azimuthal order  $m$ .  $A_{lm}(\mathbf{x})$  is the product of two contributions. First, the normalized spherical harmonics  $Y_l^m(\theta, \phi)$  that decompose the magnetic activity effect over each modes. These are spherical-polar coordinates defined in the inertial frame with a polar axis pointing in the direction of the rotation axis. Second, the weight distribution (or the shape of the active region) is defined by  $F(\theta|\mathbf{x})$ . Here,  $\mathbf{x}$  refers to the parameters that are necessary to describe the function  $F(\theta|\mathbf{x})$ .

On the Sun, spots remain apparent at the surface for over a 1-2 rotation periods and appear to be randomly distributed across the longitude. **As shown in Papini & Gizon (2019), this theoretically implies that additional split components should densely populate the spectrum. In our case, we assume that this additional splitting is small and that we have a steady perturbation that persists over the observation time.**

In theory, a third integral over the radius is necessary to describe the dependence of the magnetic activity to the stellar depth. However, p modes are weakly sensitive to the deep structure inside stars. Here, the radial integral is replaced by  $\epsilon_{nl}$ , the overall activity intensity. The frequency  $\nu_{nl}$  allows a dimensionless  $\epsilon_{nl}$ , that can be compared between stars. Section 3.2 further develops the required assumptions in order to obtain a reliable information content on the active region in asteroseismology.

### 2.3. Modeling the frequencies

Combining equation (1), (7), (8), (13) and (14) leads to,

$$\nu_{nlm} = \nu_{nl} \left( 1 + \frac{3\pi}{(2l+3)G\rho_\odot} \frac{\Delta\nu_\odot^2}{\Delta\nu^2} a_1^2 + \epsilon_{nl} A_{lm}(\mathbf{x}) \right) + \frac{m}{2\pi} \int_0^R \int_0^\pi K_{nlm}(r, \theta) \mathcal{Q}(r, \theta) r dr d\theta. \quad (15)$$

Using equation (1) and (2),  $\nu_{nlm}$  can also be expressed using the a-coefficients for  $l \leq 3$  without loss of generality,

$$\nu_{nlm} = \nu_{nl} + \sum_{j=1}^6 \mathcal{P}_j^{(l)}(m) a_j(n, l). \quad (16)$$

It is possible to use equation (15) in order to measure the activity parameters  $\mathbf{x}$ , by directly fitting the power spectrum. However, there are multiple benefits to use instead a two-step approach consisting in first using equation (16) for the fitting of the power spectrum to evaluate the a-coefficients (method detailed in Appendix B). And then, fitting the coefficients obtained during the first step, using solely the equations (13) and (14) (detailed in Appendix C). Firstly and in the general case, the evaluation of  $A_{lm}(\mathbf{x})$  requires the precise computation of a double integral. This is a slow process, that increases the time necessary to fit the power spectrum<sup>1</sup>. Postponning to an ulterior step the computation of  $A_{lm}(\mathbf{x})$  reduces drastically the parameter space from a few tens of parameters to only a few<sup>2</sup>, effectively making faster the convergence rate of the fitting algorithm and making it easier to explore

<sup>1</sup> Using a MCMC method, from a day to a couple of weeks for a single star on a present-day CPU

<sup>2</sup> With our choice of  $A_{lm}(\mathbf{x})$ , three parameters are fitted. See section 3.2.

the assumptions that have to be made on equation (15) in order to have a functional approach in real cases (see the discussions in Section 3). Secondly, it allows us to decouple the observables from the physics, enabling to test various physical assumptions without having to re-perform the lengthy power spectrum analysis. Finally, it eases the evaluation of the reliability of the activity determination, by enabling us to pinpoint the cause of biases (if any) on either the observables (eg.  $a_2, a_4$ , see Section 4) or on the interpretation of these observables in terms of physical parameters (see Section 3.3). The disadvantage of the two-step approach is that it requires more statistical assumptions, such as negelecting correlations and assuming Gaussian parameters. Although one can argue that it is possible to perform a hierarchical Bayesian analysis (eg. Hogg et al. 2010; Campante et al. 2016) to partly alleviate this issue, such an approach is generally slow and may cancel the benefits of the two-step approach.

### 3. INFORMATION CONTENT IN LOW ORDER A-COEFFICIENTS

The general theoretical formulations of Section 2, demonstrate how pulsation frequencies are expressed as function of the rotation, the centrifugal distortion and the activity of stars. However, observational limitations need to be accounted to enable a viable, robust model that extract as much possible the information content within currently existing data. This necessarily requires additional assumptions, for which the rational is detailed hereafter.

#### 3.1. Factors contributing to the splitting accuracy and precision

As explained in Kamiaka et al. (2018), the pulsation height (H) and the noise background (N) are important factors that reduce the capabilities of asteroseismic analyses. Those are a complex function of the global stellar characteristics (mass, radius, age), of the convective motion at the surface of Sun-like stars, and also depend on the instrumental limitations. The mode height and the noise background are in fact difficult to evaluate *a priori*. However, the Height-to-Noise (HNR), defined as the ratio  $H/N$  can be used to asses the quality of the spectrum of a pulsation mode. In fact, Kamiaka et al. (2018) show that the HNR at maximum of mode height  $\widehat{HNR}$ , can be used to study biases on the stellar inclination. This is because all main sequence Sun-like stars show a similar dependency of the height and of the width as a function of the frequency (see e.g. Appourchaux et al. 2014). In Section 4, we propose the same, but on a-coefficients.

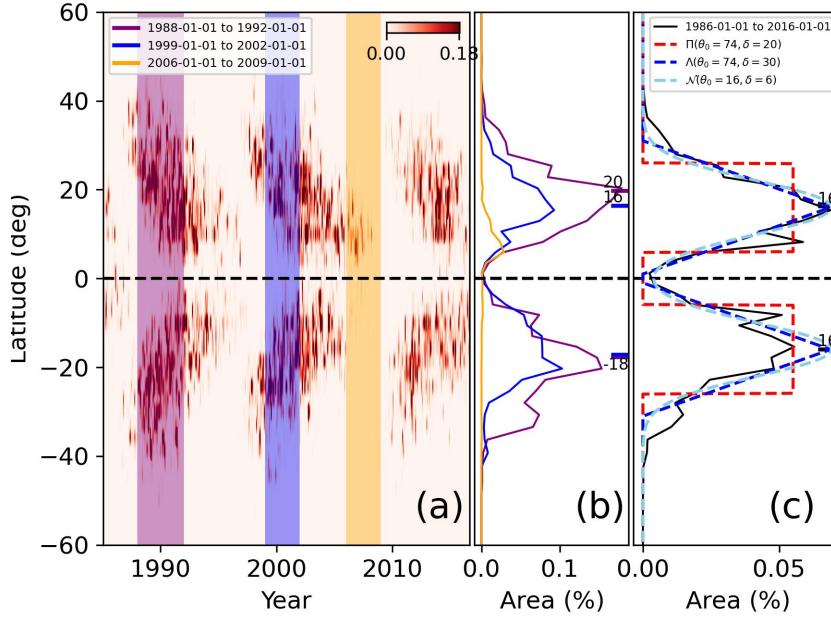
The  $\widehat{HNR}$  not only defines how many modes can be observed, but also the maximum degree of the modes that is observed. Considering specifically the Kepler observations, stars exhibit a  $\widehat{HNR}$  for  $l = 0$  modes of up to 30, the highest HNR being observed for 16 Cyg A and 16 Cyg B. Kamiaka et al. (2018) showed that the capability of distinguishing rotationally split components is of great importance if one wants to obtain a reliable asteroseismic inference of the rotation characteristics and of the stellar inclination. In particular, the ratio  $a_1/\Gamma_{v_{max}}$  between the  $a_1$  coefficient and the mode width determined at the maximum of mode amplitude  $\Gamma_{v_{max}}$  (see Figure 6), determines the expected bias on the stellar inclination. As shown in Section 4,  $a_1/\Gamma_{v_{max}}$  also controls the importance of the bias for the other low order a-coefficients. Finally, another important factor is the spectrum resolution. The higher the resolution, the more resolved are the modes. Thus, observations  $T_{obs}$  of several years are the most suitable in order to resolve and measure rotationally split components. Broadly speaking, observations exceeding a year and  $a_1/\Gamma > 0.4$  are preferable to ensure a reliable measurement of  $a_1$ .

All of the limitations discussed above incite us to introduce assumptions to ensure robust measurements of a-coefficients (ie, mitigate biases).

#### 3.2. Latitudinal profile of the activity

Noting that it is challenging to measure low-degree a-coefficients for the Sun (eg. Toutain & Kosovichev 2001), we present here a minimal set of assumptions on even a-coefficients allowing us to constrain the asphericity of stellar cavities. One of the first aspects that has to be considered is the form of  $F(\theta|x)$ , the function characterising the activity latitude  $\theta$  (see equation (14)).

**The butterfly diagram of the Sun (Figure 2) is used as a reference for this latitudinal dependence.** The data are from the Greenwich USAF/NOAA observatory (<https://solarscience.msfc.nasa.gov/greenwch.shtml>) and provide the daily area of the spots, counted manually over the period 1874-2016. The panel (a) on Figure 2 shows the butterfly diagram with colors representing the area covered by the spots in unit of percent of visible hemisphere. **It focuses on the observations after 1985 and covers two full solar cycles.** Vertical color bands highlight three time intervals : 1985-1989 (Purple), 1999-2002 (Blue) and 2006-2009 (Yellow). The two first are during a maximum of solar activity while the last one is for a minimum of activity. Due to the gradual migration of the spots over time, the longer observation period (1985-1989, 4 years) leads to a broader activity zone than the period 1999-2002. This indicates that the extension of the active region may not be trivial to measure in other stars because it will depend on the fraction of time the star is observed relative to the duration of its activity cycle. The activity cycles of Sun-like stars (if any) is a priori unknown, but Ca II H+K line emission and photometric studies (Oláh et al. 2009, 2016)



**Figure 2.** (a) Butterfly diagram for the Sun. Vertical color bands highlight periods of maximum of 1988-1992, 1999-2002 and minimum of activity of 2006-2009. (b) averaged spot area for the highlighted periods. Colored ticks on (b) are the weighted mean for the activity latitude. (c) **Filters  $F(\theta|X)$  used for this study superimposed to solar data (1986-2016). Parameters  $\theta_0$  and  $\delta$  are the latitude and the extension of the active region, respectively.**

suggest that it is of durations of roughly a few years to decades, as for the Sun. It indicates that over the course of several years, an active band as large as  $\approx 40^\circ$  may be expected.

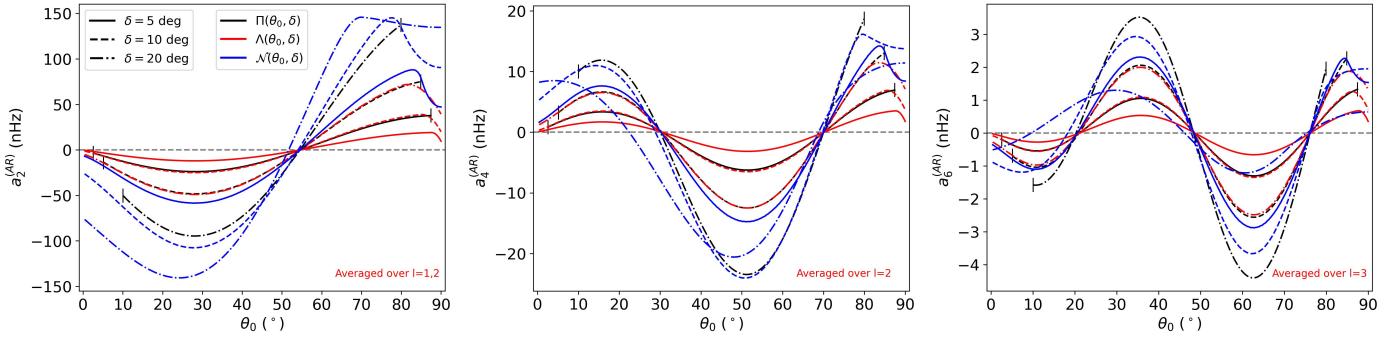
Panel (b) of Figure 2 shows the cumulative area of spots as a function of the latitude and for the three considered periods. As noted earlier, the extension of the activity band is larger for the longest timeframe. The area of the spots during the active solar phase are symmetrical towards the equator. This suggests that when the activity is strong,  $F(\theta|x)$  is mirrored between the north and south hemisphere. This may be inaccurate for low activity phase, as shown for the period 2006-2009 **but because**  $|Y_l^m(\theta, \phi)|^2$  **is also symmetrical towards the equator**, this has no incidence on the  $a_j^{(AR)}(n, l)$  coefficients.

During the minimum, the total average area of the spots is a few times lower than during the maximum. It is also narrower, such as the integral in equation (14) is small, reducing  $a_j^{(AR)}(n, l) \approx 0$ . During the phase of minimum of activity, the activity can effectively be considered as nonexistent (see Section 5 for the analysis on solar data), so that  $a_2(n, l)$  is dominated by the centrifugal term  $a_2^{(CF)}(n, l)$  and the other even a-coefficients are null.

The overall latitudinal profile is seemingly following a bell-shape with sharp slopes **during periods of activity**. Figure 2c compares data between 1986 and 2016 with three models for the active latitudes: a model using the gate function  $F(\theta|x) = \Pi(\theta_0, \delta)$ , a triangular function  $\Lambda(\theta_0, \delta)$  and a Gaussian function  $N(\theta_0, \delta)$ . The parameters of these were adjusted manually to approximately match the solar spot active latitudes profile. The triangular and the Gaussian functions describe equivalently the data, while the gate function initially proposed by Gizon (2002) is roughly fitting the data. Because spots are strictly appearing at latitudes below  $45^\circ$ , the  $\Lambda$  function may be the most suitable for the Sun. Nevertheless, these three functions are retained and compared hereafter.

### 3.3. Assumptions on the a-coefficients

The most direct method for measuring the asphericity is to evaluate it directly for each mode  $(n, l)$ , that is, measuring the terms  $a_j(n, l)$ . This being already difficult for low-order a-coefficients of the Sun (Chaplin et al. 2003), it seems unreasonable to expect an accurate measurement of all individual  $a_j(n, l)$  using asteroseismic data. These have a lower signal-to-noise ratio and severely reduced visibility at  $l \geq 3$  due to the integrated photometry. **After a trial and error process, jointly with power spectra simulations, we could identify a set of assumptions** ensuring reliable and precise measurement of a-coefficients.



**Figure 3.** Average  $a_j^{(AR)}$  (no centrifugal effect accounted for) with  $\epsilon_{nl} = 5.10^{-4}$  in function of  $\theta$  and  $\delta$  and for the gate ( $\Pi$ ), triangle ( $\Lambda$ ) and gaussian ( $\mathcal{N}$ ) filter functions. Uniqueness is guaranteed for  $\theta$  only if at least two a-coefficients are measured.

We first consider a fictitious star rotating as a solid-body with a solar activity level ( $\epsilon_{nl} \simeq 5.10^{-4}$ , Gizon (2002)) with  $a_1 = 1000$  nHz. The observed oscillation frequencies of 16 Cyg A are used here (Davies et al. 2015; Kamiaka et al. 2018). Split frequencies are derived using equations (1), (7), (13), (14), and converted into a-coefficients using equations (A8-A19). The a-coefficients are linear functions of the frequency, which is expected as  $F(\theta|\theta_0, \delta)$  describes a single active region. Thus, a reasonable assumption is to consider those as pure first order polynomial functions of frequencies. However, tests on artificial spectra showed that it is often difficult to evaluate the slope of the a-coefficients. This is because the uncertainty on any  $a_j$  is at least of the same order as its variations within the range of observed frequencies. This suggests that current asteroseismic data lack the resolution and the signal-to-noise to reliably measure the frequency dependence on the a-coefficients.

Figure 3 shows  $a_2^{(AR)}$ ,  $a_4^{(AR)}$ ,  $a_6^{(AR)}$  for activity described by  $F = \Pi$  (black) or  $\Lambda$  (red) or  $\mathcal{N}$  (blue). These are the mean coefficients for the fictitious active star, as a function of  $\theta_0$  and  $\delta$ . The figure indicates that there is a simple relationship between the a-coefficients, the co-latitude  $\theta_0$  and the extension of the activity zone  $\delta$ , independently of the shape of activity. In the case of  $F = \Pi$ , the lines are cut near the pole and the equator due to the condition  $\theta_0 \geq \delta/2$  and  $\theta_0 \leq \pi - \delta/2$ . In the other profiles, edges effects (truncation) has noticeable effects near the equator. The definition of  $\delta$  differs between the three profiles, which explains that a factor of a few in  $a_j^{(AR)}$  is noticeable between  $F = \Pi$ ,  $\Lambda$  and  $\mathcal{N}$  for a given  $\delta$ . Note that adding the centrifugal effect reduces  $a_2(n, l) = a_2^{(AR)}(n, l) + a_2^{(CF)}(n, l)$ , because the coefficient  $a_2^{(CF)}(n, l)$  is always negative.

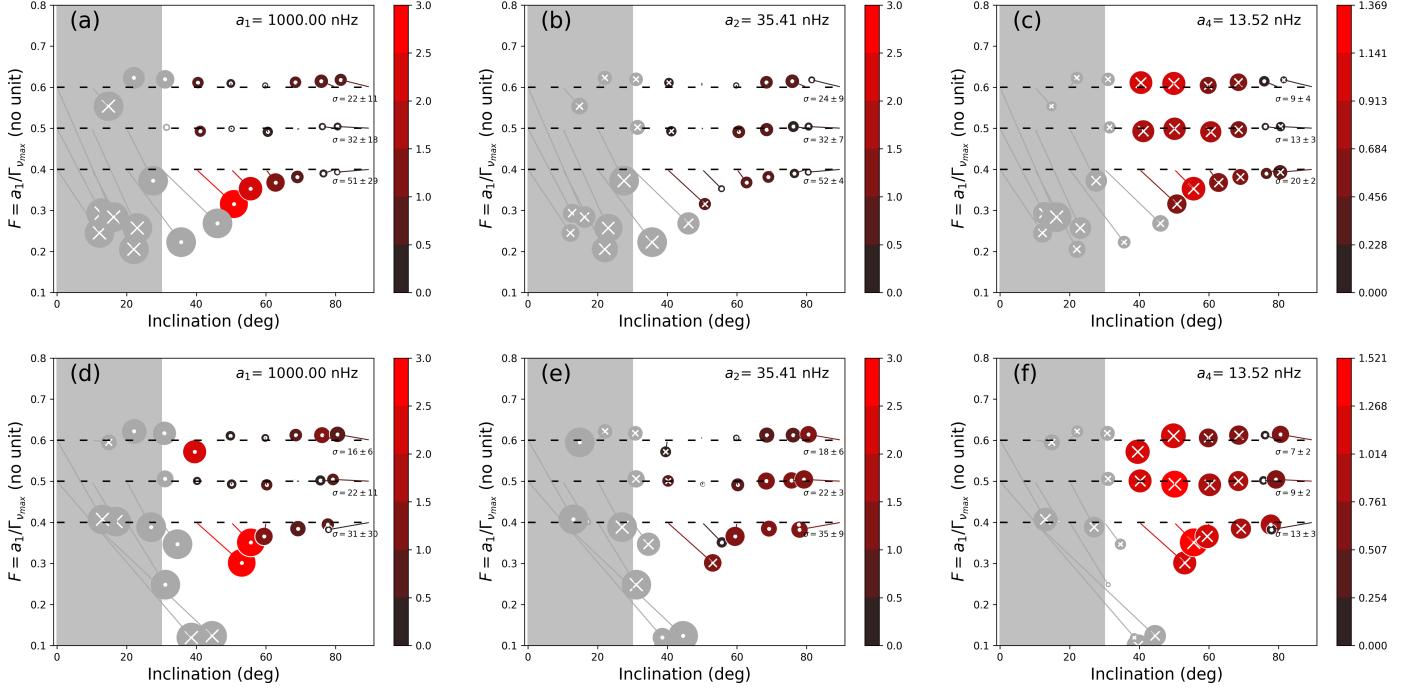
If an observation can constrain only a single a-coefficient (e.g.  $a_2$ ), there often exists a degeneracy in  $\theta_0$  as multiple value of  $a_j$  can be obtained for a given  $\theta_0$ . As the figure shows, measuring two a-coefficients alleviates this degeneracy, provided that uncertainties are small enough. In other words, the accuracy on the inference of the active region using a-coefficients is ensured only if we can simultaneously constrain two a-coefficients (e.g.  $a_2$  and  $a_4$ ). This is essential to distinguish an activity near the pole ( $\theta_0 \lesssim 30^\circ$ ) from mid-latitude ( $30^\circ \lesssim \theta_0 \lesssim 60^\circ$ ) or from near the equator ( $\theta_0 \gtrsim 60^\circ$ ).

