Butterfly diagram of a Sun-like star observed using asteroseismology

M. Bazot, M. B. Nielsen, D. Mary, J. Christensen-Dalsgaard, O. Benomar, P. Petit, L. Gizon, K. R. Sreenivasan, T. R. White

▶ **To cite this version:**

M. Bazot, M. B. Nielsen, D. Mary, J. Christensen-Dalsgaard, O. Benomar, et al.. Butterfly diagram of a Sun-like star observed using asteroseismology. Astronomy and Astrophysics - A&A, 2018, 619, 10.1051/0004-6361/201834251 . insu-03678166

HAL Id: insu-03678166

https://insu.hal.science/insu-03678166

Submitted on 30 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A&A 619, L9 (2018)
https://doi.org/10.1051/0004-6361/201834251
© ESO 2018

Astronomy & Astrophysics

LETTER TO THE EDITOR

Butterfly diagram of a Sun-like star observed using asteroseismology

M. Bazot[1,2], M. B. Nielsen[2], D. Mary[3], J. Christensen-Dalsgaard[4], O. Benomar[2], P. Petit[5], L. Gizon[6,7,2], K. R. Sreenivasan[2,8], and T. R. White[4]

[1] Division of Sciences, New York University Abu Dhabi, UAE
e-mail: mb6215@nyu.edu
[2] Center for Space Science, NYUAD Institute, Abu Dhabi, UAE
[3] Laboratoire Lagrange, Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Boulevard de l'Observatoire, CS 34229, 06304 Nice Cedex 4, France
[4] Stellar Astrophysics Centre, Department of Physics and Astronomy, Aarhus University, Ny Munkegade 120, 8000 Aarhus C, Denmark
[5] IRAP (Institut de Recherche en Astrophysique et Planétologie), Université de Toulouse, CNRS, CNES, UPS, 31400 Toulouse, France
[6] Max-Planck-Institut für Sonnensystemforschung, Göttingen, Germany
[7] Institut für Astrophysik, Georg-August-Universität Göttingen, Göttingen, Germany
[8] New York University, NY 10012, USA

Received 14 September 2018 / Accepted 19 October 2018

ABSTRACT

Stellar magnetic fields are poorly understood, but are known to be important for stellar evolution and exoplanet habitability. They drive stellar activity, which is the main observational constraint on theoretical models for magnetic field generation and evolution. Starspots are the main manifestation of the magnetic fields at the stellar surface. In this study we measured the variation in their latitude with time, called a butterfly diagram in the solar case, for the solar analogue HD 173701 (KIC 8006161). To this end, we used *Kepler* data to combine starspot rotation rates at different epochs and the asteroseismically determined latitudinal variation in the stellar rotation rates. We observe a clear variation in the latitude of the starspots. It is the first time such a diagram has been constructed using asteroseismic data.

Key words. stars: activity – stars: rotation – stars: oscillations – starspots – stars: solar-type – stars: individual: HD 173701

1. Introduction

Solar activity is characterized by an 11 year cycle of the number and area of sunspots (Schwabe 1844). Its monitoring is important in many fields, such as Earth climate (Haigh 2007 or space travel studies (Pulkkinen 2007). Sunspots are regions of high concentration of the solar magnetic field (Solanki 2003), indicating that the field is the main driver of the activity cycle. In order to understand the solar magnetic behaviour, dynamo models have been developed (Charbonneau 2010). Their aim is to explain how an initially weak magnetic field can be amplified to the values observed in the Sun. A traditional observational test to which these models ought to comply is to reproduce the butterfly diagram which describes the evolution of the latitude of the solar active regions with time (Maunder 1904). Activity has also been observed in other stars. It is important in order to understand stellar evolution, for instance through magnetic braking (Thompson et al. 2003), and exoplanet habitability (Vidotto et al. 2013).

Recovering stellar butterfly diagrams is however a difficult task that requires locating individual spots or groups of spots on the stellar surface. Spot mapping using photometric data alone is known to be hampered by degeneracies in light curve models (e.g. Walkowicz et al. 2013), so that spectroscopic or interferometric data are usually favoured in order to recover stellar brightness maps (Vogt & Penrod 1983; Roettenbacher et al. 2016). A number of active targets have been monitored through long-term Doppler mapping campaigns, although due to sparse temporal sampling these time series have rarely led to actual butterfly diagrams (e.g. Hackman et al. 2011, for II Peg). However, Doppler mapping or long baseline interferometric imaging can only detect very large stellar spots, and are therefore limited to the most active stars, quite far from the Sun in terms of magnetic properties. We propose here an original method that uses *Kepler* time series to measure the latitudes of the active regions of stars and construct a stellar butterfly diagram. We apply it to the sun-like star HD 173701 (KIC 8006161).

2. Method

Our approach uses *Kepler* photometric time series (Jenkins et al. 2010) and is divided into two stages. First we wanted to obtain information on the large-scale rotational flow in the stellar interior, in particular in the convective envelope below the surface, using the information contained in the oscillation frequencies of the global acoustic pulsation modes (p modes) of the star (Sect. 2.1). To that end, we used the recent results of

Benomar et al. (2018). Second, we tried to measure the rotation rates of active regions, carried by the surface rotational flow (Sect. 2.2). Third, we constructed the butterfly diagram of HD 173701 over the duration of the *Kepler* mission by inverting the rotation-rate measurements thanks to the asteroseismically derived rotation profile (Sect. 2.3).

We note that these measurements are uncorrelated in Sun-like stars since the observed pulsation modes of the stars are only affected by large-scale flows. In the frequency space, these two contributions are well separated, the modes having characteristic frequencies of the order of a few thousand μHz, while the rotation rates inferred from the starspots are of the order of a few hundred nHz.

2.1. Asteroseismic rotation profile

Our starting point is the derivation of a rotational profile of a Sun-like star. It is well established that the solar radiative interior rotates as a solid body, while the rotation rate of its outer convective envelope varies with latitude (Thompson et al. 2003). For the Sun, the equator rotates faster (~476 nHz) than the pole (~320 nHz). This differential rotation is most probably the result of an interplay between rotation and convection in the solar envelope (Rüdiger 1974; Gilman & Glatzmaier 1981).