It indicates that the loss of information resulting from the averaging does not affect the accuracy<sup>3</sup>. Figure 3 also indicates that for a star with an activity intensity and an activity zone of extension commensurate with the one of the Sun ( $\epsilon_{nl} \simeq 5.10^{-4}$ ,  $\delta \simeq 10^\circ$  when  $F = \Pi$ ), the uncertainty on  $a_2^{(AR)}$  must be approximately  $\lesssim 25$  nHz, the one for  $a_4^{(AR)} \lesssim 10$  nHz and  $a_6^{(AR)} \lesssim 1.5$  nHz to be able to detect significant departures of the coefficients. This is to be compared with the spectral resolution of 7-14nHz for 2-4 years of observation, typical of the longest Kepler observations. In the Figure 3,  $a_2^{(AR)}$  is obtained by averaging  $l = 1, 2$ ,  $a_4^{(AR)}$  by averaging  $l = 2$ , while  $a_6^{(AR)}$  is only constrained by  $l = 3$  modes. Averaging  $a_2^{(AR)}$  and  $a_4^{(AR)}$  over  $l = 1, 2, 3$  increases the maximum range of the a-coefficients by a factor  $\approx 1.5$  for  $a_2^{(AR)}$  and  $\approx 2$  for  $a_4^{(AR)}$ , respectively, without changing the overall shape of the function. In Sun-like stars, the  $l = 3$  modes have a HNR at least ten times lower than  $l = 0$  modes because the height ratio between  $l = 3$  and  $l = 0$  modes is around 0.08 for the Sun (Toutain & Gouttebroze 1993; Toutain et al. 1998). These modes are therefore difficult to observe. Due to all of the above, it is extremely challenging to measure  $a_6$  with current existing data. In the following, we will thus focus on assessing the reliability domains of  $a_2$  and  $a_4$  only.

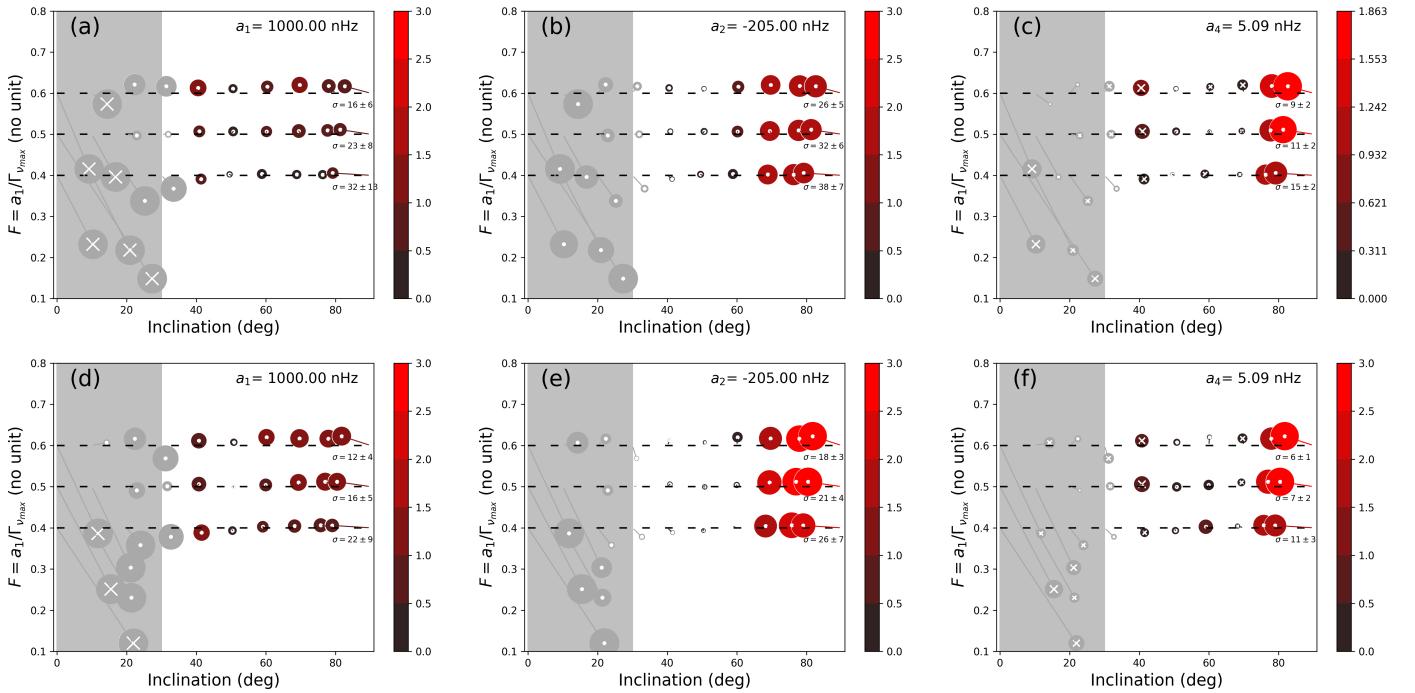
#### 4. BIAS ANALYSIS ON $A_1$ $A_2$ AND $A_4$

In order to understand the reliability of the inference on low-degree a-coefficient, it is necessary to perform a bias analysis. This requires to fit an ensemble of emulated spectra that are representative of Sun-like stars and to compare the results with the true inputs.

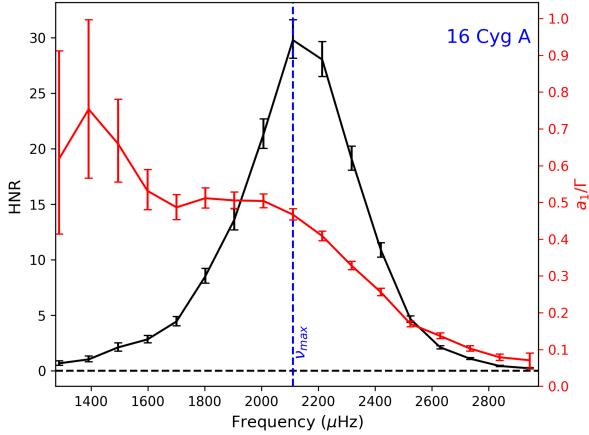
<sup>3</sup> It does however increase the uncertainty by a factor of a few.



**Figure 4.** Bias analysis for  $\widehat{HNR} = 30$  for an equatorial activity band ( $\theta_0 = 85^\circ, \delta = 10^\circ$ ) of similar intensity to the Sun ( $\epsilon_{nl} = 5.10^{-4}$ ), for  $T_{obs} = 2$  years (top) and  $T_{obs} = 4$  years (bottom). Color and size of the circles indicate the modulus of the bias. A white cross is for an underestimation. A white dot is for an overestimation. Below  $30^\circ$  of inclination (gray area and symbols), the results are not reliable.



**Figure 5.** Bias analysis for  $\widehat{HNR} = 30$  for a large polar activity cap ( $\theta_0 = 22.5^\circ, \delta = 45^\circ$ ) of similar intensity to the Sun ( $\epsilon_{nl} = 5.10^{-4}$ ), for  $T_{obs} = 2$  years (top) and  $T_{obs} = 4$  years (bottom).



**Figure 6.** HNR (black) and  $a_1/\Gamma$  ratio (red) for 16 Cyg A, which is the reference star used to make simulations.

#### 4.1. Test setup

The synthetic spectra use the frequencies, heights, widths and the noise background profile of 16 Cyg A as a template for the simulations. We want to specifically study the impact of the mode blending (effect of  $a_1/\Gamma_{\nu_{max}}$ ), the stellar inclination ( $i$ ), observation duration ( $T_{obs}$ ) and of the maximum Height-to-Noise background ( $\widehat{HNR}$ ) on the accuracy of  $a_1$ ,  $a_2$  and  $a_4$ . Grids of spectra are built in the case of an equatorial band of activity and of a large polar activity. The grid parameters and their ranges are provided in Table 4.1 and discussed further below. The spectra are made using a spectrum simulator code<sup>4</sup> that take the reference star (or template) and modify its properties to match the requirement of the user. The template is altered in terms of HNR following a similar approach as in Kamiaka et al. (2018). The main difference is that we considered a frequency-dependent noise background. Heights are rescaled according to,

$$H_{n,l=0} = \frac{\widehat{HNR}}{\widehat{HNR}_{ref}} H_{ref}(n, l=0) \quad (17)$$

where  $\widehat{HNR}$  is the maximum HNR of the synthetic star,  $\widehat{HNR}_{ref}$  the maximum HNR of the reference star and  $H_{ref}(n, l=0)$  the  $l=0$  heights of the reference star.

As for the mode blending factor  $f = a_1/\Gamma_{\nu_{max}}$ , it is calculated fixing  $a_1$  and altering  $\Gamma_{\nu_{max}}$  such that,

$$\Gamma_{\nu_{max}} = \frac{a_1}{F}. \quad (18)$$

A fiducial value of  $a_1 = 1000$  nHz ( $\simeq 2.4$  the solar rotation) is used in the simulation. Note that this differs from Kamiaka et al. (2018) where the splitting was modified in order to obtain the desired mode blending factor.

As shown in Table 4.1, three HNR cases are investigated, ranging from 10 to 30. Two observation durations are considered: 2 years and 4 years. This is representative of the best Kepler observations (Davies et al. 2015) and of future observations from PLATO (Rauer et al. 2014). Similarly to the test cases of Gizon (2002),  $a_2$  and  $a_4$  are determined for an equatorial activity of extension of the same order as in the Solar case ( $\delta \simeq 10^\circ$ ), and in the case of a large polar cap ( $\delta = 45^\circ$ ). Both situations assume an activity of the same order as for the Sun. The priors are set in a similar manner as they would if the power spectrum was from a real star. The evaluation of the mode parameters is performed using a MCMC method, like for actual stars (see Appendix B for further details), but on the limit-spectrum (no noise). Fitting the limit spectrum allows to asses the systematic errors by calculating the expectation value of the probability density function. The expected uncertainty can also be known by computing the standard deviation.

#### 4.2. Results

<sup>4</sup> The tool suite used here along with the data inputs/outputs are available at [https://github.com/OthmanB/Benomar2022/tree/version\\_2](https://github.com/OthmanB/Benomar2022/tree/version_2).

Variable	Values	
HNR	10, 20, 30	
$T_{\text{obs}}$ (years)	2, 4	
$a_1/\Gamma_{\nu_{\text{max}}}$	0.4, 0.5, 0.6	
$\epsilon_{nl}$	$5.10^{-4}$	
$a_1$ (nHz)	1000	
Active Region	$\theta_0 = 85^\circ, \delta = 10^\circ$ (Eq.)	$\theta_0 = 22.5^\circ, \delta = 45^\circ$ (Pol.)
$a_2$ (nHz)	35.41	-205.06
$a_4$ (nHz)	13.52	5.09

**Table 1.** Parameters set to construct artificial power spectra.  $a_2$  and  $a_4$  are determined assuming active regions located either in an equatorial band or in a polar cap.

	Active Sun	Quiet Sun	16 Cyg A	16 Cyg B
$\widehat{HNR}$	100	70	30	29
$\Delta\nu$ ( $\mu$ Hz)	135.1	135.1	103.35	116.92
$a_1$ (nHz)	$421 \pm 10$	$410 \pm 9$	$614 \pm 37$	$607 \pm 78$
$a_1/\Gamma$	$0.45 \pm 0.03$	$0.40 \pm 0.02$	$0.45 \pm 0.04$	$0.58 \pm 0.1$
inclination ( $^\circ$ )	90 (fixed)	90 (fixed)	$45 \pm 4$	$35 \pm 3$
$a_2$ (nHz)	$80 \pm 19$	$11 \pm 21$	$19.6 \pm 8.6$	$18.5 \pm 23.5$
$a_4$ (nHz)	$5.0 \pm 10.5$	$2.1 \pm 9.8$	$2.9 \pm 8.9$	$-27.9 \pm 6.5$ or $-1.0 \pm 6.7$
$d_2^{(\text{CF})}$ (nHz)	$-6.5 \pm 0.3$	$-6.1 \pm 0.3$	$-17.0 \pm 2.2$	$-15.5 \pm 4.0$
$d_2^{(\text{AR})}$ (nHz)	$86.2 \pm 18.7$	$16.7 \pm 20.4$	$39.0 \pm 20.6$	$34.0 \pm 21.8$
$b(a_2)$ (nHz)	/	/	-10	-10
$b(a_4)$ (nHz)	/	/	-10	-10

**Table 2.** Summary of the main parameters for the Sun and 16 Cyg A and B used to infer the activity. fiducial biases  $b(a_2)$  and  $b(a_4)$  for 16 Cyg A and B are set after inspection of Figure 4. The lowest (highest) solution of  $a_4$  for 16 Cyg B is has a probability of 30% (70%).

To appreciate the level of inaccuracy achieved when measuring  $a_j, j \in [1, 2, 4]$  coefficients, the bias is calculated for  $a_j$ ,  $f = a_1/\Gamma_{\nu_{\text{max}}}$  and  $i$ ,

$$\begin{aligned} b(a_j) &= E[a_j^{\text{meas}}] - a_j^{\text{true}}, \\ b(F) &= E[F^{\text{meas}}] - F^{\text{true}}, \\ b(i) &= E[i^{\text{meas}}] - i^{\text{true}}. \end{aligned} \quad (19)$$

Here,  $X^{\text{meas}}$  and  $X^{\text{true}}$  refer to the measured and the true value of the parameters  $X = \{a_j, F, i\}$ . The letter E refers to the expectation value (median) from the probability density function.

Figure 4 shows the resulting bias maps for a-coefficients as a function of the blending factor  $F$  and of the stellar inclination for a-coefficients corresponding to an equatorial activity band. Figure 5 is the same but for a large polar activity cap. The bias is represented as a projected three dimensional vector using the three quantities defined by equation (19) and for  $j \in [1, 2, 4]$ . To evaluate the importance of the bias on  $a_j$ , the plot shows  $b(a_j)/\langle \sigma_{a_j}(i, F) \rangle_i$ , with  $\langle \sigma_{a_j}(i, F) \rangle_i$  the average standard deviation of the probability density functions obtained by the MCMC sampling over  $i \in [30^\circ, 90^\circ]$ . The value of  $\langle \sigma_{a_j}(i, F) \rangle_i$  (noted  $\sigma$  for convenience hereafter) and its variance along the inclination axis is also shown on the plots. A value of  $b(a_j)/\sigma$  greater than one indicates that the inaccuracy exceeds the typical expected uncertainty at  $1\sigma$  and may lead to significantly biased results during the subsequent analyses of the a-coefficients. A negative bias (underestimation) for  $a_j$  is indicated by a cross within the circle while a positive bias (overestimation) by a dot. The darker the color, the smaller the norm of the normalised bias. The size of the circle symbols is also proportional to the bias, but capped to 3, to avoid excessively large symbols when  $i < 30^\circ$  (see next paragraph for explanations). The case of observations of duration of 2 years are in the top figures (a, b, c) while the 4 years observation case is shown in bottom figures (d, e, f).

We first note that the biases on inclination and on  $F$  are consistent with results from Kamiaka et al. (2018). Regions of stellar inclination below  $30^\circ$  are found to provide very unreliable results and the median of the inclination is biased toward  $\simeq 80^\circ$  when the true input is  $i = 90^\circ$ . Meanwhile, biases in  $a_1$  are found to often exceed 50% in the gray area<sup>5</sup>, indicating that for inclination below  $\simeq 30^\circ$ , the median estimator of the probability density function might not be trusted. This justifies also the range for the calculation of  $\sigma$ . In the following we therefore focus the discussion of the figures for the region above  $30^\circ$ .

In the case of an equatorial zone of activity, we note that measurements of  $a_1$  may have negligible inaccuracies when the blending factor  $F$  is above 0.4, but starts to be significantly deteriorated at (and probably below)  $f = 0.4$  because  $|b(a_1)|/\sigma$  exceeds 1.5 when  $i < 50^\circ$ . Split components are then overlapping significantly leading to important degeneracies between a-coefficients, mode widths and the stellar inclination. Meanwhile, we note that the biases on  $a_2$  and  $a_4$  remain mild compared to the uncertainty ( $|b(a_2)|/\sigma < 1$  and  $|b(a_4)|/\sigma < 1$ ), even for  $f = 0.4$ .

In the case of a polar active region,  $|a_2|$  represents  $\simeq 20\%$  of the  $a_1$  coefficient. In these conditions, we note that  $b(a_1)/\sigma$  does not exceed the unity (see (a) and (d)), indicating a good accuracy. The terms  $|b(a_2)|/\sigma$  and  $|b(a_4)|/\sigma$  are also lower than 0.5, provided that the stellar inclination is between  $30^\circ$  and  $70^\circ$ . This indicates that a large  $a_2$  coefficient is generally associated with a higher accuracy for all a-coefficients. The plot also demonstrates that a large zone in the parameter space has moderate to small bias.

Interestingly, figures for  $\widehat{HNR}$  of 10 and 20 (Figures 16-19 in Appendix D) lead to similar conclusions. Therefore, even in less favorable  $\widehat{HNR}$  conditions, the expected uncertainty on the measurement generally encompass the bias provided that the stellar inclination exceeds  $\simeq 30^\circ$  (ensuring that  $m \neq 0$  have significant amplitudes) and that the mode blending factor is above 0.4. This indicates that for a majority of stars observed by Kepler, the accuracy may not be a major issue when measuring  $a_1$ ,  $a_2$  and  $a_4$  and provided that a careful assessment of  $F$  and  $i$  is performed. Note however that the average uncertainty  $\sigma$  does usually increase when the  $\widehat{HNR}$  and/or the observation duration decrease, which will have an impact on the precision of the determination of the active region.

## 5. ANALYSIS OF SOLAR DATA

The Sun has a well-known 11-year activity cycle that makes it ideal for testing further the accuracy of the method. The Sun activity cycle is analysed at two instants, highlighted in Figure 2 and for which high-quality asteroseismic data are available: (a) the maximum of activity between January 1999 and January 2002; and (b) the minimum of activity between January 2006 and January 2009. Data are from the Variability of Solar Irradiance and Gravity Oscillations instrument aboard SOHO spacecraft (Frohlich et al. 1997) and presents very few gaps.

### 5.1. $a$ -coefficients of the Sun

As for the simulated data, the analysis consists in fitting the power spectrum for cases (a) and (b), to measure the  $a_1$ ,  $a_2$  and  $a_4$  coefficients using the Bayesian modeling and the MCMC method described in the Appendix B. **Figures of the best fits and their discussions are provided in Appendix E.** The goal being to evaluate the accuracy of the inference of the activity zone from a-coefficients, the presented results are for a stellar inclination fixed to  $90^\circ$ , instead of having it as a free parameter. This value correspond approximately to the stellar inclination as seen by the SoHo satellite. This alleviates biases on  $a_j$  coefficients that may arise due to the systematic underestimation of  $i$  when it is close to 90 degrees. However, carried tests with a free inclination during the maximum of activity<sup>6</sup> lead to a difference of  $-12\text{nHz}$  in  $a_2$  and to  $-1.3\text{nHz}$  in  $a_4$ . This is consistent with the expected difference from the bias map of Figure 4. In agreement with the discussion of Section 4, the difference does not have a significant impact on the activity inference because the uncertainties for those parameters are larger than the observed measurement shift.

Figure 7 shows the measured probability distribution function of the relevant parameters along with their correlations during the maximum of solar activity (left) and the minimum (right). The  $a_2^{(CF)}$  distribution in red represents the expected centrifugal effect on  $a_2$  as derived from frequency shifts  $\delta_{nlm}^{(CF)}$  of equation (13). The distributions are Gaussians and show a weak correlation, reflecting the quality of the data. Between the maximum and the minimum of activity, the  $a_2$  coefficient drifted significantly, from  $a_2 = 80 \pm 19 \text{nHz}$  to  $a_2 = 11 \pm 21$  falling within the  $1\sigma$  confidence interval of the centrifugal term. Although  $a_4$  may have changed, the effect is below uncertainty levels and remain close to 0. As shown in Figure 15, other time intervals may lead to  $a_4$  departing from 0. The figures demonstrate that a comparison of  $a_2$  and  $a_2^{(CF)}$  may reveal the activity of a Sun-like star. At the maximum of activity  $a_2^{(CF)}$  is inconsistent with  $a_2$  at  $4.5\sigma$ , but it is in agreement at  $1\sigma$  during the minimum of activity. Interestingly, Chaplin et al. (2003) also studied in details the frequency asymmetry of  $l = 2$  modes using BiSON and GOLF

<sup>5</sup> This is likely due to the fit mis-identifying the  $l = 2, m = \pm 1$  with  $l = 2, m = \pm 2$ , effectively doubling or halving the  $a_1$  coefficient.

<sup>6</sup> The measured stellar inclination is  $i = 74 \pm 3$  degrees at the maximum of activity of 1999-2002.

	Active Sun	Quiet Sun	16 Cyg A		16 Cyg B				
	.	.	No bias	bias corr.	No bias	bias corr.			
$\ln(P(\mathbf{O} \mathbf{M}_{CF})$	-10.809	-0.367	-1.856	-3.889	-10.596	-1.228	-5.910	-2.951	
$\Pi(\theta_0, \delta)$	$\epsilon_{nl}$ ( $10^{-4}$ )	$7.6^{+25.5}_{-5.3}$	$3.7^{+19.6}_{-3.1}$	$5.3^{+19.6}_{-4.1}$	$6.4^{+24.8}_{-4.8}$	$12.5^{+31.9}_{-7.5}$	$4.6^{+20.8}_{-3.7}$	$11.5^{+39.4}_{-7.9}$	$5.1^{+23.8}_{-3.9}$
	$\theta_0$ (deg)	$76^{+8}_{-7}$	$60^{+19}_{-41}$	$71^{+13}_{-14}$	$79^{+8}_{-10}$	$58^{+3}_{-3}$	$67^{+14}_{-25}$	$61^{+3}_{-4}$	$78^{+8}_{-12}$
	$\delta$ (deg)	$7^{+16}_{-5}$	$2^{+11}_{-2}$	$4^{+16}_{-3}$	$5^{+15}_{-4}$	$8^{+17}_{-6}$	$3^{+13}_{-2}$	$6^{+14.8}_{-4.6}$	$4^{+15}_{-3}$
$\Pi(\theta_0, \delta)$	$\ln(P(\mathbf{O} \mathbf{M}_{AR})$	-0.685	-0.347	-0.482	-0.656	-0.878	-0.449	-0.781	-0.611
Significance (%)		> 99.99	50.5	79.8	96.2	> 99.9	68.6	99.3	91.2
$\Lambda(\theta_0, \delta)$	$\epsilon_{nl}$ ( $10^{-4}$ )	$14.2^{+33.0}_{-9.8}$	$17^{+42.9}_{-14.2}$	$15.4^{+39.5}_{-12.2}$	$13.6^{+34.1}_{-4.8}$	$21.1^{+35.2}_{-12.6}$	$15.1^{+38.9}_{-12.4}$	$17.3^{+34.0}_{-11.4}$	$13.3^{+36.5}_{-10.3}$
	$\theta_0$ (deg)	$76^{+9}_{-7}$	$49^{+27}_{-36}$	$66^{+16}_{-41}$	$78^{+9}_{-21}$	$58^{+2}_{-3}$	$58^{+21}_{-40}$	$61^{+3}_{-4}$	$75^{+11}_{-43}$
	$\delta$ (deg)	$7^{+16}_{-5}$	$6^{+19}_{-5}$	$6^{+17}_{-4}$	$5^{+15}_{-4}$	$10^{+16}_{-6}$	$6^{+17}_{-5}$	$8^{+16}_{-5}$	$5^{+16}_{-4}$
$\Lambda(\theta_0, \delta)$	$\ln(P(\mathbf{O} \mathbf{M}_{AR})$	-0.698	-0.296	-0.492	-0.699	-0.930	-0.427	-0.847	-0.639
Significance (%)		> 99.99	51.8	79.6	96.0	> 99.99	69.0	99.4	91.0
$\mathcal{N}(\theta_0, \delta)$	$\epsilon_{nl}$ ( $10^{-4}$ )	$3.7^{+16.0}_{-2.2}$	$15.1^{+42.8}_{-13.5}$	$12.2^{+39.5}_{-10.5}$	$6.6^{+24.3}_{-5.0}$	$13.3^{+24.1}_{-7.0}$	$12.6^{+41.0}_{-11.0}$	$6.6^{+24.3}_{-5}$	$8.9^{+35.4}_{-7.5}$
	$\theta_0$ (deg)	$73^{+11}_{-10}$	$48^{+28}_{-34}$	$63^{+18}_{-38}$	$75^{+11}_{-27}$	$56.6^{+3}_{-7}$	$55^{+23}_{-37}$	$75^{+11}_{-27}$	$72^{+14}_{-42}$
	$\delta$ (deg)	$9^{+26.6}_{-8}$	$7^{+20}_{-6}$	$6^{+22}_{-5}$	$4^{+20}_{-3}$	$6^{+22}_{-5}$	$7^{+21}_{-6}$	$4^{+20}_{-3}$	$4^{+19}_{-3}$
$\mathcal{N}(\theta_0, \delta)$	$\ln(P(\mathbf{O} \mathbf{M}_{AR})$	-0.708	-0.384	-0.584	-0.811	-0.943	-0.516	-0.811	-0.759
Significance (%)		> 99.99	49.6	78.1	95.6	> 99.99	67.1	99.4	90.0
Average	$\epsilon_{nl}$ ( $10^{-4}$ )	$7.7^{+25.7}_{-5.6}$	$11.4^{+41.6}_{-10.0}$	$9.2^{+34.2}_{-7.5}$	$8.0^{+29.4}_{-6}$	$15.5^{+30.7}_{-9.1}$	$10.5^{+37.2}_{-9.0}$	$11.8^{+31.5}_{-7.5}$	$9.1^{+34.7}_{-7.4}$
	$\theta_0$ (deg)	$75^{+10}_{-8}$	$53^{+25}_{-37}$	$68^{+14}_{-36}$	$78^{+9}_{-14}$	$58^{+3}_{-4}$	$61^{+18}_{-34}$	$60^{+4}_{-6}$	$75^{+11}_{-38}$
	$\delta$ (deg)	$7^{+22}_{-6}$	$5^{+18}_{-4}$	$5^{+18}_{-4}$	$5^{+17}_{-4}$	$8^{+18}_{-6}$	$6^{+18}_{-5}$	$7^{+19}_{-6}$	$4^{+17}_{-3}$
Average	$\ln(P(\mathbf{O} \mathbf{M}_{AR})$	-0.697	-0.342	-0.519	-0.722	-0.917	-0.464	-0.824	-0.670
Significance (%)		> 99.99	50.6	79.2	96.0	> 99.99	68.2	99.4	90.7

**Table 3.** Statistical summary for the activity parameters and their model log-marginal likelihood  $\ln(P(\mathbf{O}|\mathbf{M}_{AR})$  when  $F(\theta|\mathbf{x}) = \Pi(\theta_0, \delta), \Lambda(\theta_0, \delta)$  or  $\mathcal{N}(\theta_0, \delta)$ .  $\ln(P(\mathbf{O}|\mathbf{M}_{CF})$  is the log-marginal likelihood for a pure centrifugal effect. The Activity significance is the probability that the activity is necessary to explain the data and derived from the log-marginal likelihoods. Uncertainty on the activity significance is less than 0.25%.

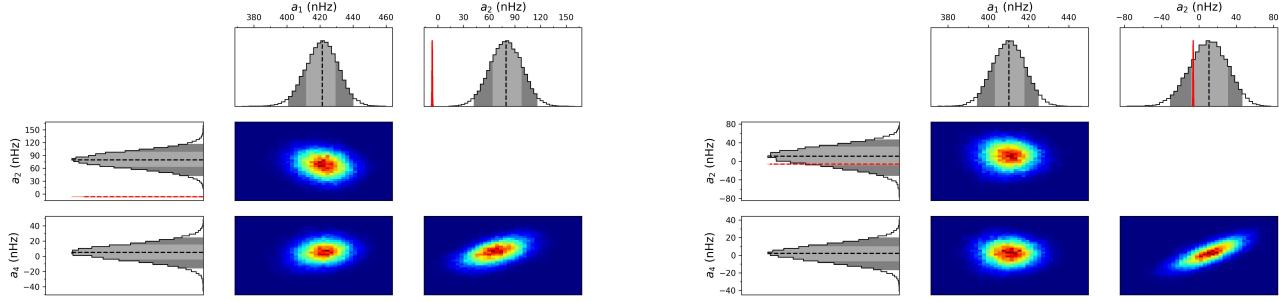
data with a different methodology, and over period that encompasses the maximum of 1999-2002. Their measure consider only  $T_{n22}$  so a direct comparison is not straightforward. However, we note that with a 844 day-long timeseries starting in Feb 1999<sup>7</sup>, they detect a frequency shift  $T_{n22} \simeq 190$  nHz at a similar significance ( $\simeq 4\sigma$ ) than us when averaging over all modes between  $\simeq 2000 - 3300$   $\mu$ Hz (to be compared to our range of  $\simeq 2300 - 3600$   $\mu$ Hz). Their Figure 8 also show that the global effect of the activity cycle between 1994 and 2000 is evident on the averaged  $T_{n22}$  while the frequency dependence of  $T_{n22}$  may have too large uncertainties to ascertain a frequency-trend. This is in line with our own findings (see our Section 3.3).

A rigorous statistical evaluation of our significance requires the joint use of  $a_2$  and  $a_4$ . This is discussed in Section 5.2, along with other activity results. Finally note that for inclination of  $\simeq 90$  degrees, it is not possible to determine  $a_3$ , because the amplitudes of the azimuthal components for  $l = 1, 2$ , are not favorable (Gizon & Solanki 2004).

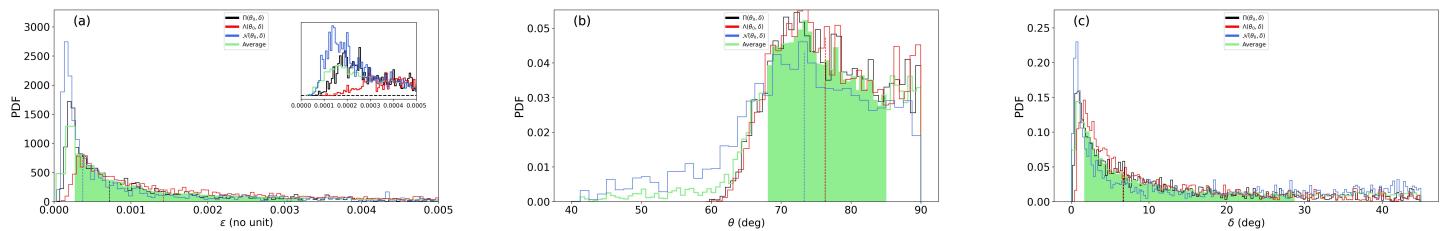
## 5.2. Activity of the Sun

The activity intensity and its latitudinal coverage is derived from a fit of the  $a_2^{(AR)}$  and  $a_4^{(AR)}$  coefficients. Technical details are in Appendix C. Statistical summary of the results are listed in Table 3. The statistical significance of the activity exceeds 99.99% (highly significant) at the maximum of activity, but is around 50.6% (not significant) during the minimum of activity,

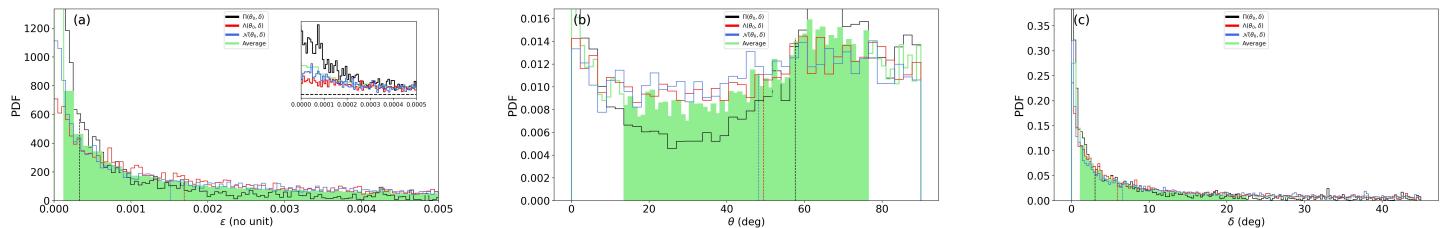
<sup>7</sup> To be compared to our timeserie of 1095 days starting in Jan 1999.



**Figure 7.** Probability Density Functions and their correlations, obtained by MCMC for coefficients  $a_1, a_2, a_3, a_4$  for the Sun at Maximum of activity (1999-2002, left) and at minimum of activity (2006-2009, right). The red curve is the expected  $a_2^{CF}$  coefficient for a pure centrifugal distortion. The light and dark gray PDF filling is for the  $1\sigma$  and  $2\sigma$  confidence interval, respectively.



**Figure 8.** Inferred pdf for (a)  $\epsilon$ , (b)  $\theta_0$  and (c)  $\delta$ , during the Sun maximum activity of 1999-2002. The inset of (a) is a zoom into the near-zero  $\epsilon$  values with smaller binning.



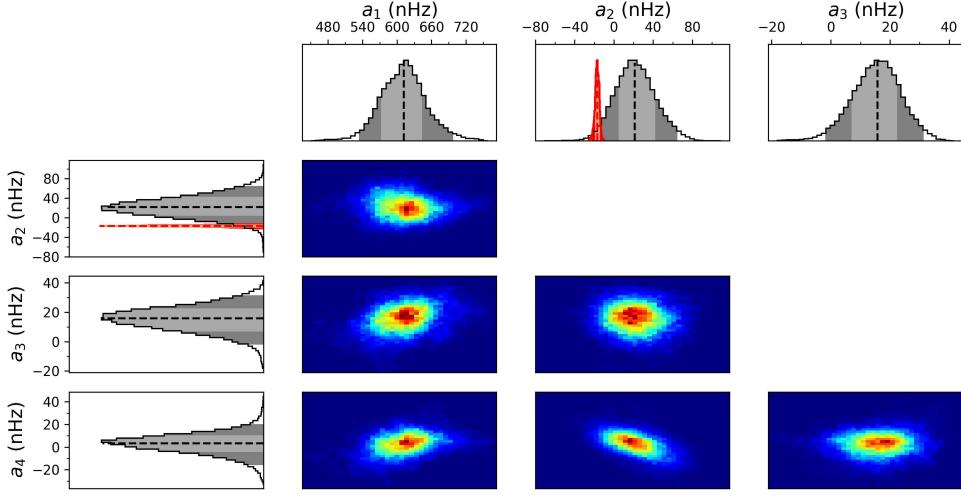
**Figure 9.** Inferred pdf for (a)  $\epsilon$ , (b)  $\theta_0$  and (c)  $\delta$ , during the Sun minimum activity of 2006-2009. The inset of (a) is a zoom into the near-zero  $\epsilon$  values with smaller binning.

demonstrating the possibility of detecting activity in Sun-like stars. **The choice of the function  $F$  describing the active region changes the detection significance by only a few percents. The three explored models fit equivalently the data implying that the shape of the active region cannot be determined with currently available  $a_2$  and  $a_4$  constraints.**

Figure 8 and 9 show  $\epsilon_{nl} = \epsilon$ ,  $\theta_0$  and  $\delta$  for the Sun at its maximum of activity (1999-2002) and its minimum (2006-2009), respectively. **There is no major differences between the three activity profiles, although it is noted that uncertainties are larger in the case of a triangular and a Gaussian activity zone.** During the maximum of activity,  $\epsilon \simeq 0$  is clearly excluded (see inset). **With  $F = \Pi$ , we observe a log-normal distribution, with median  $7.6 \cdot 10^{-4}$ , consistent with the value reported by Gizon (2002).** However, the uncertainty is large, **suggesting that only the order of magnitude can be constrained in Sun-like stars.** At the minimum of activity, the log-normal distribution morphs into a  $1/x$  law, similar to the Jeffreys prior. This is a sign of weaker statistical significance for the activity.