Rotation affects stellar pulsations. The classical framework to describe stellar oscillations is built on the perturbed equations for the static stellar structure (Aerts et al. 2010). It can be shown that superimposing a small-amplitude rotational velocity field on the hydrostatic stellar structure will lift a degeneracy of the eigenfrequencies of the p modes, an effect usually termed "rotational splitting" (Lynden-Bell & Ostriker 1967). The degenerate frequency then becomes a multiplet whose distribution depends on the solid-body component of rotation rate of the star and on the magnitude of the differential rotation in the convective zone. If one defines a theoretical model $\Omega = \Omega(\theta)$ for the latitudinal differential rotation, with θ being the co-latitude, it is possible to relate it to the frequency splitting. This means that one can infer a differential-rotation profile if there are precise enough measurements of the frequency splitting.

In a recent work, Benomar et al. (2018) measured contributions from latitudinal differential rotation to the above-mentioned rotational splitting for a set of Sun-like stars. It includes the solar analogue HD 173701, which has a mass $M = 0.95\,M_\odot$, with $M_\odot$ the solar mass, and an age $t_\star = 4.49$ Gyr (Silva Aguirre et al. 2017). This star rotates on average 1.33 times faster than the Sun, with a measured bulk rotation rate of 566 nHz. It has been observed for approximately four years by *Kepler*. Detailed modelling of its acoustic power spectrum led to a significant detection of non-zero latitudinal differential rotation. In the following we explain how to extend such results with rotation-rate measurements in order to construct the butterfly diagram of the HD 173701. For the sake of completeness we review the work of Benomar et al. (2018) in more detail in Appendix A.

At this point we have to make a further assumption relative to Benomar et al. (2018). It has been shown by Kiefer et al. (2017), Santos et al. (2018), and Salabert et al. (2018) that the oscillation frequencies of HD 173701 shift with the activity cycle. We therefore need to take these shifts to be constant over frequency ranges of the order of a rotational splitting.

2.2. Photometric characterization of active regions

Once obtained, a relation $\Omega(\theta)$ can be used to infer the co-latitude, θ_a, of a stellar active region such as a group of spots

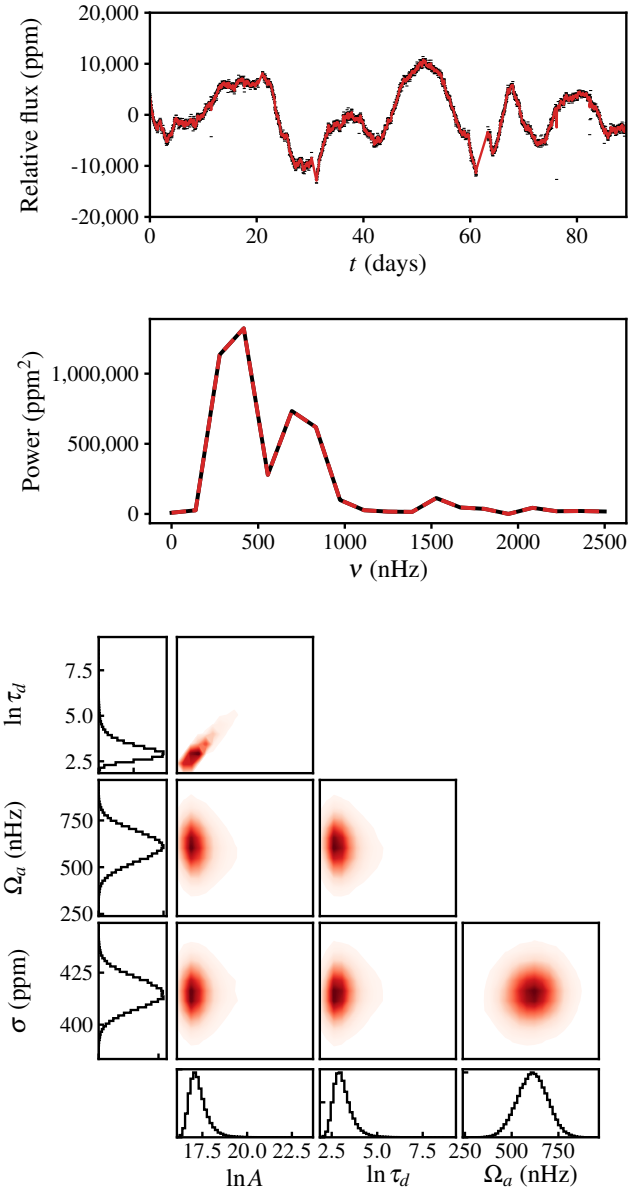

Fig. 1. *Upper panel*: time series for the relative flux HD 173701 during Q3. The black dots mark the measurements and associated errors. The red line shows the inferred mean value of the posterior predictive density conditional on the inferred MAP of the parameters (Gelman 2004). *Middle panel*: corresponding power spectrum computed using the Lomb-Scargle periodogram (Scargle 1982). The black line represents the observed power spectrum. The red line is the power spectrum of the mean value of the conditional posterior predictive density. *Lower panel*: marginalized 1D and 2D PDFs for the parameters of the correlation function of the Gaussian process used to model the time series.

provided its rotation rate, $\Omega_a = \Omega(\theta_a)$, can be determined. The method for rotation rate measurements is inspired by solar observations and uses photometry. The total solar irradiance varies during the activity cycle due to bright regions (plages) and dark regions (spots) moving across the solar disk (Fröhlich & Lean 2004). A similar behaviour is observed for other stars using photometric measurements. During its lifetime, an active region will produce a quasi-periodic signal in the photometric time series. The long-cadence *Kepler* time series allow us to measure such modulations for HD 173701.