The co-latitude  $\theta_0 \simeq 75^\circ$  is consistent with the butterfly diagram, which suggests  $\theta_0 \simeq 80^\circ$ . At the minimum of activity and despite a non-significant detection, the probability distribution of  $\theta_0$  shows a weak indication of activity at mid and high co-latitudes. Despite the high significance of the detection during the maximum of activity,  $\delta$  is poorly constrained, showing that this parameter is challenging to measure on the Sun and on other Sun-like stars.

## 6. ANALYSIS OF 16 CYB A AND B



**Figure 10.** Probability Density Functions and their correlations, obtained by MCMC of the power spectrum of 16 Cyg A and for coefficients  $a_1$ ,  $a_2$ ,  $a_3$ ,  $a_4$  for 16 Cyg A. The red curve shows the expected  $a_2^{CF}$  coefficient for a pure centrifugal distortion of the star. The light and dark gray PDF filling is for the  $1\sigma$  and  $2\sigma$  confidence interval, respectively.

Due to their brightness (magnitudes  $V=5.95$  and  $6.20$ ) 16 Cyg A and B have modes with the highest HNR among all of Sun-like stars observed asteroseismically so far. They constitute ideal candidates to evaluate the activity. The two stars are wide binaries, with a confirmed planet around 16 Cyg B (Cochran et al. 1997) and were extensively studied (e.g. Neckel 1986; King et al. 1997; Deliyannis et al. 2000; Schuler et al. 2011; Takeda 2005; Metcalfe et al. 2012; Lund et al. 2014b; Verma et al. 2014; Metcalfe et al. 2016; Roxburgh 2017; Bellinger et al. 2017; Maia et al. 2019; Bazot et al. 2019; Bazot 2020; Nsamba et al. 2022). The 2.5 years (13 September 2010 to 8 March 2013) observation by the Kepler space-borne instrument is used in this section to measure pulsation parameters. The data are the same as those used in Bazot et al. (2019) for which instrumental issues (outliers, jumps, trends) and the quarter stitching is performed using the procedure described in García et al. (2011b). The binary system has precisely measured angular diameters, making it two of the few Sun-like stars with known interferometric radii. Their measured radius are  $1.22 \pm 0.02 R_\odot$  and  $1.12 \pm 0.02 R_\odot$  for 16 Cyg A and B, respectively (White et al. 2013). Their spectroscopic parameters are very close to those of the Sun: Their effective temperatures are  $T_{\text{eff}} = 5825 \pm 50$  K,  $T_{\text{eff}} = 5750 \pm 50$  K and their metallicity are  $[M/H] = 0.10 \pm 0.09$ ,  $[M/H] = 0.05 \pm 0.06$  (Ramírez et al. 2009) for 16 Cyg A and B, respectively. With an estimated age of around 7 Gyrs for both stars (Metcalfe et al. 2016; Bazot 2020, e.g.), they are significantly older than the Sun.

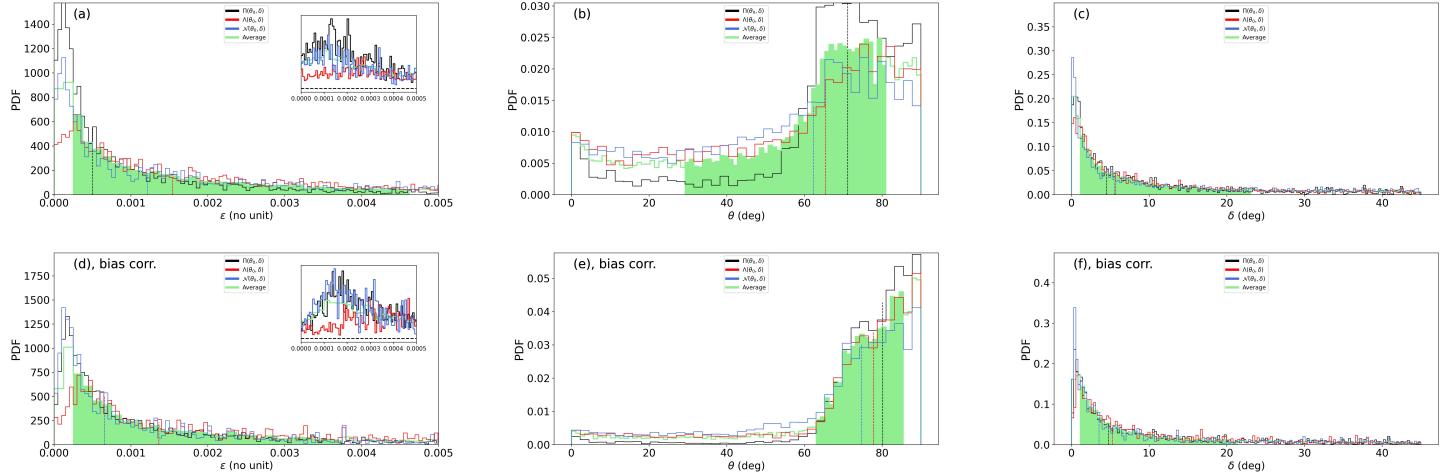
### 6.1. Seismic constraints for 16 Cyg A

Earlier studies of 16 Cyg A revealed around 60 modes of pulsations with significance in the power spectrum<sup>8</sup>. The asteroseismic analysis of Davies et al. (2015) found a stellar inclination of  $56^{+6}_{-5}$  and a rotation  $23.8^{+1.5}_{-1.8}$  ( $\langle \delta\nu_{nlm}/m \rangle_{nl} \approx 486^{+40}_{-29}$  nHz). Using a refined power spectrum modeling that account for  $a_1$ ,  $a_3$  and parametrises the cavity asphericity  $(R_{\text{eq}} - R_{\text{pol}})/R_{\text{eq}}$ , Bazot et al. (2019) reported values ( $i = 58.5 \pm 6.8$ ,  $a_1 = 464 \pm 43$  nHz), consistent with Davies et al. (2015).

Figure 10 shows the probability density function for  $a_1$ ,  $a_2$ ,  $a_3$  and  $a_4$  for 16 Cyg A, along with their correlation and with  $a_2^{CF}$  superimposed in the  $a_2$  quadrant. Table 2 synthetizes the inferred values of the coefficients. The probability density functions are near-Gaussian and thus the quadratic mean of the asymmetrical uncertainties is reported in the table. The  $a_1$  coefficient is significantly lower than past estimates but remain consistent at a  $2\sigma$  confidence level. The difference may be due to the lower stellar inclination  $i = 45 \pm 4$ , that is also only consistent with earlier determination at  $2\sigma$ . The  $a_3$  coefficient is marginally greater but with smaller uncertainty than Bazot et al. (2019) (they reported  $a_3 = 11.15 \pm 10.95$  nHz).

More importantly,  $a_2$  is positive at  $1\sigma$ , which is consistent with the assertion of star prolateness from Bazot et al. (2019) and is inconsistent with the centrifugal distortion term,  $a_2^{CF}$ . Finally,  $a_4$ , includes zero within  $1\sigma$ .

<sup>8</sup> The exact value may differ from author to author, due to different choice for the significance.



**Figure 11.** Inferred pdf for 16 Cyg A for raw measures (a,b,c) or with bias correction (d,e,f). The green curve is for the average pdf with  $F = \Pi, \Lambda$  and  $\mathcal{N}$ . The shaded area is its  $1\sigma$  confidence interval. The inset of (a) and (d) is a zoom into the near-zero  $\epsilon$  values with smaller binning.

With respect to Figures 4–5,  $a_1/\Gamma_{v_{max}} = 0.45 \pm 0.04$  and  $i = 45 \pm 4^\circ$  place 16 Cyg A in a parameter space where the expected magnitude of the bias for  $a_1$ ,  $a_2$  and  $a_4$  is of the order of  $0.5\sigma$ ,  $0.5\sigma$  and  $0.8\sigma$  respectively. This remains accurate even when accounting for the slightly higher estimates of inclination from previous publications. Translated into absolute units, this corresponds to approximately  $b(a_1) \simeq +10\text{nHz}$  (overestimation),  $b(a_2) \simeq -10\text{nHz}$  (underestimation),  $b(a_4) \simeq -10\text{nHz}$  (underestimation). These fiducial values are used in Section 6.3 to evaluate the effect of the potential bias on the estimates of the activity zone.

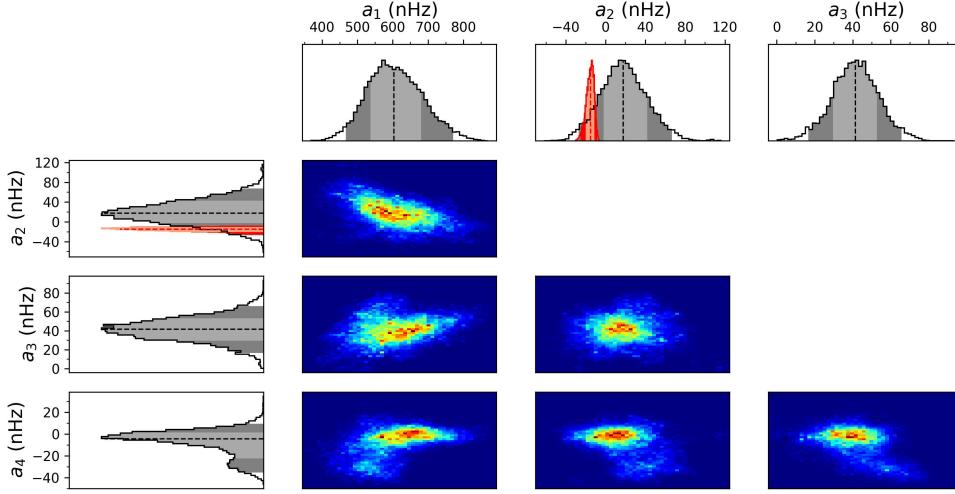
## 6.2. Seismic constraints for 16 Cyg B

Similary to 16 Cyg A, 16 Cyg B has around 60 observed modes. However, the reported precision for the past determination of the seismic parameters is less accurate than for 16 Cyg A, despite a similar HNR. Davies et al. (2015) and Bazot et al. (2019) both note a large degeneracy between the stellar inclination and the average rotation rate, with even a clear bi-modality in the distributions obtained by Bazot et al. (2019). Their global solution of  $i = 36^{+17}_{-7}$  is associated to two separate solutions for  $a_1$ , centered around  $\simeq 300\text{nHz}$  and  $\simeq 550\text{nHz}$ , that they use to infer the latitudinal differential rotation profile. Meanwhile, as for 16 Cyg A, the star is found to be prolate, indicating a surface activity.

Figure 12 shows the measured a-coefficients. It appears that including the  $a_4$  coefficient remove the degeneracy issue observed by previous publications. The stellar inclination,  $a_1$  and  $a_2$  are precisely determined and have approximately Gaussian probability distributions and these are reported as such in Table 2. The found rotation rate  $a_1$  correspond to the higher solution of rotational splitting of Davies et al. (2015) and Bazot et al. (2019) (see their Figures of probability density functions). Meanwhile, the degeneracy observed in  $a_1$  in earlier studies is moved to  $a_4$ : It exhibits two solutions. As our method of activity inference assumes Gaussian distributions, the bi-modality of  $a_4$  requires to separate the two apparent solutions. Using Gaussian mixture modeling (Bishop 1995; Bazot et al. 2019), the mean and standard deviation of both solutions are measured, enabling their separate analysis. The lowest estimate ( $a_4 = -27.9 \pm 6.5\text{nHz}$ ) weights 30% and the highest estimate ( $a_4 = -1.0 \pm 6.7\text{nHz}$ ) is more significative as it weights 70%. The inference of the activity from the two solutions is discussed in Section 6.4. Note that  $a_3 = 45^{+13}_{-14}\text{nHz}$  is significantly higher than the reported values from Bazot et al. (2019) ( $a_3 = 13.89 \pm 13.95\text{nHz}$ ) and may have an impact on the rotational profile.

## 6.3. Activity inference for 16 Cyg A

Figure 11 shows the results from the inference of the activity parameters and a statistical summary is in Table 3. **Triangular and Gaussian descriptions of active latitudes give larger uncertainties than the simple gate model. However, they all suggest a near-equatorial activity, with a similar detection significance level of  $\simeq 79\%$  (without bias correction) or  $\simeq 96\%$  with fiducial correction. As for the Sun and because the shape of the active region is *a priori* unknown, the average distribution of the parameters are discussed.** The activity intensity has a large uncertainty, but according to the median of  $\epsilon$ , may be bettwen the maximum and the minimum of solar activity.



**Figure 12.** Probability Density Functions and their correlations, obtained by MCMC of the power spectrum of 16 Cyg B and for coefficients  $a_1, a_2, a_3, a_4$  for 16 Cyg A. The red curve shows the expected  $a_2^{CF}$  coefficient for a pure centrifugal distortion of the star. The light and dark gray PDF filling is for the  $1\sigma$  and  $2\sigma$  confidence interval, respectively.

The probability distribution of  $\theta_0 \approx 68^\circ$  ( $\theta_0 \approx 78^\circ$  with fiducial bias correction) indicates that the active region is likely near the equator and below  $45^\circ$  of latitude. Accounting for the bias leads to an activity closer to the equator. The precision of the inference is unfortunately insufficient to know whether the activity is concentrated at mild-latitudes ( $\approx 30^\circ$  as for the Sun at maximum of activity) or if the activity is closer to the equator.

The posterior probability distribution of  $\delta$  does not allow us to precisely constrain the extension of the activity region. As already noted in the case of the Sun, this parameter requires stringent constraint on both the  $a_2$  and  $a_4$  coefficient in order to inform us about the size of the active region.

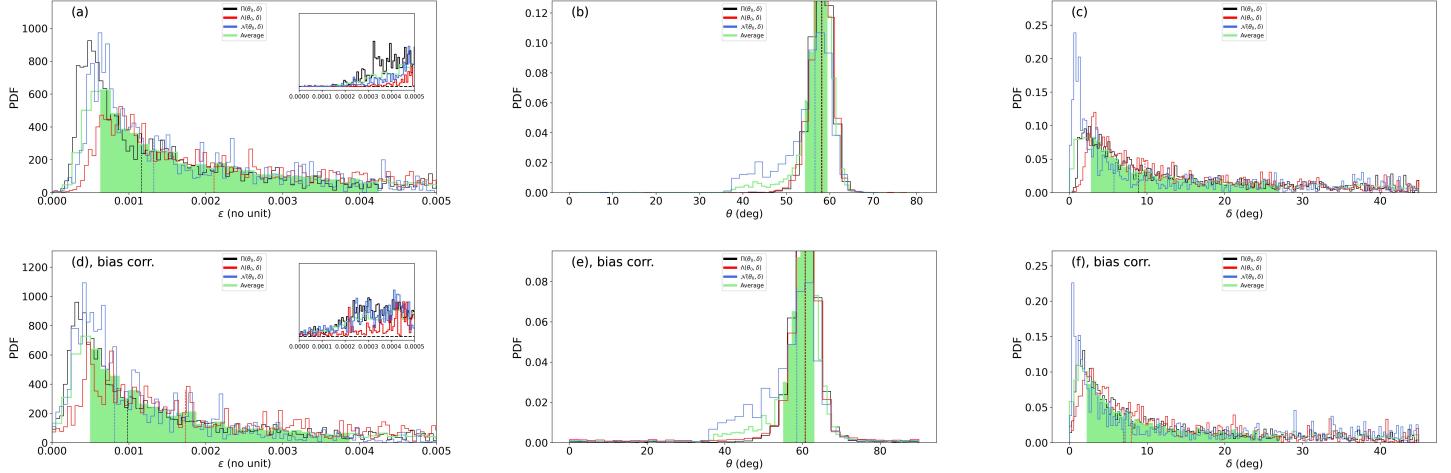
#### 6.4. Activity inference for 16 Cyg B

Figure 13 and 14 shows the results for the activity for 16 Cyg B in the two possible scenarios of  $a_4$  discussed in Section 6.2 (see Table 3 for the statistical summary). In the case of  $a_4 = -27.9 \pm 6.5\text{ nHz}$  (Figure 13), the activity is highly significant (greater than 99.3% with or without bias correction, **whatever is the activity zone model**) and has stronger intensity than in the case of the Sun. The active region is then located at  $\theta_0 \approx 58^\circ$  ( $\theta_0 \approx 60^\circ$ , after bias correction), ie. at comparable latitudes seen during a maximum of activity of the Sun. The bias correction has negligible effect on the inferred activity latitude. Due to  $a_4$  significantly departing from 0, the extension of active region is better constrained than in 16 Cyg A or the Sun, but remains weakly informative.

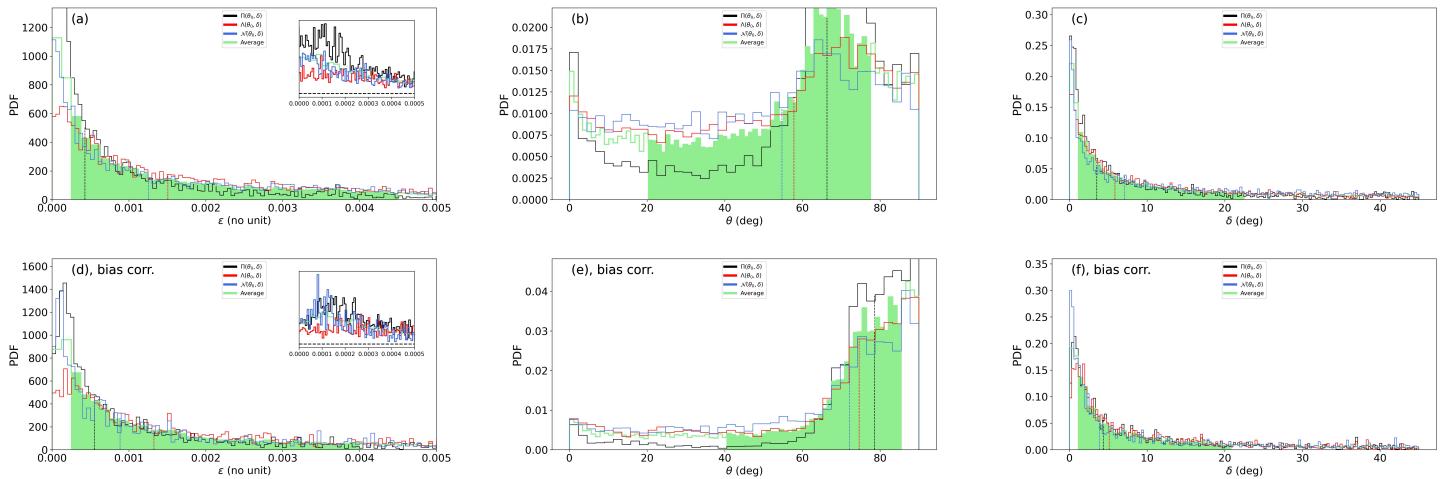
The second more likely solution (probability of 70%) corresponds to  $a_4 = -1.0 \pm 6.7\text{ nHz}$  (Figure 14). It is associated to an activity at co-latitudes above  $\approx 40^\circ$  and of lower activity. Its statistical significance is low, and the activity intensity may be of the same order or lower than 16 Cyg A. **In fact, it looks similar to the solar case when approaching its minimum of activity.** Accounting for the fiducial bias, the solution is more concentrated to the equatorial region and differs significantly from the lower probability  $a_4$  solution. Here, the uncertainty on  $a_2$  and  $a_4$  is again too large to provide a stringent constraint on the extension of the activity zone  $\delta$ .