To measure the period of this modulation, we model the time series using Gaussian processes. Owing to the intermittent

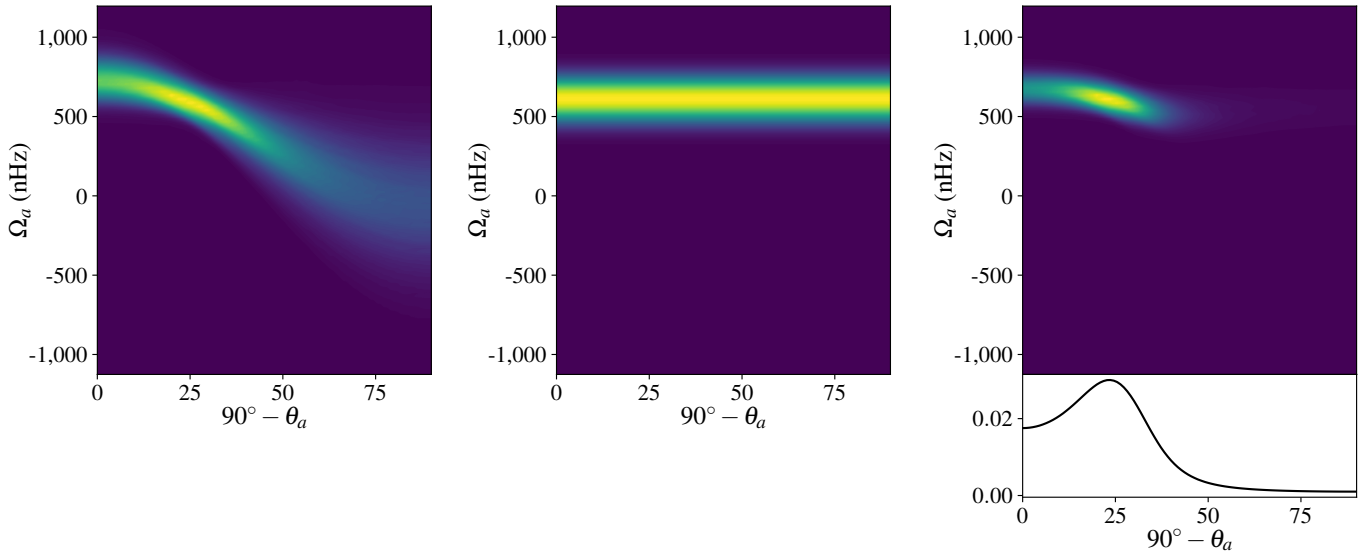

Fig. 2. Probability density for the latitude of an active region at median time of Q3. *Left panel*: theoretical density for the couple (Ω_a, θ_a). *Central panel*: prior density for the couple (Ω_a, θ_a); it is Gaussian for Ω_a and uniform for θ_a. *Upper right panel*: posterior density for the parameters. *Lower right panel*: marginal density for θ_a, obtained after integration over Ω_a.

influence of the plages and spots, and to the stellar limb darkening, this signal departs significantly from pure harmonic oscillations. In addition, the noise impacting the photometric times series has a frequency-dependent part mainly caused by surface granulation. The time series model depends on several parameters which are estimated using an MCMC algorithm. These parameters are the amplitude of the modulation induced by the active region A (in ppm), the lifetime of the active regions τ_d (in days), the rotation rate Ω_a (in nHz), and an additional noise component σ (in ppm), describing residuals not captured by the other component of the Gaussian process (see Appendix B for details).

We model independently the time series for each *Kepler* quarter which span from Q0 to Q17. The fit obtained for Q3, representative of our results, is shown in Fig. 1; it reproduces well the time series and its frequency spectrum. Our analysis is restricted to a region in the frequency spectrum surrounding the low-frequency activity peak (Fig. 1, middle panel). In fact, we observe multiple peaks potentially corresponding to several active regions. We consider that each of these peaks corresponds to a spot or a group of spots that rotates at the same latitude. We treat them as a single active region and consider the rotation rate of its barycentre. This is motivated by the fact that, when counting sunspots, a larger weight is given to groups (Hathaway 2010). The independent modelling, quarter by quarter, allows us to measure the evolution of the rotation rate with time. We note that quarters Q0, Q1, and Q17 were left out because the corresponding time series are too short to obtain robust results.

The Gaussian processes reproduce the temporal variation of the intensity curve very well as can be seen in the upper panel of Fig. 1. Using the marginal density of the rotation rate Ω_a, we can estimate its maximum a posteriori (MAP). Such a density is seen for Q3 in the lower panel Fig. 1, corresponding to a MAP of 704 nHz. The MAP estimates of the rotation rates for all quarters are in the range 330–985 nHz, which already indicates that the spots are migrating along the stellar surface. In general the amplitudes of the signal and the lifetimes of the active regions are poorly constrained. Since they are strongly correlated, a wide range of values for the lifetime and the amplitude can lead to good models. Consequently, their estimated values cover several orders of magnitude. This does not impact the final result since these parameters are not correlated with the others.

2.3. Inversion for the latitude of the active regions

Our goal is now to invert for the latitude of the active region at each for each quarter we selected. Since the latitudinal rotation profile has been derived using the entire *Kepler* time series, it is necessary to assume that the properties of differential rotation do not vary on timescales comparable to the activity cycle. In the solar case, this is verified to a very good approximation (Thompson et al. 2003).

In order to invert for the latitude of the active regions, we need to take into account all the errors that may affect our final estimate of $\theta_a(t)$. These are of two kinds. The first is the error on the measurement of the rotation rate, the second is the uncertainty in the theoretical model for the rotation rate, $\Omega(\theta)$, as obtained by (Benomar et al. 2018, see also Appendix A).

A generic framework to solve an inverse problem that takes into account these two sources of error is given by the concept of conjunction of states of information (Tarantola & Valette 1982, see Appendix C). In the following we use the more compact notation $\Omega_a(t_i) := \Omega_i$ and $\theta(t_i) := \theta_i$, with $i = 1, \ldots, N$ and N the number of rotation rate measurements. The posterior density for (Ω_i, θ_i), which represents the state of information (or state of knowledge, see Appendix C) we have once all the errors have been considered, can be written

$$\sigma_i(\Omega_i, \theta_i) = \frac{\rho_i(\Omega_i, \theta_i)\Theta(\Omega_i, \theta_i)}{\mu(\Omega_i, \theta_i)}. \quad (1)$$

The theoretical density $\Theta(\Omega_i, \theta_i)$ describes the error on the model. It is basically the relation $\Omega_i = \Omega_i(\theta_i)$ with an associated error bar. The prior density $\rho_i(\Omega_i, \theta_i)$ is the combination of the information obtained from the observation on Ω_i (see Fig. 1, lower panel) and the prior information on the parameter θ_i (which has been chosen uniform over the range 0°–90°). The posterior density $\sigma_i(\Omega_i, \theta_i)$ results from the conjunction of these two pieces of information. In Fig. 2 we show these three