## 7. DISCUSSION AND CONCLUSION

The stellar activity is a complex phenomenon emmerging from the interplay between the stellar plasma, rotation and magnetism. The activity distorts the shape of mode cavity, perturbing the pulsation frequencies. This perturbation is here observationally evaluated using a-a-coefficient decomposition. This allows us to separate observables (the a-coefficients) from their physical interpretation (the activity modeling). The separation enables us to demonstrate that the measurements of the average low a-coefficient,  $a_2$  and  $a_4$ , under some observational conditions, are sufficient to reveal the presence of a statistically significant activity of similar intensity to the Sun and to determine its latitude. The required observational conditions are analysed using a methodology similar to Kamiaka et al. (2018), that is, by constructing a grid of artificial power spectra, that allows us to determine the bias for  $a_1, a_2$  and  $a_4$ . It is found that if the height-noise-ratio exceeds 10, the mode blending factor  $f = a_1/\Gamma_{v_{max}}$  is greater



**Figure 13.** Inferred pdf of  $\epsilon$ ,  $\theta_0$  and  $\delta$  for raw results of 16 Cyg B with the lower solution  $a_4 = -29.2 \pm 5.6\text{nHz}$ , without bias correction (a, b, c) or with it (d, e, f). The green curve is for the average pdf with  $F = \Pi$ ,  $\Lambda$  and  $\mathcal{N}$ . The shaded area is its  $1\sigma$  confidence interval. The inset of (a,d) is a zoom into the near-zero  $\epsilon$  values with smaller binning.

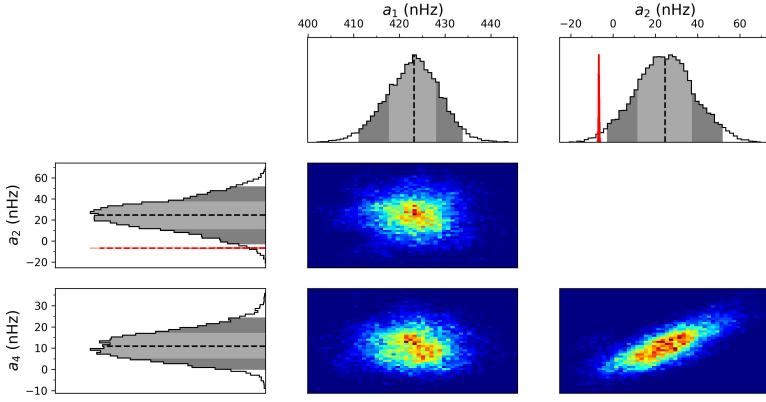


**Figure 14.** Same as Figure 13 but for the higher solution  $a_4 = -0.9 \pm 6.8\text{nHz}$ .

than 0.4, the inclination is above  $30^\circ$  and that the observation is longer than 2 years, the inaccuracy remains mild and generally smaller than the  $1\sigma$  uncertainty. The uncertainty and/or the inaccuracy may however become too large to reliably detect any activity beyond the above-specified conditions. In particular and in agreement with Kamiaka et al. (2018), the stellar inclination is a decisive variable to ensure the accuracy of the measurement. Below  $30^\circ$ ,  $a_1$  is often wrong by a factor 2. This is likely due to the fit mistakenly identifying  $l = 2, m = \pm 1$  as  $l = 2, m = \pm 2$  and vice-versa. Other a-coefficients are then severely inaccurate. Therefore, such a bias analysis suggests that rotation studies of stellar ensembles require a careful star selection.

A method that uses the average  $a_2$  and  $a_4$  coefficients is proposed to perform a subsequent analysis of the activity, considering a geometrical model of the activity effect on the pulsations frequencies and accounting for the stellar asphericity due to the centrifugal effects. This two-step approach is tested on the case of the Sun and shows that it is effectively able to detect the change of activity between the solar maximum of activity around 1999-2002 and the minimum of activity around 2006-2009. Although the use of averaged a-coefficients makes it difficult to evaluate the extension of the activity zone, the model successfully retrieves the mean latitude of activity during the maximum of the solar cycle.

The method is then applied to the brightest stars observed during the initial observational phase of the Kepler space instrument, 16 Cyg A and B. These stars were selected as a test-bed due to the fact that they are well studied and present the highest mode signal-to-noise ratio of all the currently known main-sequence stars. Davies et al. (2015) suggested that these stars have mild



**Figure 15.** Probability Density Functions and their correlations, obtained by MCMC for coefficients  $a_1, a_2, a_3, a_4$  for the Sun in-between two activity cycles (2006-2011). The red curve is the expected  $a_2^{CF}$  coefficient for a pure centrifugal distortion. The light and dark gray PDF filling is for the  $1\sigma$  and  $2\sigma$  confidence interval, respectively.

to no activity. However, [Bazot et al. \(2019\)](#), using a parametric model for describing the asphericity  $(R_{\text{eq}} - R_{\text{pol}})/R_{\text{eq}}$  found an asphericity significant at  $1\sigma$ . Our current analysis, performed using the same data set, confirms this asphericity and found a mild (relatively to the Sun) to moderate activity for both stars. It is found that 16 Cyg A has a near-equatorial band of activity during the period of observation (13 September 2010 to 8 March 2013), with a significance of the detection greater than 79.8%.

The case of 16 Cyg B is more ambiguous. A bi-modality on the average splitting  $\langle \delta\nu_{nlm}/m \rangle \simeq a_1$  and on the stellar inclination is already reported in [Davies et al. \(2015\)](#) and [Bazot et al. \(2019\)](#). Our refined model suggests that this bi-modality is in fact related to the  $l = 2$  modes. Indeed, as we account for  $a_1, a_2, a_3$  and  $a_4$ , we note that the bi-modality previously seen on  $a_1$  in earlier studies is displaced to  $a_4$ . The solutions of  $a_4$  are separated using a Gaussian process algorithm, which found that the weight (or importance) of the highest solution, close to 0 nHz is of 70%. The lower solution, close to -28 nHz has a weight of 30%. The associated probability distribution for stellar inclination  $i = 35 \pm 3$  is unique (instead of being bi-modal in past studies). The two solutions of  $a_4$  are analysed independently. The lower solution (with lower weight) is associated to an overall activity that is stronger than the Sun, localized at latitudes of approximately  $32^\circ$  (with an uncertainty of  $\simeq 3^\circ$ ). The higher solution is linked to a lower activity and is weakly significant (probability greater than **67.1%**). Although the uncertainty is large, the study suggests an activity closer to the equatorial region. In the Sun and as evident in its butterfly diagram, the quiet phase is associated to a magnetic activity closer to the equator, while the transition to the active Sun is abrupt and characterised by the appearance of magnetic spots at latitudes of  $30\text{--}40^\circ$ . In that context and although it is not possible to rule out the possibility of a statistical fluke, an interpretation of the bi-modality is that the star was transitioning from a period of low activity to a more active period during the observation time of the Kepler instrument. To evaluate that hypothesis, we selected solar-data between Jan 2006 and Jan 2011, including the end of a cycle and the start of a new one. The measured a-coefficients for that analysis are shown in Figure 15. There is no visible bi-modality on either  $a_2$  or  $a_4$ . Because 16 Cyg B is evidently different than the Sun, this does not refute the hypothesis, but weakens it. A more firm verification would require a follow-up observation of the star, ie with PLATO ([Rauer et al. 2014](#)). Or extensive simulations in order to attempt to reproduce the bi-modality.

This work demonstrates that it is possible to determine the latitude and intensity of the activity for Sun-like stars, when this activity is similar to, or exceed the one of the Sun. The limited observed bias suggests that the analysis of the stellar activity could also be extended to a larger set of stars. The Kepler LEGACY stars ([Lund et al. 2017](#)) are ideal candidates as these generally satisfy the criteria of reliability that are discussed in this conclusion and detailed in Section 4.

Although the a-coefficient analysis has the benefit to allow to simplify the bias studies and to make the analysis faster, a further axis of improvement would consist in using a single analysis step, ie, fitting directly the power spectrum with the model of activity. Contrary to the two-step analysis, such an approach does not require assumptions on the properties of probability distribution of the a-coefficients as it is currently the case. It would also make use of the full set of a-coefficients (not only their average), which may lead to smaller uncertainties. However, the two approaches would certainly be needed in tandem, as the stability of the solution (and its accuracy) is harder to ascertain when performing a single step fitting approach.

We thank Saskia Hekker for her comments and advices.

**All of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. The specific observations analyzed can be accessed via [10.17909/T9059R](https://doi.org/10.17909/T9059R).**

*Facilities:* SoHo Telescope, Kepler Telescope

*Software:* Analysis tool suite for this paper <https://github.com/OthmanB/Benomar2022>

## APPENDIX

### A. POLYNOMIALS, A-COEFFICIENTS AND SPLITTINGS

#### A.1. *The first six used polynomials*

The  $\mathcal{P}_j^{(l)}(m)$  coefficients, which are introduced by Schou et al. (1994), can be obtained by normalizing the  $H_j^m(l)$  coefficients by Ritzwoller & Lavelle (1991) so that  $\mathcal{P}_j^{(l)}(m = l) = l$ .

$$\mathcal{P}_j^{(l)}(m) = \frac{l(2l-j)!}{(2l)!} H_j^m(l) = (-1)^{-l+m} \frac{l \sqrt{(2l-j)!(2l+j+1)!}}{(2l)!} \begin{pmatrix} j & l & l \\ 0 & m & -m \end{pmatrix}, \quad (\text{A1})$$

in which the last factor means the Wigner  $3j$  symbol.

**They can also be obtained by recurrence using equation (2) and (3), together with the normalisation condition,  $\mathcal{P}_j^l(m) = l$  and starting with  $\mathcal{P}_0^l(m) = l$ .**

**The first six polynomials are** as follows:

$$\mathcal{P}_1^{(l)}(m) = m \quad (\text{A2})$$

$$\mathcal{P}_2^{(l)}(m) = \frac{3m^2 - l(l+1)}{2l-1} \quad (\text{A3})$$

$$\mathcal{P}_3^{(l)}(m) = \frac{5m^3 - (3l(l+1)-1)m}{(l-1)(2l-1)} \quad (\text{A4})$$

$$\mathcal{P}_4^{(l)}(m) = \frac{(35m^4 - 5(6l(l+1)-5)m^2) + 3l(l+1)(l(l+1)-2)}{2(l-1)(2l-1)(2l-3)} \quad (\text{A5})$$

$$\mathcal{P}_5^{(l)}(m) = \frac{252m^5 - 140(2L-3)m^3 + (20L(3L-10)+48)m}{8(4l^4 - 20l^3 + 35l^2 - 25l + 6)} \quad (\text{A6})$$

$$\mathcal{P}_6^{(l)}(m) = \frac{924m^6 - 420m^4(3L-7) + 84m^2(5L^2 - 25L + 14) - 20L(L^2 - 8L + 12)}{64l^5 - 480l^4 + 1360l^3 - 1800l^2 + 1096l - 240} \quad (\text{A7})$$

where  $L = l(l+1)$ .

#### A.2. *Relationship between a-coefficients and splittings*

The relationships between a-coefficients and symmetric/anti-symmetric splittings can be derived using equations (4) and (5). These are used to convert model frequencies  $\nu_{nlm}$  into a-coefficients.

##### A.2.1. *Explicit form for low aj-coefficients and l=1*

$$a_1(n, l = 1) = S_{n11} = \frac{\nu_{n,1,1} - \nu_{n,1,-1}}{2} \quad (\text{A8})$$

$$a_2(n, l = 1) = T_{n11}/3 = \frac{(\nu_{n,l,-1} + \nu_{n,l,1})/2 - \nu_{n,l,0}}{3} \quad (\text{A9})$$

595 A.2.2. *Explicit form for low aj-coefficients and l=2*

596 
$$a_1(n, l = 2) = \frac{S_{n21} + 4S_{n22}}{5} \quad (\text{A10})$$

596  
597  
598  
599  
600  
601  
602

598 
$$a_2(n, l = 2) = \frac{2T_{n22} - T_{n21}}{7} \quad (\text{A11})$$

598  
599  
600  
601  
602

599 
$$a_3(n, l = 2) = \frac{S_{n22} - S_{n21}}{5} \quad (\text{A12})$$

599  
600  
601  
602

600 
$$a_4(n, l = 2) = \frac{T_{n22} - 4T_{n21}}{70} \quad (\text{A13})$$

600  
601  
602

603 A.2.3. *Explicit form for low aj-coefficients and l=3*

604 
$$a_1(n, l = 3) = \frac{S_{n31} + 4S_{n32} + 9S_{n33}}{14} \quad (\text{A14})$$

604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614

605 
$$a_2(n, l = 3) = \frac{-15T_{n31} + 25T_{n33}}{126} \quad (\text{A15})$$

605  
606  
607  
608  
609  
610  
611  
612  
613  
614

606 
$$a_3(n, l = 3) = \frac{-S_{n31} - 2S_{n32} + 3S_{n33}}{9} \quad (\text{A16})$$

606  
607  
608  
609  
610  
611  
612  
613  
614

607 
$$a_4(n, l = 3) = \frac{T_{n31} - 7T_{n32} + 3T_{n33}}{77} \quad (\text{A17})$$

607  
608  
609  
610  
611  
612  
613  
614

608 
$$a_5(n, l = 3) = \frac{5S_{n31} - 8S_{n32} + 3S_{n33}}{126} \quad (\text{A18})$$

608  
609  
610  
611  
612  
613  
614

609 
$$a_6(n, l = 3) = \frac{15T_{n31} - 6T_{n32} + T_{n33}}{1386} \quad (\text{A19})$$

609  
610  
611  
612  
613  
614

## 615 B. SPECTRUM ANALYSIS

616 B.1. *Acoustic-spectrum modelling*

617 The analysis of the asteroseismic data is very often performed by fitting the power spectrum (eg. Appourchaux et al. 1998,  
 618 2008). This requires to devise a likelihood function and a model for the observed spectrum. As our fitting involves a Bayesian  
 619 approach, the priors are also required.

620 B.1.1. *Likelihood*

621 The likelihood function is determined by the noise statistics of the power spectrum, which is a  $\chi^2$ -squared with two degree  
 622 of freedom. It is also based on the assumption that the frequency bins are independent and uncorrelated. This implies that the  
 623 observation duration is assumed to be much greater than the typical lifetime of the pulsation modes and that the duty cycle is  
 624 sufficient to allow us to neglect any leakage induced by the window function. The observation timeframe for the LEGACY sample  
 625 is of the order of years, which is significantly longer than the mode lifetime. Furthermore, the duty cycle is above 95%, which  
 626 was shown to be sufficient to neglect leakage (Stahn 2010). Under these conditions, the likelihood function is (Anderson et al.  
 627 1990),

628 
$$L_1(S(\nu_i)|X, M) = \sum_{i=0}^{N-1} \frac{1}{S(\nu_i)} \exp\left(-\frac{S(\nu_i)}{M(\nu_i, X)}\right) \quad (\text{B20})$$

628

629 where  $S(\nu_i)$  is the power at the central frequency  $\nu_i$  of the  $i^{th}$  bin. Here,  $N$  is the total number of bins and  $M(\nu_i, X)$  is the model  
 630 of the power spectrum with the set of variable  $X$  for the model  $M$ .

631 From Bayes theorem, the posterior probability density function is defined as,

632 
$$\pi_1(X|S(\nu), M) = \frac{\pi_1(X|M) L_1(S(\nu)|X, M)}{\pi_1(S(\nu)|M)} \quad (\text{B21})$$

632

633 where  $\pi_1(X|M)$  is the prior knowledge on the parameters  $X$ .  $\pi_1(S(\nu)|M)$  is a normalization constant, essential only for comparing  
 634 the significance of models (eg. Gregory 2005; Benomar et al. 2009). Variables are assumed to be independent from each

others. This implies that the joint prior is the product of individual priors. The posterior is evaluated using the Tempered Adaptive MCMC code<sup>9</sup> described in Atchadé (2006) and implemented by Benomar (2008).

### B.1.2. Acoustic spectrum model

The model used for the power spectrum fitting, involves a sum of asymmetrical Lorentzian, superimposed to a monotonically decreasing function of frequency (pink noise). The asymmetric Lorentzian is commonly used in helioseismology (Duvall et al. 1993; Nigam & Kosovichev 1998; Georgobiani et al. 2000; Toutain et al. 1998) and in asteroseismology for Sun-like stars (Benomar et al. 2018b),

$$M(\nu) = \sum_{n=n_0}^{n_{\max}} H_{nlm} \frac{(1 + B_{nlm} z)^2 + B_{nlm}^2}{1 + z^2} + N(\nu), \quad (\text{B22})$$

with  $z = 2(\nu - \nu_{nlm})/\Gamma_{nlm}$ . Each asymmetrical Lorentzian is defined by a height  $H_{nlm}$ , width  $\Gamma_{nlm}$ , a central frequency  $\nu_{nlm}$  and an asymmetry  $B_{nlm}$ . As explained by Gizon (2006); Benomar et al. (2018b), the asymmetry coefficient depends on the mode width and frequency. The normalized asymmetry coefficient  $\chi_{nlm} = 2 \frac{B_{nlm} \nu_{nlm}}{\Gamma_{nlm}}$  is fitted instead as it is nearly constant over the range of fitted modes.

Several prescriptions exist for describing the noise background  $N(\nu)$ . Here, it is assumed to be a sum of two generalized Lorentzian, sometimes referred as Harvey-like profiles (Harvey 1985), and of a white noise,

$$N(\nu) = \frac{A_1}{1 + (\tau_1 \nu)^{p_1}} + \frac{A_2}{1 + (\tau_2 \nu)^{p_2}} + N_0. \quad (\text{B23})$$

Here  $N_0$  is the white noise and  $A_k$  is the maximum heights of the  $k^{\text{th}}$  generalized Lorentzian. The  $\tau_k$  parameter is timescale that is the inverse of the full width at half maximum of the Lorentzians, and  $p_k$  is a power exponent.

The fit is performed globally over all the statistically significant peaks visible in the power spectrum of a star. This can lead to a very large number of fitted parameters (up to a few thousands), which is practically unsuitable. A model simplification is therefore preferred, similar to eg. Appourchaux et al. (2008); Benomar et al. (2009); Campante et al. (2011); Handberg & Campante (2011) with main simplifications recalled hereafter,

- The  $m$  dependence on heights is dropped due its relationship with the stellar inclination (Gizon & Solanki 2003). This saves several hundreds of parameters.
- The relative height of the different degree  $l$  is constant across the fitted range and for a given  $l$ , hence  $V_l^2 = H_{n,l}/H_{n,l=0} = \text{const.}$ . Only the  $l = 0$  heights  $H_{n,l=0}$  are variables. The mode visibility  $V_l^2$  replaces the  $H_{n,l>0}$  as a variable.
- Given a degree  $l$ ,  $\Gamma_{n,l,m} = \Gamma_{n,l}$  is imposed. This is justified by the fact that the width depends weakly on the frequency: all splitted components are assumed to have the same width.
- Because  $\Gamma_{n,l}$  is nearly independent of the degree, it is possible to fit  $\Gamma(\nu) = \Gamma_{n,l=0}$  and interpolate it to the frequencies of the modes with degree  $l > 0$ .