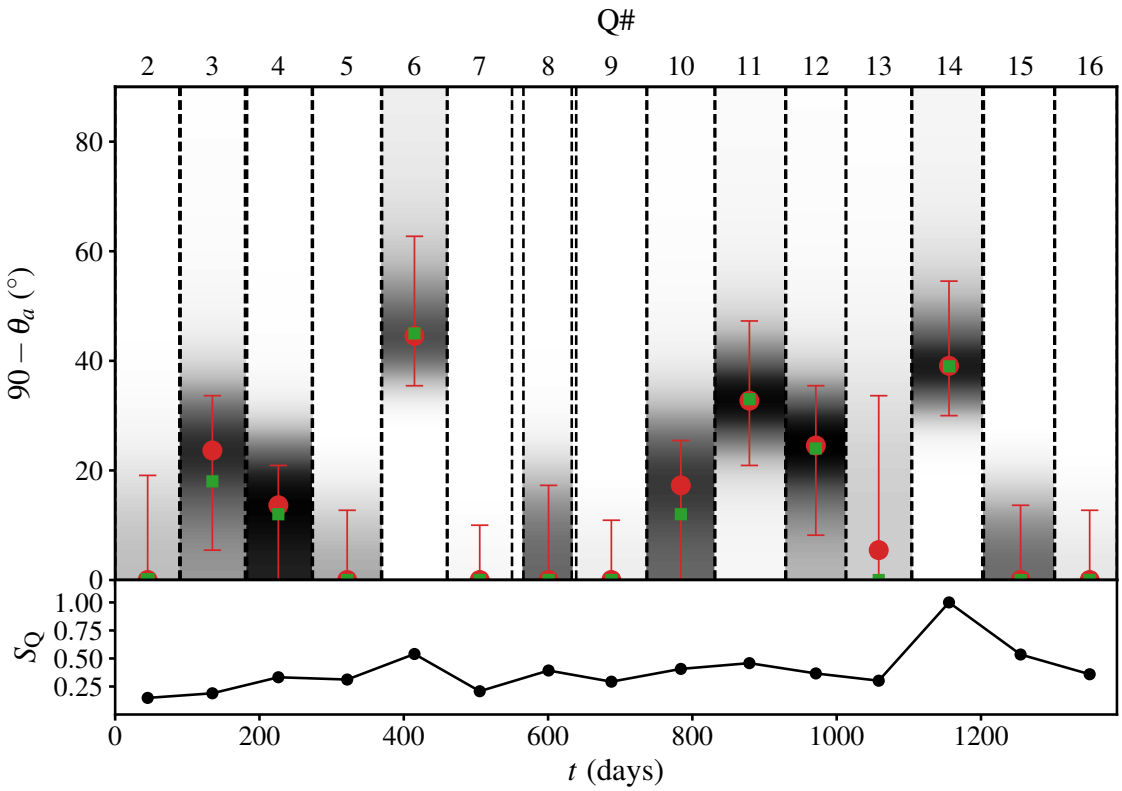

Fig. 3. Asteroseismic butterfly diagram of HD 173701, giving the latitude of the active region as a function of the median time of each quarter (labelled on the upper axis). The grey shades show the posterior densities obtained for each of the selected quarters and normalized to the global maximum; greyscale from 0 (white) to 1 (black). Red dots mark the maximum a posteriori estimates of the latitude. The vertical bars give the limits of the corresponding 68.3% credible intervals. The green squares show the maximum of the density of the 1024 estimates obtained from artificial time series. *Lower panel*: variation in amplitude of the signal, S_Q, normalized to its maximum.

densities for quarter Q3, respectively in the left, central, and right panel. The marginal posterior on θ_i is shown in the lower right panel, and is obtained from the integration over all possible rotation rates of the joint posterior. The meaning of the homogeneous density $\mu(\Omega_i, \theta_i)$ is given in Appendix C.

3. Result

We use the marginal posteriors of $\theta_a(t)$ to derive the butterfly diagram pictured in Fig. 3 for all selected quarters. The latitude of the active region is restricted to the range [0°−90°] since we do not resolve the stellar disk and thus do not know on which hemisphere the spot is located.

The credible intervals were computed as the smallest intervals that encompass the MAP and over which the probability mass is 0.683. There is a clear variation in the data. Half the active regions are found at the equator. Among the other half, five clearly exclude the 0° latitude at a 68.3% level. The active regions seen in Q6 and Q14 exclude it at a 99.7% level. It is thus ruled out that the signal seen could be due to equatorial active regions, with the high-latitude regions being outliers driven by intrinsic stochastic variability.

The estimated co-latitude (red dots in Fig. 3) rely on the parameters of the Gaussian process, which were estimated from the time series. This means that different data sets of the same star in the same activity configuration would lead to different inferred latitudes. In order to investigate the variability of the angle estimation procedure (and hence check for the robustness of these results), we made Monte Carlo simulations. We simulated 1024 time series using the MAP estimates of the parameters of the Gaussian processes for each quarter[1]. We applied the latitude-estimation process to each artificial time series, resulting in a sample of 1024 MAP estimates for each such latitude of the active region. The distribution of this MAP obtained from artificial time series can be compared to the MAP obtained directly from the actual data. The location of the maxima for these distributions are shown in Fig. 3 as green squares. The MAP locations are essentially not impacted by the noise realization, which supports the fact that the co-latitude extracted from the *Kepler* data for this star must be close to the values shown in Fig. 3. Slightly larger deviations are seen when the uncertainty on θ_a is large, which is to be expected.

Over the four years of measurement, the active regions remain located at latitudes below 50°. The butterfly diagram clearly shows an alternance between equatorial and high-latitude active regions.

In the lower panel of Fig. 3 we plotted S_Q, which we define as the standard deviation of the light curve evaluated over quarter Q. We used this quantity as a proxy for the activity index $S_{\rm ph}$ rather than the definition given in García et al. (2014) for two reasons. First it involves an evaluation over a period that is five times the rotation rate, which becomes hard to interpret when we take into account the variability of the rotation rate. Second, if we retain the average rotation rate of KIC 8006161, then a quarter is approximately 4.5 times larger, which is commensurate with

[1] To generate samples we used the predictive probability density (Rasmussen & Williams 2005, Sect 2.1.1).

the definition of $S_{\rm ph}$ and preserves the statistical independence of the points in our time series.

We first see that low-amplitude modulations are observed for active regions at high latitude and near the equator. Let us assume that the amplitude of the signal follows the variations of the fraction of the stellar surface covered by the active regions. Then this observation in qualitative agreement with what is already known for the Sun (Hathaway 2010, Fig. 28) where small spots are observed at all latitudes (within the band in which they are confined). The two highest-amplitude modulations are seen for the regions with the highest latitudes, during Q6 and Q14. If our assumption that the amplitude relates to the surface covered by the active regions, then this departs from what is seen in the solar case. There the largest spots are seen at mid-latitudes.

Extracting a periodicity for the activity cycle from these measurements remains a difficult task. The periodogram of the signal does not show any significant peaks at low frequencies. Time series with longer spans would be needed to obtain such an estimate. It is interesting to note that the two highest-latitude spots also correspond to the longest estimated lifetimes (in the sense of the MAP), with $\tau_{\rm d} \gtrsim 1$ yr. This is of course an overestimation, since the signal lasts less than this characteristic time (see Appendix B). It should be noted that the lifetime of the active region and its rotation rate are uncorrelated, we can therefore safely assume that these biases do not impact the estimates of the latitude for these epochs (as can be checked in Fig. 1 for Q3). Such values nevertheless indicate that active regions are producing a coherent signal over a large fraction of Q6 and Q14. This could correspond to one or several long-lived large spots. Again this seems to depart from what is known in the Sun, for which spot lifetimes are much smaller.

Other activity measurements have been obtained for HD 173701. Karoff et al. (2018) have obtained Ca II H and K line measurements that can track the activity cycle. Kiefer et al. (2017), Santos et al. (2018), and Salabert et al. (2018) have all measured consistent values for the activity-induced frequency shifts over the last eleven quarters of the of the *Kepler* mission. The Salabert et al. measurements correlate well with $S_{\rm ph}$. Further work and larger data sets are needed to understand precisely how the butterfly diagram correlates with these other indicators.

4. Conclusion

In this letter we presented the first stellar butterfly diagram derived obtained by combining information inferred from asteroseismic and photometric analyses. Provided an approximate stellar model is known, the only data required to perform the inversion is a photometric times series (collected from *Kepler* in this study). This approach identifies a powerful link between asteroseismology and other branches of stellar physics studying stellar magnetism, for instance Zeeman-Doppler Imaging (Semel 1989). Stellar butterfly diagrams, out of reach of Zeeman-Doppler mapping when it is applied to solar analogues, nicely complement the long-term monitoring of large-scale magnetic geometries still accessible for low activity stars. Both proxies of stellar cycles offer complementary views to understand the underlying dynamo processes. The technique itself requires only moderate computational time and can be envisaged as a systematic processing for surveys of starspots in the perspective of the forthcoming space mission PLATO.

Acknowledgements. This material is based upon work supported by the NYU Abu Dhabi Institute under grant G1502. Funding for the Stellar Astrophysics Centre is provided by The Danish National Research Foundation (Grant DNRF106). In memory of Albert Tarantola.

References

Aerts, C., Christensen-Dalsgaard, J., & Kurtz, D. W. 2010, Asteroseismology (Netherlands: Springer)
Angus, R., Morton, T., Aigrain, S., Foreman-Mackey, D., & Rajpaul, V. 2018, MNRAS, 474, 2094
Benomar, O., Bazot, M., Nielsen, M. B., et al. 2018, Science, 361, 1231
Brown, T. M., Christensen-Dalsgaard, J., Dziembowski, W. A., et al. 1989, ApJ, 343, 526
Charbonneau, P. 2010, Liv. Rev. Sol. Phys., 7, 3
Christensen-Dalsgaard, J. 2008a, Ap&SS, 316, 13
Christensen-Dalsgaard, J. 2008b, Ap&SS, 316, 113
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Foreman-Mackey, D., Agol, E., Ambikasaran, S., & Angus, R. 2017, AJ, 154, 220
Fröhlich, C., & Lean, J. 2004, A&ARv, 12, 273
García, R. A., Ceillier, T., Salabert, D., et al. 2014, A&A, 572, A34
Gelman, A. 2004, Bayesian Data Analysis, Texts in Statistical Science (Chapman & Hall/CRC)
Gilman, P. A., & Glatzmaier, G. A. 1981, ApJS, 45, 335
Gizon, L., & Solanki, S. K. 2003, ApJ, 589, 1009
Gough, D. O., & Thompson, M. J. 1990, MNRAS, 242, 25
Hackman, T., Mantere, M. J., Jetsu, L., et al. 2011, Astron. Nachr., 332, 859
Haigh, J. D. 2007, Liv. Rev. Sol. Phys., 4, 2
Hansen, C. J., Cox, J. P., & van Horn, H. M. 1977, ApJ, 217, 151
Hathaway, D. H. 2010, Liv. Rev. Sol. Phys., 7, 1
Jaynes, E. T. 1968, IEEE Trans. Syst. Sci. Cybern., 4, 227
Jenkins, J. M., Caldwell, D. A., Chandrasekaran, H., et al. 2010, ApJ, 713, L120
Karoff, C., Metcalfe, T. S., Santos, Â. R. G., et al. 2018, ApJ, 852, 46
Kiefer, R., Schad, A., Davies, G., & Roth, M. 2017, A&A, 598, A77
Lund, M. N., Lundkvist, M., Silva Aguirre, V., et al. 2014, A&A, 570, A54
Lynden-Bell, D., & Ostriker, J. P. 1967, MNRAS, 136, 293
Maunder, E. W. 1904, MNRAS, 64, 747
Pijpers, F. P. 1997, A&A, 326, 1235
Pulkkinen, T. 2007, Liv. Rev. Sol. Phys., 4, 1
Rasmussen, C. E., & Williams, C. K. I. 2005, Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning) (The MIT Press)
Ritzwoller, M. H., & Lavely, E. M. 1991, ApJ, 369, 557
Roettenbacher, R. M., Monnier, J. D., Korhonen, H., et al. 2016, Nature, 533, 217
Rüdiger, G. 1974, Astron. Nachr., 295, 229
Salabert, D., Régulo, C., Pérez Hernández, F., & García, R.A. 2018, A&A, 611, A84
Santos, A. R. G., Campante, T. L., Chaplin, W. J., et al. 2018, ApJS, 237, 17
Scargle, J. D. 1982, ApJ, 263, 835
Schou, J., Christensen-Dalsgaard, J., & Thompson, M. J. 1994, ApJ, 433, 389
Schwabe, M. 1844, Astron. Nachr., 21, 233
Semel, M. 1989, A&A, 225, 456
Silva Aguirre, V., Lund, M. N., Antia, H. M., et al. 2017, ApJ, 835, 173
Solanki, S. K. 2003, A&ARv, 11, 153
Tarantola, A. 2004, Inverse Problem Theory and Methods for Model Parameter Estimation (Philadelphia, PA, USA: Society for Industrial and Applied Mathematics)
Tarantola, A., & Valette, B. 1982, J. Geophys., 50, 159
Thompson, M. J., Christensen-Dalsgaard, J., Miesch, M. S., & Toomre, J. 2003, ARA&A, 41, 599
Vidotto, A. A., Jardine, M., Morin, J., et al. 2013, A&A, 557, A67
Vogt, S. S., & Penrod, G. D. 1983, PASP, 95, 565
Walkowicz, L. M., Basri, G., & Valenti, J. A. 2013, ApJS, 205, 17