These assumptions reduce the number of variables to a few tens in the case of CoRoT, Kepler or TESS observations. The frequencies of the modes follow equation (14) in Section 2.3.

### B.2. Priors

Priors are fundamental of a Bayesian method. This section explicits  $\pi_1(\mathbf{X}|M, I)$ , the prior used during the power spectrum fitting. Parameters of the vector  $\mathbf{X}$  are assumed independent to each others, such that the product rule is used to define  $\pi_1(\mathbf{X}|M, I)$ .

Main mode parameters						
Height ( $\text{ppm}^2/\mu\text{Hz}$ )	Width ( $\mu\text{Hz}$ )	Frequency ( $\mu\text{Hz}$ )	Asymmetry ( $ \chi $ )	$V_{l=1}^2$	$V_{l=2}^2$	inclination
$\mathcal{J}(1, 1000)$	$\mathcal{J}(0.1, 45)$	$\mathcal{U}(\nu_{\min}, \nu_{\max})$	$\mathcal{J}(5, 100)$	$\mathcal{G}(1.5, 0.15)$	$\mathcal{G}(0.53, 0.053)$	$\mathcal{U}(0, 90)$

**Table 4.** Mode priors used for the fit of the power spectrum. Prior parameters on frequencies are visually determined, see Section B.2.

<sup>9</sup> <https://github.com/OthmanB/Benomar2022/Programs/TAMCMC-1.81>

Heights, widths, frequencies, asymmetry and inclination use non-informative priors. These are either Jeffreys priors for scale parameters, noted  $\mathcal{J}(x_{\min}, x_{\max})$  or uniform prior, noted  $\mathcal{U}(x_{\min}, x_{\max})$ . Priors on frequencies are uniform and require a visual inspection of the power spectrum in order to assign the lower and upper bound of the prior for each mode that follow the expected pattern for main sequence Sun-like stars (equally spaced p modes) and that show an excess of power relative to the background that exceeds 80%. The excess of power is determined using a smooth spectrum for which the noise statistics is derived by Appourchaux (2003). Mode visibilities are defined by Gaussian priors (noted  $\mathcal{G}(x_0, \sigma)$ ), with mean  $x_0$  set as the solar value (see Ballot et al. 2011) and the standard deviation  $\sigma$  is 10% of the mean. Table 4 lists the type and the prior characteristic values that are used for the parameters of the modes.

The prior on  $a_1$  is uniform between 0 and 1500 nHz. For 16 Cyg A and B, an uniform fixed prior is set on  $|a_3|$  over the range [0, 100] nHz. However at each iteration of the optimisation process,  $|a_3|$  is not allowed to exceed 20% of  $a_1$ . Preventing extremely large relative value of  $a_3/a_1$ , improves the fit stability and ensure a faster convergence of the algorithm. For the Sun analysis and because it is not possible to measure  $a_3$  for  $i \approx 90^\circ$  (due to the lack of amplitude of  $l = 2, m = \pm 1$  at that inclination),  $a_3$  is fixed to 0.

The priors on  $a_2$  and  $a_4$  are also uniform. The range is defined by using the maximum range of Figure 3 (showing  $a_2^{(AR)}$  and  $a_4^{(AR)}$ ), increasing it by 50% and adding the expected centrifugal term  $a_2^{(CF)}$ , assuming  $a_1 = 400$ nHz for the Sun and  $a_1 = 600$ nHz for 16 Cyg A/B. The  $\Delta\nu$  reported in Table 2 is also used. For the same reason as to  $a_3$ ,  $a_2/a_1$  cannot exceed 50% and  $a_4/a_1$  cannot exceed 20% at each iteration step of the optimisation process.

The noise priors are obtained from a global MAP approach similar to Benomar et al. (2012). This provides the best fit values and  $1\sigma$  uncertainties that are used as priors. The noise background of the individual model analysis is described by equation (B23). During the global MAP fit, the model is made of that same background model, plus a Gaussian envelope to account for the power excess due to the modes (eg Mathur et al. 2010; Huber et al. 2011).

### C. INFERENCE OF THE ACTIVE LATITUDES

The determination of the activity is performed as a second step using as observables the marginalised posterior distribution of the average a-coefficients. This section describes the method and assumptions to determine the posterior distribution and to compute the significance of the detection.

#### C.1. Posterior distribution and likelihood

Similarly as to the power spectrum fitting described in Section B, the determination of the most likely latitudes for the activity and its significance necessitate the computation of a posterior distribution,

$$\pi_2(\mathbf{X}|\mathbf{O}, M) = \frac{\pi_2(\mathbf{X}|M) L_2(\mathbf{O}|\mathbf{X}, M)}{\pi_2(\mathbf{O}|M)}, \quad (\text{C24})$$

with  $\pi_2(\mathbf{X}|\mathbf{O}, M)$ ,  $\pi_2(\mathbf{X}|M)$ ,  $L_2(\mathbf{O}|\mathbf{X}, M)$  are the posterior distribution, the prior, the likelihood, respectively. The denominator  $\pi_2(\mathbf{O}|M)$  is the normalisation constant used for model comparison (marginal likelihood). The posterior distributions of the a-coefficients  $a_{2,o}$  and  $a_{4,o}$  obtained by power spectrum fitting are the observables contained in  $\mathbf{O}$ . The class of the model is identified by the variable  $M$ .

For simplicity, the observables are assumed to be distributed according to un-correlated gaussians functions, such that  $\mathbf{O} = \{a_{2,o}, a_{4,o}, \sigma_{2,o}, \sigma_{4,o}\}$ , where  $a_{2,o}$  and  $a_{4,o}$  are the mean for the distributions of  $a_2$  and  $a_4$  while  $\sigma_{2,o}, \sigma_{4,o}$  are the standard deviations. This leads to a log-Likelihood in the form of a  $\chi^2$ -squared,

$$\ln L_2(\mathbf{X}|\mathbf{O}, M) = -\frac{(a_{2,o} - a_{2,m}(\mathbf{X}))^2}{2\sigma_{2,o}^2} - \frac{(a_{4,o} - a_{4,m}(\mathbf{X}))^2}{2\sigma_{4,o}^2}. \quad (\text{C25})$$

In Dziembowski et al. (2000), a-coefficients are weighted using the inverse of the mode inertia (see their equation 4). This is due to the fact that the inertia account for most of the frequency-variations of the coefficients. It is essentially relevant if the uncertainties are small enough to observe a trend when modes are fitted individually. However, in the case of the Sun and with either GOLF or BiSON data, it is difficult to perceive a frequency-trend (Chaplin et al. 2003). Our own trials on VIRGO/SPM using a linear fit to describe the frequency dependence of a-coefficient did not detect a slope that is significant at more than  $1\sigma$ . A similar analysis on 16 Cyg A and B and on simulations showed evidence of large uncertainties when attempting to determine frequency-variations for stars with HNR typical of Kepler observations. In fact, only the average the a-coefficient over  $l$  and  $n$  is shown to be robustly determined (see also our discussion Section 3.3) so that inertia effects are here neglected.

The activity model  $M_{AR}$  depends on the variables  $X = \{\epsilon, \theta_0, \delta\}$  and  $a_{2,m}(X)$ ,  $a_{4,m}(X)$  denote the modeled a-coefficients. The log-likelihood is then,

$$\ln L_2(a_{2,o}, a_{4,o}, \sigma_{2,o}, \sigma_{4,o} | a_{2,m}, a_{4,m}, M_{AR}) = -\frac{(a_{2,o} - a_{2,m}(\epsilon, \theta_0, \delta))^2}{2\sigma_{2,o}^2} - \frac{(a_{4,o} - a_{4,m}(\epsilon, \theta_0, \delta))^2}{2\sigma_{4,o}^2}. \quad (\text{C26})$$

**In our case, the observables are** the average of the fitted modes coefficients,  $a_{2,o} = a_2^{(CF)} + a_2^{(AR)}$  and  $a_{4,o} = a_4^{(AR)} = 0$ . However, to conveniently propagate all errors on the parameters of  $a_2^{(CF)}$  ( $a_1, \Delta\nu, \Delta\nu_\odot, \rho_\odot, v_{nl}$ ), it is preferable to use  $a_{2,o} = a_2^{(AR)}$ , obtained by subtracting the  $a_2^{(CF)}$  from the  $a_2$  measured by power spectrum fitting. The distribution of  $\Delta\nu$  is computed by linear fitting of the group-wise ensemble of samples ( $n, l = 0$ ) for each measured frequencies  $v_{nl}$ . The term  $a_2^{(CF)}$  is finally obtained by weighted average of all independent  $a_2^{(CF)}(n, l)$  computed at the star's posterior frequencies  $v_{nl}$ . The subtraction of  $a_2^{(CF)}$  is then again performed using the samples of its posterior, enabling to construct of the posterior probability distribution function of  $a_2^{(AR)}$ , from which we deduce its mean  $a_{2,o}$  and its standard deviation  $\sigma_{2,o}$ . To evaluate the relevance of the activity, two models are considered. First, a model ( $M = M_{CF}$ ) without activity, that accounts only for the centrifugal effects and second, a model ( $M = M_{AR}$ ) with activity. Their details and the choice of the priors is described in the following sections for each of them.

### C.2. Model without activity $M_{CF}$

The model  $M_{CF}$  has no unknown variable that require minimisation, see equation (13). For the sake of the model comparison with the model  $M_{AR}$ , it is however important to determine the marginal likelihood of  $M_{CF}$ . From that perspective, it is necessary to calculate the denominator of equation (C24),  $P(\mathbf{O}|M_{CF})$  which is defined as an integral,

$$P(\mathbf{O}|M_{CF}) = \int \pi(X|M_{CF}) L_2(a_{2,o}, a_{4,o}, \sigma_{2,o}, \sigma_{4,o} | a_{2,m}, a_{4,m}, M_{CF}) dX. \quad (\text{C27})$$

An absence of activity correspond to the limit case where the a-coefficients of the activity are exactly 0. This corresponds to setting dirac priors on  $a_{2,m} = \delta_0$  and  $a_{4,m} = \delta_0$  and implies that the marginal likelihood is the local value of the likelihood at  $X = \{0, 0\}$ . The model comparison is commonly performed in log-space such that,

$$\ln P(a_{2,o}, a_{4,o}, \sigma_{2,o}, \sigma_{4,o} | M_{CF}) = -\frac{a_{2,o}^2}{2\sigma_{2,o}^2} - \frac{a_{4,o}^2}{2\sigma_{4,o}^2}. \quad (\text{C28})$$

### C.3. Model with activity $M_{AR}$

Regarding the model with activity of equation (C26), priors must then be set on the variables subject to optimisation. The  $\epsilon$  and  $\delta$  parameters are indispensable intensive parameters and the adequate non-informative prior is then the (truncated) Jeffreys prior (Jeffreys 1961), uniform in the log-space,

$$J(x) = \frac{\ln(1 + x_{max}/x_{min})}{x + x_{min}}, \quad (\text{C29})$$

where  $x_{min}$ ,  $x_{max}$  are upper and lower bounds such that if  $x > x_{max}$  or  $x < x_{min}$  then  $J(x) = 0$ . This guarantees that the probability density is proper (the integral over x is finite). In case of weak information content in the observables, this prior is more weighted toward a null-value. We set  $\epsilon_{min} = 5.10^{-4}$  and  $\epsilon_{max} = 10^{-2}$ . This embraces the solar value  $\epsilon \approx 5.10^{-4}$  (Gizon 2002). Because at the maximum of solar activity,  $\delta \approx 10^\circ$ , we set  $\delta_{min} = 10^\circ$ . Note also that consistently with the discussion on Figure 2 of Section 3,  $\delta_{max} = 45^\circ$  is required. Finally, the location parameter  $\theta_0$  has an uniform prior in the range  $[0, 90]^\circ$ .

The marginal likelihood of the model  $M_{AR}$  requires us to evaluate the triple integral,

$$P(a_{2,o}, a_{4,o}, \sigma_{2,o}, \sigma_{4,o} | M) = \int \int \int \pi(\epsilon) \pi(\theta) \pi(\delta) L_2(a_{2,o}, a_{4,o}, \sigma_{2,o}, \sigma_{4,o} | a_{2,m}, a_{4,m}, M) d\epsilon d\theta d\delta. \quad (\text{C30})$$

The MCMC process used here involves the use of parallel Metropolis-Hastings tempered chains. The chains are mixing each other in order to enhance the sampling. As explained in Benomar et al. 2009, see their Section A.3, these parallel chains can be used to approximate the equation (C30). This technique is here used with 10 parallel chains following a geometrical temperature law  $T_k = 1.7^{k-1}$ , with  $k$ , the chain index such that  $k = 1$  is the target distribution, given by equation (C24).

753

754

755

756

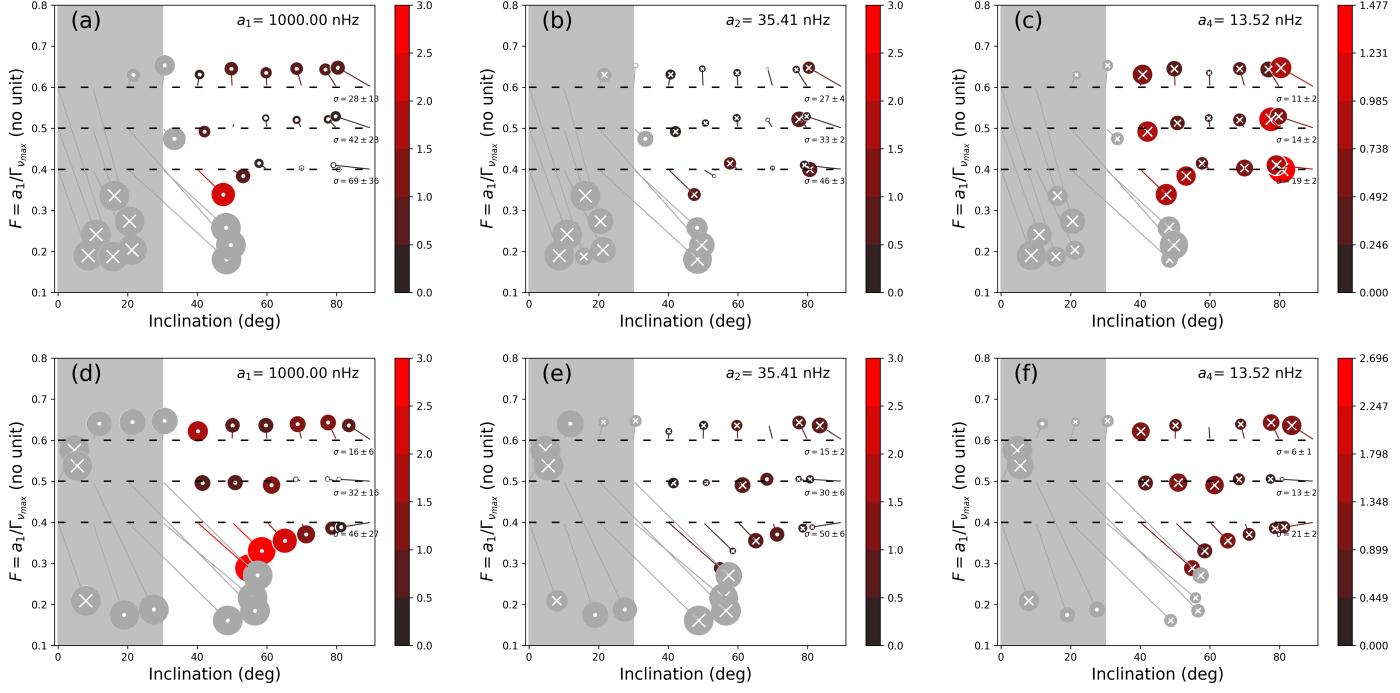
757

758

759

760

This section shows the bias on the  $a_1$ ,  $a_2$  and  $a_4$  coefficients in the case of  $\widehat{HNR} = 10$  or  $\widehat{HNR} = 20$ . The figure 16 and 18 are for the case of an equatorial band of activity and figure 17 and 19, for a polar cap. It is noticeable that the inaccuracy of the fit remains smaller than standard deviation in the majority of the parameter space and for  $i > 30^\circ$ ,  $a_1/\Gamma_{\nu_{max}} > 0.4$ . Interestingly, the relative-to-error bias on  $a_1$  is more pronounced in the case of an equatorial band of activity and for  $\{a_1/\Gamma_{\nu_{max}} \leq 0.4, i = [30, 60]\}$  (Figure 16d and 18d) and can exceed 3 times the uncertainty. The accuracy of the inference of activity based on the a-coefficients is expected to weakly depend on the  $\widehat{HNR}$  (although potentially with large uncertainty) for most stars, except within this regime (in addition to the gray area), where the centrifugal effect will be overestimated.



**Figure 16.** Bias analysis for  $\widehat{HNR} = 10$  for an equatorial activity band ( $\theta_0 = 85^\circ, \delta = 10^\circ$ ) of similar intensity to the Sun ( $\epsilon_{nl} = 5.10^{-4}$ ), for  $T_{obs} = 2$  years (top) and  $T_{obs} = 4$  years (bottom).