Appendix A: Inference of the differential rotation profile

In this section we describe briefly the principles of asteroseismic inversion that allow us to estimate the parameters of our model for the stellar rotation rate $\Omega(\theta)$. In generic form it can be written as an expansion over a basis of functions $W_s(\theta)$ depending on even powers of $\cos(\theta)$ (Brown et al. 1989)

$$\Omega(r,\theta) = \sum_{s=0}^{s_{\max}} \Omega_s(r) W_s(\theta). \tag{A.1}$$

The functions $\Omega_s(r)$ are often chosen as piecewise continuous functions in order to account for the change in rotational regime between the radiative interior and the convective envelope. In the case of HD 173701, the $\Omega_s(r)$ are chosen constants, and these are the parameters we estimate.

The rotational splitting can also be expressed as a basis expansion of the form (Brown et al. 1989)

$$\delta\nu_{n,l,m} = \sum_{j=0}^{j_{\max}} a_{2j+1}(n,l)\zeta_j^{(l)}(m), \tag{A.2}$$

where the $\zeta_j^{(l)}(m)$ form an orthogonal basis obeying $\sum_m \zeta_i^{(l)}(m)\zeta_j^{(l)}(m) = 0$ if $i \neq j$. Finally, the a_j and Ω_s are related through the integral equation

$$\delta\nu_{n,l,m} = \int_0^{\pi}\int_0^{R_\star} K_{n,l,m}(r,\theta)\Omega(r,\theta) r \mathrm{d}\theta \mathrm{d}r, \tag{A.3}$$

where $K_{n,l,m}(r,\theta)$ is a kernel that depends on the equilibrium stellar structure and the eigenfunction of the corresponding p modes (Hansen et al. 1977). To obtain these kernels we numerically solved the equations for the stellar structure and pulsations, using, respectively, ASTEC (Christensen-Dalsgaard 2008a) and `adipls` (Christensen-Dalsgaard 2008b).

A judicious choice of the basis functions $\zeta_j^{(l)}$ will ensure that there is a one-to-one relation between the coefficients a_{2s+1} of the splitting expansion and the Ω_s of the rotation rate expansion (Ritzwoller & Lavely 1991; Schou et al. 1994). The resulting orthogonality condition on the W_s allows us to derive their functional form (Pijpers 1997), which is $W_s(\theta)=P^1_{2s+1}(\cos\theta)/\sin\theta$, with $P_l^m(\cos\theta)$ being the associated Legendre polynomial of degree l and order m. Using this expression and setting $s_{\max}=1$ in Eq. (A.1), we obtain the formula used for the rotation rate in this study:

$$\Omega(\theta) = \Omega_0 - 1.5\Omega_1(5\cos^2\theta - 1) \tag{A.4}$$

We note that the condition on $s_{\max}$ is justified by the above-mentioned one-to-one relation and the fact that the current seismic data only allow us to observe a_1 and a_3 (tests to detect a_5 splitting coefficients in the data considered in this study were inconclusive). The goal is to obtain estimates of Ω_0 and Ω_1 in order to derive the theoretical density of the couple (Ω_i, θ_i).

In practice, the coefficients a_1 and a_3 were estimated using a parametric model for the power spectrum (Gizon & Solanki 2003), namely a sum of Lorentzian components and some noise. The locations of the Lorentzians are the mode eigenfrequencies. Their distribution is given by the relation

$$\nu_{n,l,m} = \nu_{n,l,0} + \delta\nu_{n,l,m} + \beta(\nu). \tag{A.5}$$

Here $\delta\nu_{n,l,m}$ describes the rotational effects and is given by Eq. (A.2). The additional term $\beta(\nu)$ has been introduced to account for the effects of departures from strict sphericity of the star (centrifugal force, magnetic fields, tidal distortion, etc.) and is a linear function of the central frequency of the multiplet (Gough & Thompson 1990). To estimate a_3, the mode of degree $l = 2$ was used.