761

762

763

764

765

766

767

768

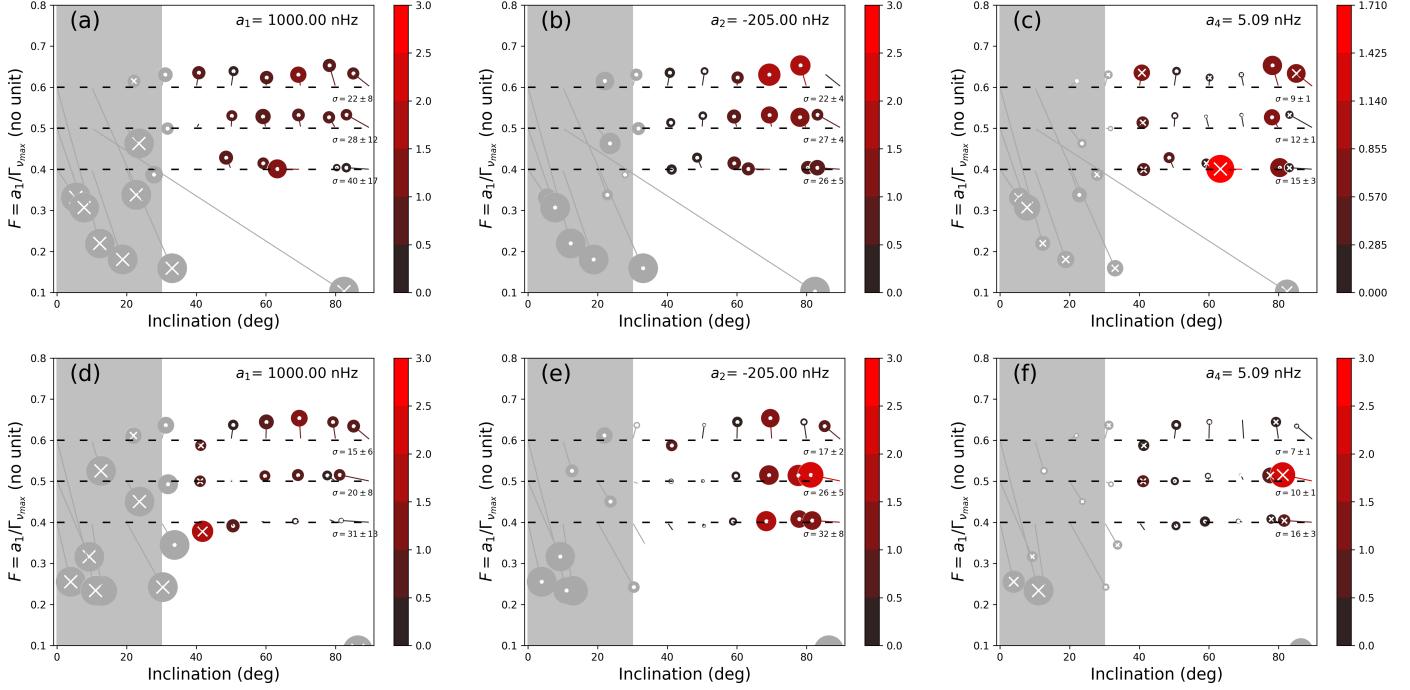
769

770

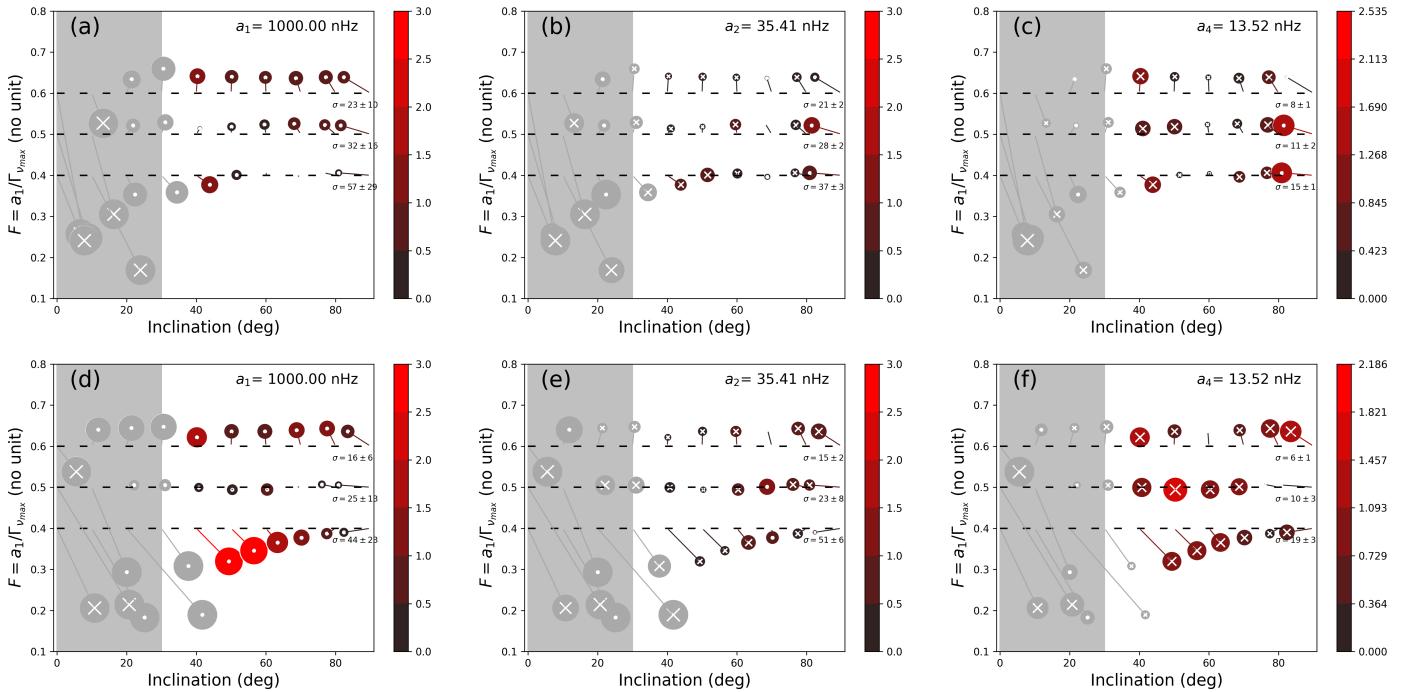
## E. FIGURES OF THE BEST FITS

This section presents visuals on the best MCMC fits for the power spectrum of the active Sun (Figure 20), quiet Sun (Figure 21), 16 Cyg A (Figure 22) and for the two  $a_4$  solutions of 16 Cyg B (Figures 23 and 24). For 16 Cyg B, the solutions are here separated by selecting the median for the samples only below or above  $a_4 = 21 \text{ nHz}$ . In the case of the Sun, the impact of the activity on the profile of  $l = 2$  is very clear, as the shifts of the m-components introduces an asymmetry in power for the overall  $l = 2$  mode profile during the active phase. During the quiet phase such an asymmetry is not visible.

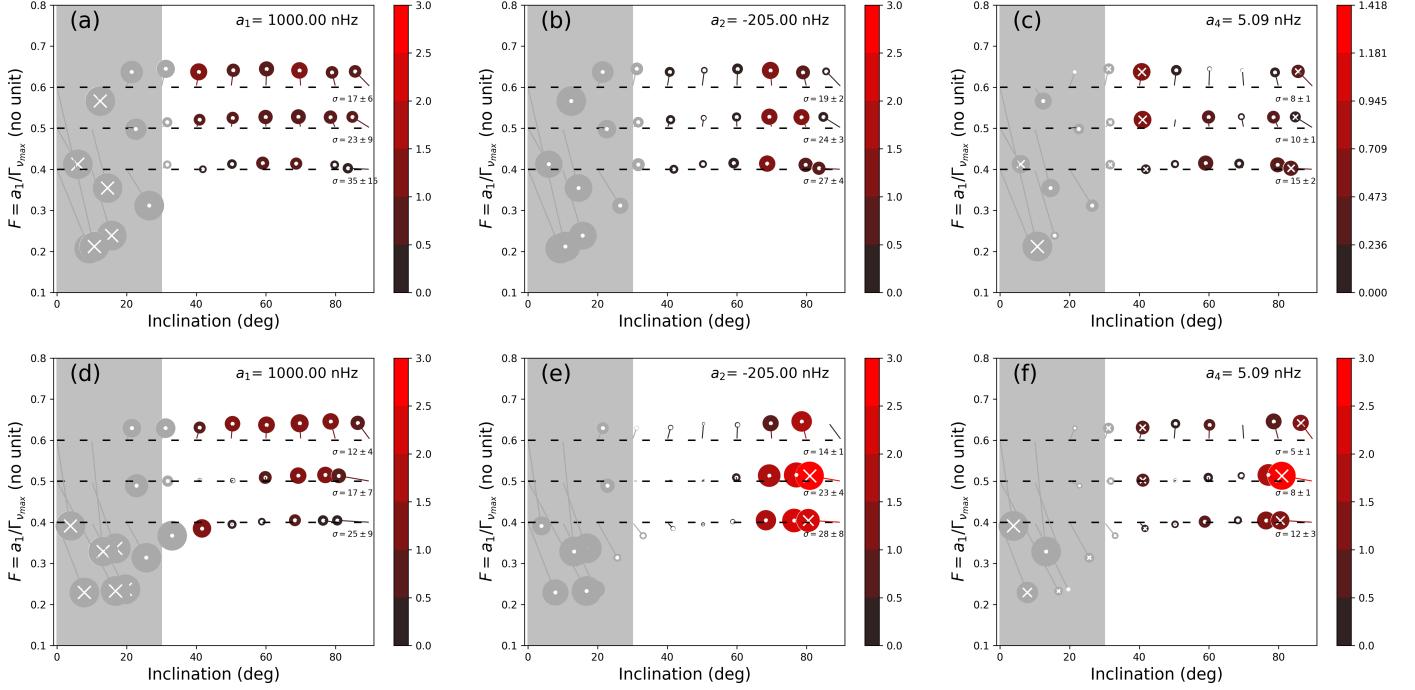
Similarly and although it is less pronounced than in the Sun, in the case of 16 Cyg A and B, a weak asymmetry is visible on  $l = 2$  modes. For 16 Cyg B, the two solutions of  $a_4$  provide sensibly similar mode profiles as the difference is weakly apparent to the eye only for  $l = 2$ . The residuals do not show striking differences. This visually supports the fact that goodness of the fit is similar and explain that the two solutions co-exists in statistical terms.



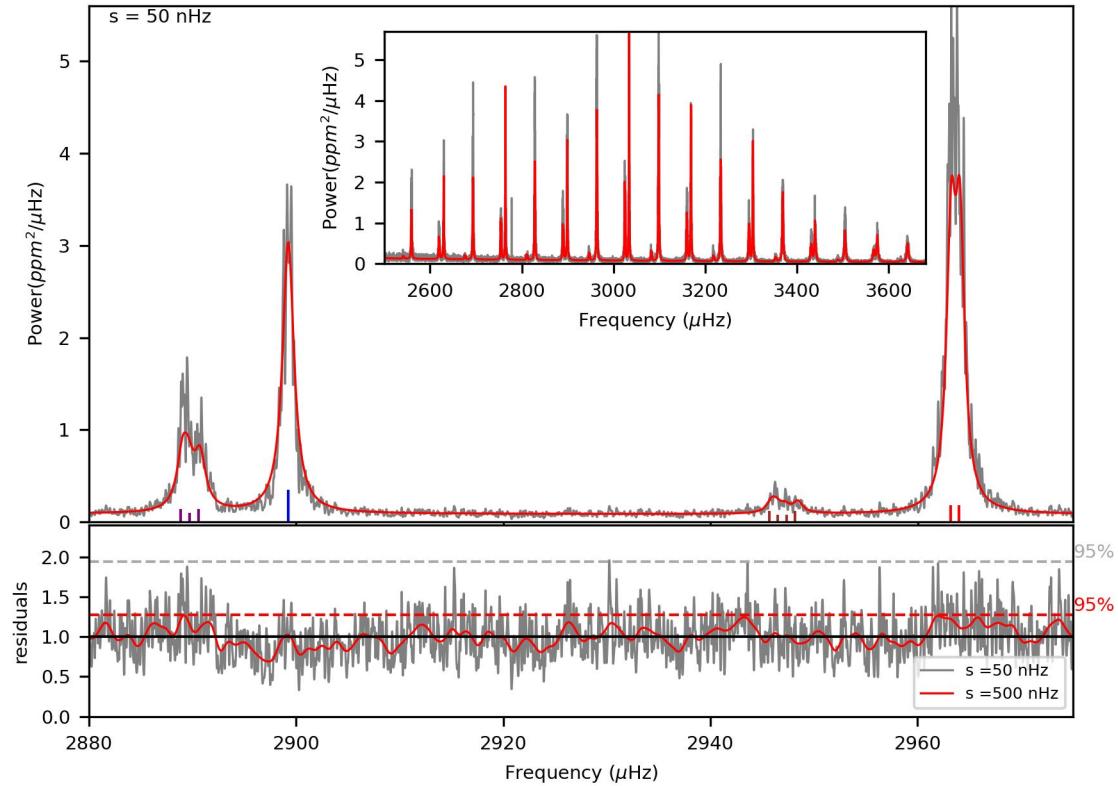
**Figure 17.** Bias analysis for  $\widehat{HNR} = 10$  for a large polar activity cap ( $\theta_0 = 22.5^\circ, \delta = 45^\circ$ ) of similar intensity to the Sun ( $\epsilon_{nl} = 5.10^{-4}$ ), for  $T_{obs} = 2$  years (top) and  $T_{obs} = 4$  years (bottom).



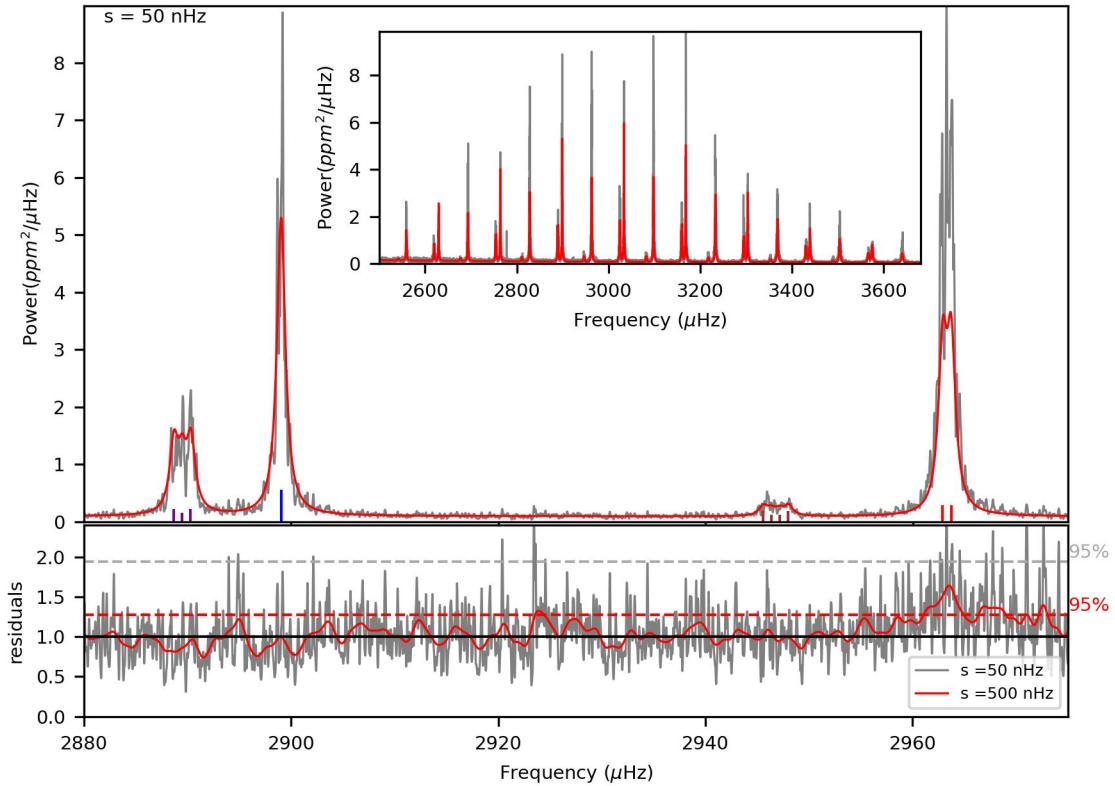
**Figure 18.** Bias analysis for  $\widehat{HNR} = 20$  for an equatorial activity band ( $\theta_0 = 85^\circ, \delta = 10^\circ$ ) of similar intensity to the Sun ( $\epsilon_{nl} = 5.10^{-4}$ ), for  $T_{obs} = 2$  years (top) and  $T_{obs} = 4$  years (bottom).



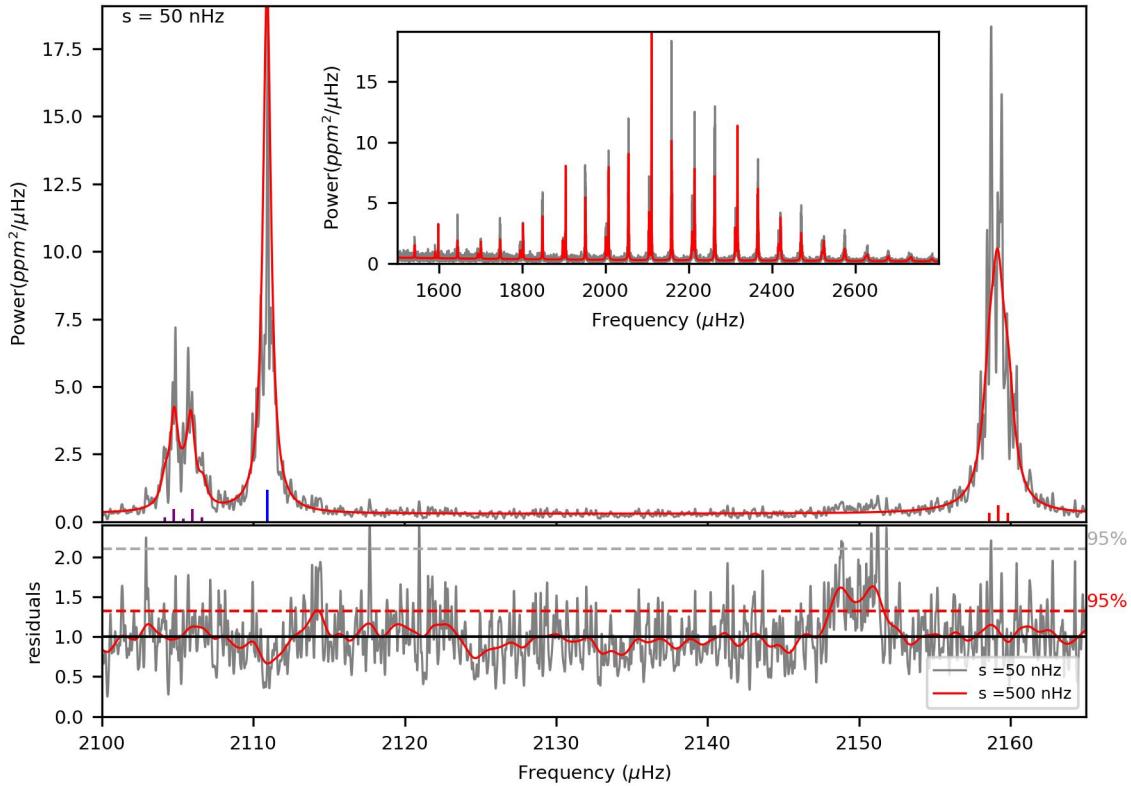
**Figure 19.** Bias analysis for  $HNR = 20$  for a large polar activity cap ( $\theta_0 = 22.5, \delta = 45$ ) of similar intensity to the Sun ( $\epsilon_{nl} = 5.10^{-4}$ ), for  $T_{obs} = 2$  years (top) and  $T_{obs} = 4$  years (bottom).



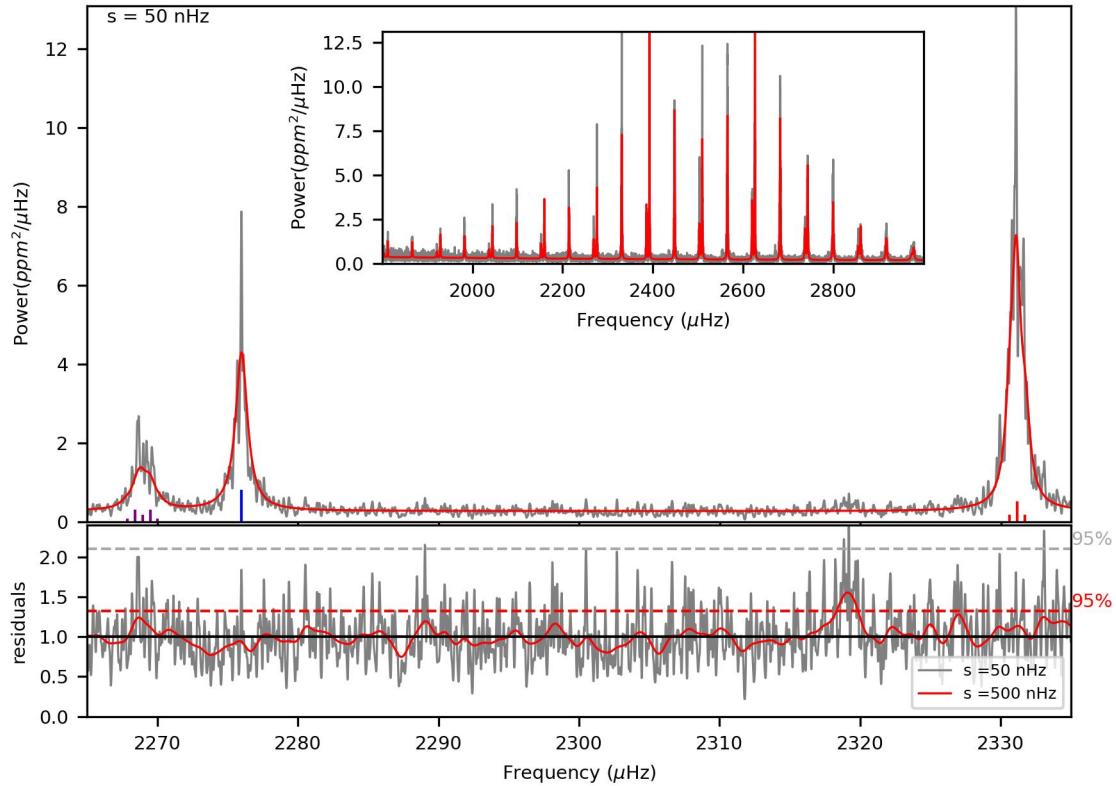
**Figure 20.** Solar power spectrum for observation between 1999 and 2002 after Gaussian smoothing over a kernel width  $s=50\text{nHz}$ . Superimposed is the best MCMC fit. **Top.** Highlight of a  $l = 2, 0, 3, 1$  mode group (left to right) and their fit (red). The power asymmetry due to non-symmetric  $m$  components is visible in the data. **Inset.** Overall view of the modes. **Bottom.** Residual of the fit with two level  $s$  of gaussian smoothing.



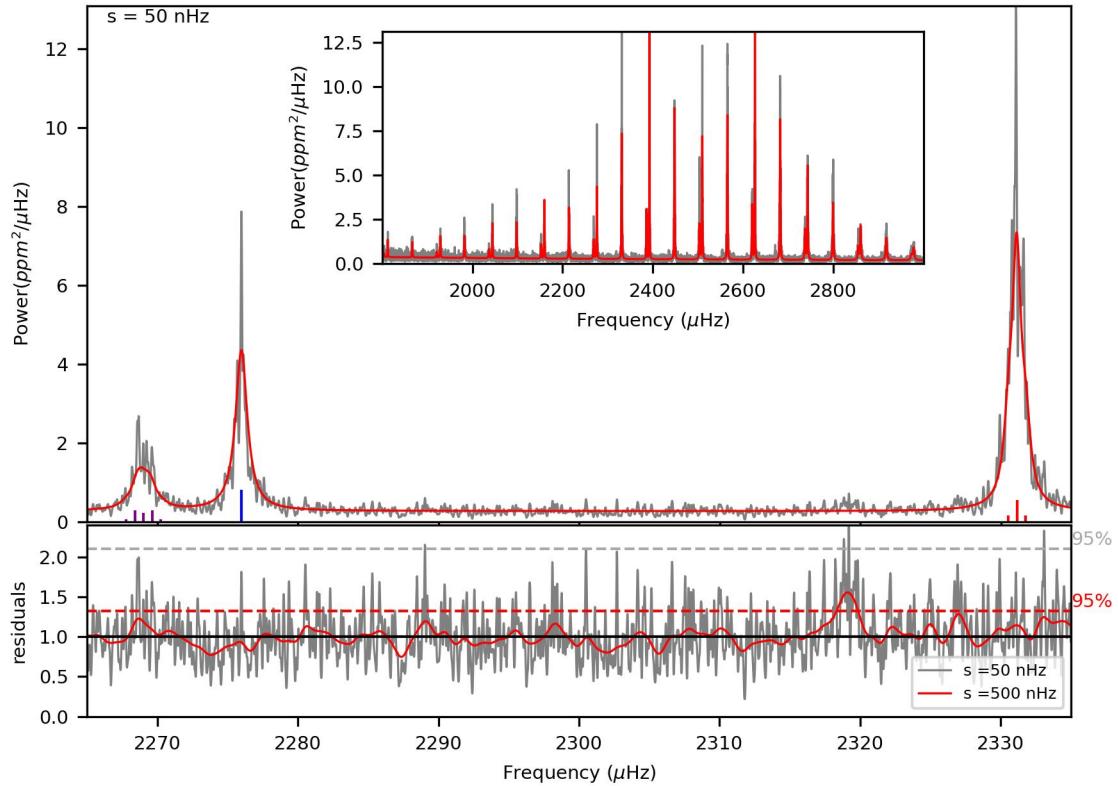
**Figure 21.** Same as Figure 20 but between 2006 and 2009. **Top.** Highlight of a  $l = 2, 0, 3, 1$  mode group (left to right) and their fit (red). The power asymmetry seen for 1999–2002 is not apparent in the data. **Inset.** Overall view of the power spectrum. **Bottom.** Residual of the fit with two level s of gaussian smoothing. The residual on  $l = 1$  is a bit high here due to the fix visibility for all modes.



**Figure 22.** Same as Figure 20 but for 16 Cyg A. **Top.** Highlight of a  $l = 2, 0, 1$  mode group (left to right) and their fit (red). A very mild power asymmetry is seen on the  $l = 2$  data. **Inset.** Overall view of the power spectrum. **Bottom.** Residual of the fit with two level  $s$  of gaussian smoothing. The residual show an excess of power due to the low HNR $\approx 1.7$   $l = 3$  (not fitted here).



**Figure 23.** Same as Figure 20 but for 16 Cyg B,  $a_4 < -21 \text{ nHz}$ . **Top.** Highlight of a  $l = 2, 0, 1$  mode group (left to right) and their fit (red). A mild power asymmetry is seen on the  $l = 2$  data. **Inset.** Overall view of the power spectrum. **Bottom.** Residual of the fit with two level  $s$  of gaussian smoothing. The residual show an excess of power due to the low HNR  $\approx 1.6$   $l = 3$  (not fitted here).



**Figure 24.** Same as Figure 20 but for  $16 \text{ Cyg B}$ ,  $a_4 > -21 \text{ nHz}$ . **Top.** Highlight of a  $l = 2, 0, 1$  mode group (left to right) and their fit (red). **Inset.** Overall view of the power spectrum. **Bottom.** Residual of the fit with two level  $s$  of gaussian smoothing. The residual show an excess of power due to the low HNR $\simeq 1.6$   $l = 3$  (not fitted here).

## REFERENCES

- 772 Anderson, E. R., Duvall, Jr., T. L., & Jefferies, S. M. 1990, ApJ,  
 773 364, 699
- 774 Antia, H. M., Chitre, S. M., & Thompson, M. J. 2000, A&A, 360,  
 775 335
- 776 Appourchaux, T. 2003, A&A, 412, 903
- 777 Appourchaux, T., Gizon, L., & Rabello-Soares, M.-C. 1998,  
 778 A&AS, 132, 107
- 779 Appourchaux, T., et al. 2008, A&A, 488, 705  
 780 —. 2014, A&A, 566, A20
- 781 Atchadé, Y. F. 2006, Methodology and Computing in Applied  
 782 Probability, 8, 235
- 783 Baglin, A., Auvergne, M., Barge, P., Deleuil, M., Catala, C.,  
 784 Michel, E., Weiss, W., & COROT Team. 2006, in ESA Special  
 785 Publication, Vol. 1306, ESA Special Publication, ed.  
 786 M. Fridlund, A. Baglin, J. Lochard, & L. Conroy, 33
- 787 Ballot, J., Barban, C., & van't Veer-Mennenet, C. 2011, A&A, 531,  
 788 A124
- 789 Bazot, M. 2020, A&A, 635, A26
- 790 Bazot, M., Benomar, O., Christensen-Dalsgaard, J., Gizon, L.,  
 791 Hanasoge, S., Nielsen, M., Petit, P., & Sreenivasan, K. R. 2019,  
 792 A&A, 623, A125
- 793 Bellinger, E. P., Basu, S., Hekker, S., & Ball, W. H. 2017, ApJ,  
 794 851, 80
- 795 Benomar, O. 2008, Communications in Asteroseismology, 157, 98
- 796 Benomar, O., Appourchaux, T., & Baudin, F. 2009, A&A, 506, 15
- 797 Benomar, O., Bedding, T. R., Stello, D., Deheuvels, S., White,  
 798 T. R., & Christensen-Dalsgaard, J. 2012, ApJL, 745, L33
- 799 Benomar, O., Takata, M., Shibahashi, H., Ceillier, T., & García,  
 800 R. A. 2015, MNRAS, 452, 2654
- 801 Benomar, O., et al. 2018a, Science, 361, 1231  
 802 —. 2018b, ApJ, 857, 119
- 803 Bishop, C. M. 1995, Neural Networks for Pattern Recognition  
 804 (Oxford University Press, Inc)
- 805 Borucki, W. J., et al. 2010, Science, 327, 977
- 806 Brun, A. S., & Browning, M. K. 2017, Living Reviews in Solar  
 807 Physics, 14, 4
- 808 Campante, T. L., et al. 2011, A&A, 534, A6  
 809 —. 2016, ApJ, 819, 85
- 810 Catura, R. C., Acton, L. W., & Johnson, H. M. 1975, ApJL, 196,  
 811 L47
- 812 Ceillier, T., et al. 2017, A&A, 605, A111
- 813 Chandrasekhar, S. 1969, Ellipsoidal figures of equilibrium
- 814 Chaplin, W. J., Elsworth, Y., Isaak, G. R., Miller, B. A., New, R.,  
 815 Thiery, S., Boumier, P., & Gabriel, A. H. 2003, Monthly Notices  
 816 of the Royal Astronomical Society, 343, 343
- 817 Cochran, W. D., Hatzes, A. P., Butler, R. P., & Marcy, G. W. 1997,  
 818 ApJ, 483, 457
- 819 Davies, G. R., et al. 2015, MNRAS, 446, 2959
- 820 Deheuvels, S., et al. 2012, ApJ, 756, 19  
 821 —. 2014, A&A, 564, A27
- 822 Deliyannis, C. P., Cunha, K., King, J. R., & Boesgaard, A. M.  
 823 —. 2000, AJ, 119, 2437
- 824 Duvall, Jr., T. L., Jefferies, S. M., Harvey, J. W., Osaki, Y., &  
 825 Pomerantz, M. A. 1993, ApJ, 410, 829
- 826 Dziembowski, W. A., Goode, P. R., Kosovichev, A. G., & Schou, J.  
 827 —. 2000, ApJ, 537, 1026
- 828 Frohlich, C., et al. 1997, SoPh, 170, 1
- 829 García, R. A., Salabert, D., Ballot, J., Sato, K., Mathur, S., &  
 830 Jiménez, A. 2011a, Journal of Physics Conference Series, 271,  
 831 012049
- 832 García, R. A., et al. 2011b, MNRAS, 414, L6
- 833 Georgobiani, D., Kosovichev, A. G., Nigam, R., Nordlund, Å., &  
 834 Stein, R. F. 2000, ApJL, 530, L139
- 835 Gizon, L. 2002, Astronomische Nachrichten, 323, 251
- 836 —. 2006, Central European Astrophysical Bulletin, 30, 1
- 837 Gizon, L., & Solanki, S. K. 2003, ApJ, 589, 1009  
 838 —. 2004, SoPh, 220, 169
- 839 Gizon, L., et al. 2016, Science Advances, 2, e1601777
- 840 Gough, D. O., & Taylor, P. P. 1984, Mem. Soc. Astron. Italiana, 55,  
 841 215
- 842 Gough, D. O., & Thompson, M. J. 1990, MNRAS, 242, 25
- 843 Gregory, P. C. 2005, Bayesian Logical Data Analysis for the  
 844 Physical Sciences: A Comparative Approach with  
 845 ‘Mathematica’ Support
- 846 Handberg, R., & Campante, T. L. 2011, A&A, 527, A56
- 847 Hansen, C. J., Cox, J. P., & van Horn, H. M. 1977, ApJ, 217, 151
- 848 Hartmann, L., & Rosner, R. 1979, ApJ, 230, 802
- 849 Harvey, J. 1985, ESA SP, 235, 199
- 850 Hogg, D. W., Myers, A. D., & Bovy, J. 2010, ApJ, 725, 2166
- 851 Huber, D., et al. 2011, ApJ, 743, 143
- 852 Jeffreys, H. 1961, Theory of Probability, 3rd edn. (Oxford,  
 853 England: Oxford)
- 854 Kamiaka, S., Benomar, O., & Suto, Y. 2018, \mnras, 479, 391
- 855 King, J. R., Deliyannis, C. P., Hiltgen, D. D., Stephens, A., Cunha,  
 856 K., & Boesgaard, A. M. 1997, AJ, 113, 1871
- 857 Libbrecht, K. G., & Woodard, M. F. 1990, Nature, 345, 779
- 858 Lignières, F., Rieutord, M., & Reese, D. 2006, A&A, 455, 607
- 859 Lund, M. N., Miesch, M. S., & Christensen-Dalsgaard, J. 2014a,  
 860 ApJ, 794, 96
- 861 Lund, M. N., et al. 2014b, A&A, 570, A54  
 862 —. 2017, ApJ, 835, 172
- 863 Maia, M. T., Meléndez, J., Lorenzo-Oliveira, D., Spina, L., &  
 864 Jofré, P. 2019, A&A, 628, A126
- 865 Mathur, S., et al. 2010, A&A, 511, A46
- 866 Metcalfe, T. S., Egeland, R., & van Saders, J. 2016, ApJL, 826, L2
- 867 Metcalfe, T. S., et al. 2012, ApJL, 748, L10
- 868 Neckel, H. 1986, A&A, 159, 175

- 869 Nielsen, M. B., Schunker, H., Gizon, L., Schou, J., & Ball, W. H.  
 870 2017, A&A, 603, A6
- 871 Nigam, R., & Kosovichev, A. G. 1998, ApJL, 505, L51
- 872 Nsamba, B., Cunha, M. S., Rocha, C. I. S. A., Pereira, C. J. G. N.,  
 873 Monteiro, M. J. P. F. G., & Campante, T. L. 2022, MNRAS
- 874 Oláh, K., Kővári, Z., Petrovay, K., Soon, W., Baliunas, S., Kolláth,  
 875 Z., & Vida, K. 2016, A&A, 590, A133
- 876 Oláh, K., et al. 2009, A&A, 501, 703
- 877 Papini, E., & Gizon, L. 2019, Frontiers in Astronomy and Space  
 878 Sciences, 6
- 879 Ramírez, I., Meléndez, J., & Asplund, M. 2009, A&A, 508, L17
- 880 Rauer, H., et al. 2014, Experimental Astronomy, 38, 249
- 881 Ricker, G. R., et al. 2014, in Society of Photo-Optical  
 882 Instrumentation Engineers (SPIE) Conference Series, Vol. 9143,  
 883 Proc. SPIE, 914320
- 884 Ritzwoller, M. H., & Lavelle, E. M. 1991, ApJ, 369, 557
- 885 Roxburgh, I. W. 2017, A&A, 604, A42
- 886 Schou, J., Christensen-Dalsgaard, J., & Thompson, M. J. 1994,  
 887 ApJ, 433, 389
- 888 Schuler, S. C., Cunha, K., Smith, V. V., Ghezzi, L., King, J. R.,  
 889 Deliyannis, C. P., & Boesgaard, A. M. 2011, ApJL, 737, L32
- 890 Silva-Valio, A., Lanza, A. F., Alonso, R., & Barge, P. 2010, A&A,  
 891 510, A25
- 892 Simon, T., Ayres, T. R., Redfield, S., & Linsky, J. L. 2002, ApJ,  
 893 579, 800
- 894 Stahn, T. 2010, PhD thesis, Georg August University of Gottingen,  
 895 Germany
- 896 Takeda, Y. 2005, PASJ, 57, 83
- 897 Thompson, M. J., Christensen-Dalsgaard, J., Miesch, M. S., &  
 898 Toomre, J. 2003, ARA&A, 41, 599
- 899 Toutain, T., Appourchaux, T., Frohlich, C., Kosovichev, A.,  
 900 Rakesh, N., & Scherrer, P. 1998, in ESA Special Publication,  
 901 Vol. 418, Structure and Dynamics of the Interior of the Sun and  
 902 Sun-like Stars, ed. S. Korzennik, 973
- 903 Toutain, T., & Gouttebroze, P. 1993, A&A, 268, 309
- 904 Toutain, T., & Kosovichev, A. G. 2001, in ESA Special  
 905 Publication, Vol. 464, SOHO 10/GONG 2000 Workshop: Helio-  
 906 and Asteroseismology at the Dawn of the Millennium, ed.  
 907 A. Wilson & P. L. Pallé, 123–126
- 908 Ulrich, R. K. 1986, ApJL, 306, L37
- 909 Unno, W., Osaki, Y., Ando, H., Saio, H., & Shibahashi, H. 1989,  
 910 Nonradial Oscillations of Stars (University Tokyo Press)
- 911 van Saders, J. L., Ceillier, T., Metcalfe, T. S., Silva Aguirre, V.,  
 912 Pinsonneault, M. H., García, R. A., Mathur, S., & Davies, G. R.  
 913 2016, Nature, 529, 181
- 914 Vaughan, A. H. 1983, in Solar and Stellar Magnetic Fields: Origins  
 915 and Coronal Effects, ed. J. O. Stenflo, Vol. 102, 113–131
- 916 Verma, K., et al. 2014, ApJ, 790, 138
- 917 Walker, G., et al. 2003, PASP, 115, 1023
- 918 White, S. M. 1999, SoPh, 190, 309
- 919 White, S. M., et al. 2017, SoPh, 292, 88
- 920 White, T. R., Bedding, T. R., Stello, D., Christensen-Dalsgaard, J.,  
 921 Huber, D., & Kjeldsen, H. 2011, ApJ, 743, 161
- 922 White, T. R., et al. 2013, MNRAS, 433, 1262