The existence of a one-to-one relation between (a_1, a_3) and (Ω_0, Ω_1) was then used to estimate the latter. The measured rotational splitting in the observed acoustic spectrum was modelled

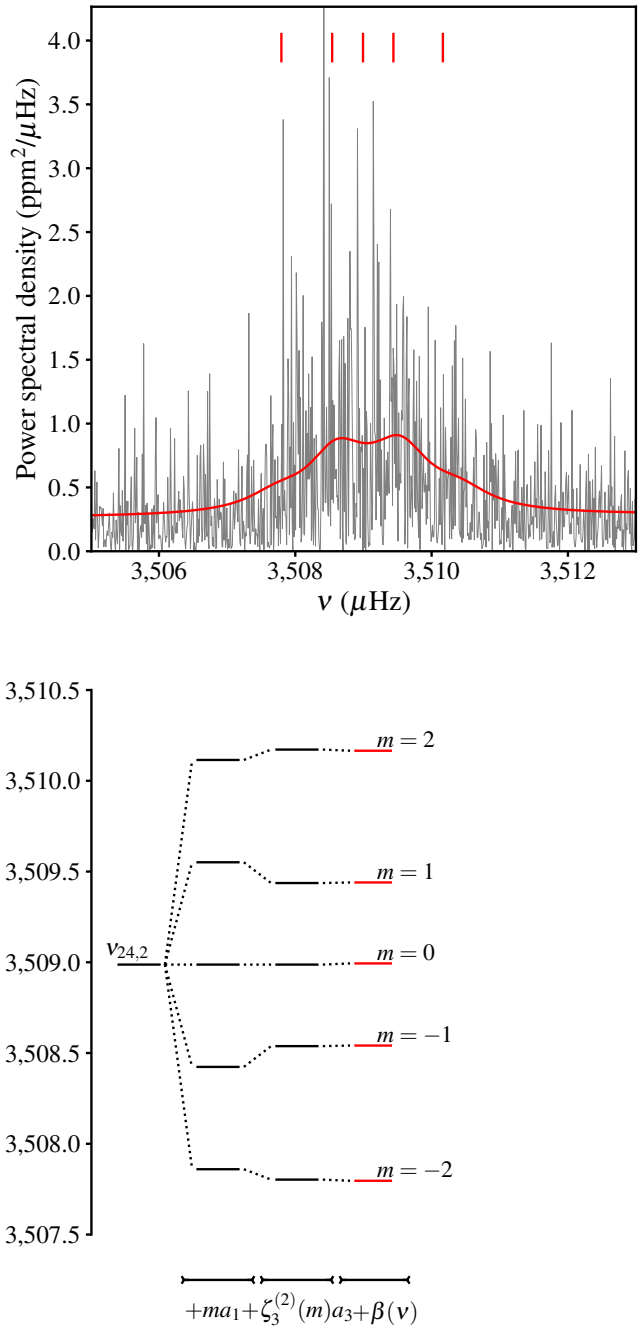

Fig. A.1. *Top panel*: power spectrum of the short-cadence *Kepler* data in the region of the eigenfrequency multiplet centred around the frequency $\nu_{n=24,l=2,m=0}$. The black and red lines show respectively the data and the best-fit model of the power spectrum. The vertical red ticks mark the frequencies of the eigenmodes of the multiplet, for $-2 \leq m \leq 2$. *Bottom panel*: splitting diagram for the multiplet. The first two terms of the splitting sequence correspond to the rotational effects, while $\beta(\nu)$ is an additional term that describes aspherical contributions to the eigenfrequencies. The red horizontal ticks correspond to those seen in the *bottom panel*.

(Benomar et al. 2018) using Eq. (A.2) with $j_{\max} = s_{\max} = 1$. The splitting coefficients were considered as free parameters to be estimated. The inferred values are $a_1 = 563 \pm 69$ nHz and $a_3 = 28.61 \pm 12.41$ nHz. An example of the modelled power spectrum is given in Fig. A.1, alongside the measured splittings. A model for the rotation rate of the convective zone of the form (A.4) was used together with kernels computed from stellar models (Christensen-Dalsgaard 2008a,b; Lund et al. 2014) to invert this model. Estimates for the rotation coefficients $\Omega_0 = 566 \pm 69$ nHz and $\Omega_1 = -104 \pm 45$ nHz were inferred.

Finally, we expand slightly on the assumption made concerning the influence of the activity-induced frequency shifts on the estimated value for a_3. The assumption we have to make is that the entire multiplet caused by rotational splitting is shifted as a block and thus the frequency splitting remains the same. This allows the activity-induced frequency shifts to depend on the frequency itself, but this variation has to remain negligible at scales of the order of $\sim 3\,\mu$Hz, which is characteristic of a rotational splitting in KIC 8006161. If this holds, then we may expect a_3 to be relatively constant with time. We note that this is related to but subtly different from the assumption of non-varying differential rotation magnitude with time made at the beginning of Sect. 2.3, which says that the magnitude of differential rotation does not change over time. The new assumption says that the measurement of this constant quantity a_3 is not biased by time-varying effects affecting the frequencies.

Appendix B: Measuring rotation from active regions

One element needed to infer the latitude of an active region is a measurement of its rotation rate. It is critical to obtain an estimate of the associated uncertainties in order to compute the conjunction of information states described in Sect. 2 (see also Appendix C).

The idea retained here is to fit the low-frequency components of the time series using a Gaussian process. This method has been applied recently to *Kepler* data with some success (Angus et al. 2018). The critical point that allowed the use of Gaussian processes to model long time series was the development of methods for the fast inversion of covariance matrices[2] (Foreman-Mackey et al. 2017).

The idea of fitting a Gaussian process to a time series with N points is to consider it as the realization of a random vector of dimension N. In principle one would then have to estimate the mean and the covariance matrix of the parent distribution, i.e. $N(N+3)/2$ parameters. However, the assumption is made that the process is stationary and thus that any given term of the covariance matrix can be determined from a correlation function, k, that only depends on the time difference, $k(t_i, t_j) = k(\tau_{ij})$, with $\tau_{ij} = |t_j - t_i|$. Adopting models developed by Foreman-Mackey et al. (2017), we chose a function of the form

$$k(\tau_{ij}) = \frac{A}{2} e^{-\tau_{ij}/\tau_{\rm d}} [\cos(2\pi\Omega_{\rm a}\tau_{ij}) + 1] + \sigma^2 \delta_{ij}. \tag{B.1}$$

The cosine term on the right-hand side models the rotation rate. The second term including σ^2 is sometimes dubbed "jitter" and may capture potential model errors or compensate for underestimated observational errors (Angus et al. 2018).

The parameters of the covariance model are $\boldsymbol{\lambda} := (A, \tau_{\rm d}, \Omega_{\rm a}, \sigma)$. We used a Bayesian statistical model to describe each time series. The posterior density of $\boldsymbol{\lambda}$ was estimated using an MCMC algorithm (Foreman-Mackey et al. 2013). An example of the resulting sampling is shown in Fig. 1. The likelihood is given by the density of the Gaussian process. The components of $\boldsymbol{\lambda}$ were assumed independent and the prior could be written as a product of uniform univariate densities. After several tests we found that we could efficiently sample the posterior density if we fit A and $\tau_{\rm d}$ in logarithmic space and $\Omega_{\rm a}$ and σ in linear space. We chose uniform priors for $\Omega_{\rm a}$ and σ and log-uniform priors for A and $\tau_{\rm d}$. The adopted boundaries on $\ln A$ and $\ln \tau_d$ are respectively $[-\infty, +\infty]$ and $[-10, 12]$. For $\Omega_{\rm a}$ and σ they are $[150, 1600]$, $[150, 4500]$. We note that the prior on $\Omega_{\rm a}$ was set so that our results do not get contaminated by very low- or very high-frequency signals that are not properly reproduced by our model (hence relating to filtering practices that can be found in other studies; Angus et al. 2018). These boundaries were set after several tests and trials.

The prior on the lifetime also demands some caution. It should be noted that its upper boundary is far greater than the duration of any time series we are using in this study. And indeed, some estimates of $\tau_{\rm d}$ from the MCMC simulation happen to be greater than a quarter duration. This can be viewed in several ways. First, it should be remembered that we are fitting a stochastic process to the time series. Obtaining long lifetimes out of the estimation process just means that high values of $\tau_{\rm d}$ are compatible with the degree of correlation between the different timescales probed by the time series and that this degree always remains high. From another perspective, the stationarity assumption under which we are working breaks down if we try to only model one active region over long series. Consequently, when we encounter large estimates for the lifetimes of the active regions, the precise value of the estimate should be seen as meaningless, or at least severely biased. What is important in these cases is that $\tau_{\rm d}$ is larger than the duration of the time series, indicating an active region that is stable over the duration of the entire subsample.

Appendix C: Conjunction of information states

The method of inversion is based on the framework described by Tarantola & Valette (1982). In this section we just summarize the main points of the approach and refer to the original paper for further details. The basic postulate is that any state of knowledge on the values of a set of parameters can be described using a density function[3,4]. The approach of the conjunction of states of information consists in expressing the posterior information on a couple $(\boldsymbol{d}, \boldsymbol{m})$, with $\boldsymbol{d}$ and $\boldsymbol{m}$ being the data and model-parameter vectors, using density functions

$$\sigma(\boldsymbol{d}, \boldsymbol{m}) = \frac{\rho(\boldsymbol{d}, \boldsymbol{m})\Theta(\boldsymbol{d}, \boldsymbol{m})}{\mu(\boldsymbol{d}, \boldsymbol{m})}, \tag{C.1}$$

where the densities σ, ρ, Θ, and μ represent states of information. They are not necessarily probability densities, which would require that they be normalizable. The density σ represents the conjunction of the two states of information described

[2] We used the `celerite` package for python. https://github.com/dfm/celerite

[3] By "state of knowledge" we mean knowledge of the quantitative characteristics of a system. Therefore, the existing information we have on this characteristic can be cast into a density function. This method is not fit to deal with the qualitative knowledge one might have on a system.

[4] This density is a probability density only if it can be normalized to one.

by ρ and Θ. The density $\mu(\boldsymbol{d}, \boldsymbol{m})$ is sometimes called a homogeneous probability density and represents the state of null information (Tarantola 2004). This means that μ is the density that is the least informative[5] on the values of the couple $(\boldsymbol{d}, \boldsymbol{m})$. It is present in Eq. (C.1) so that the conjunction of any state of information by the null state results in no loss of information. The functions ρ and Θ are, respectively, the prior probability density and the probability density of the theoretical model on $(\boldsymbol{d}, \boldsymbol{m})$. The function ρ contains the information on the system (parameter and data) before the inference process. This can be either the observational data or the prior information on the parameters of the model. The density Θ represents the uncertainties on the theoretical model.

We now switch back to the notations of Sect. 2 using the relations $\boldsymbol{d} = \{\Omega_i\}$ and $\boldsymbol{m} = \{\theta_i\}$. The data and the parameters are independent, hence we can write $\mu(\Omega_i, \theta_i) = \mu_{\Omega_i}(\Omega_i)\mu_{\theta_i}(\theta_i)$. In this work we choose μ_{Ω_i} and μ_{θ_i} as uniform densities (Tarantola 2004). Using the same argument of independence, we can write the prior probability density as a product of probability densities $\rho(\Omega, \theta) = \rho_{\Omega_i}(\Omega_i)\rho_{\theta_i}(\theta_i)$, where ρ_{Ω_i} is obtained from the Gaussian-process modelling and ρ_{θ_i} is uniform on the latitude interval $[0, 90°]$ since we have no prior information on the latitude of the active region. The range chosen is the consequence of our lack of spatial resolution between the two stellar hemispheres (see Sect. 3). Finally we write $\Theta(\Omega_i, \theta_i) = \eta_{\Omega_i}(\Omega_i|\theta_i)\mu_{\theta_i}(\theta_i)$ with $\eta_{\Omega_i}(.|\theta_i)$ the probability density on Ω_i conditional on $\theta = \theta_i$, i.e. the probability density of the theoretical rotation rate taken at θ_i (see the central panel of Fig. 2 for an illustration). These choices ensure that σ is normalizable and can be treated as a probability density.

The final step, once $\sigma(\Omega_i, \theta_i)$ has been estimated, is to obtain the posterior probability density on θ_i only instead on (Ω_i, θ_i). This is achieved by marginalizing the rotation rate, i.e. performing the integral

$$\sigma_{\theta_i}(\theta_i) = \int \sigma(\Omega_i, \theta_i) d\Omega_i, \tag{C.2}$$

over the relevant range of rotation rates. The density σ_{θ_i} is called the marginal probability density for θ_i. The σ_{θ_i} are the objects represented in Fig. 3 as a function of time t and θ_i and is what we call the butterfly diagram.

[5] As explained below, we are able to choose μ constant. However, the precise choice of the homogeneous density is a delicate problem that has been discussed for instance by Jaynes (1968).