Reply to the referee:

We would like to thank the referee for his careful reading and the comments which helped greatly to improve the paper.
We have modified several sections following the referee's comments, whenever possible. Since some modifications would have enlarged the paper significantly, we decided to also reorganize some of the figures,

- Figures 2 and 3 were merged as they address the same question regarding the validity of the model for the shape of the active region.
- Figure 4 is removed because it contains little relevant information. It is replaced by a short description in lieu of the former description (Section 3.3, line 279)

In the introduction, we have added a reference to Dziembowski et al. (2000) when mentioning the activity effect on the Sun shape (line 44).
In Section 3.3 about our assumption a reference to Chaplin et al. (2003) is also provided (line 272) and a paragraph discussing a comparison of their results with ours is provided in section 5.1.
A few additional changes were also made in Section 1 and 2 to better present the methodology. We for example explain earlier why the longitudinal dependence is neglected (this was made in Section 3 in the older version).
Furthermore, we have decided to remove the Appendix B, which was showing theoretical calculations justifying the fact that the longitude is ignored. This is replaced by a reference to Papini & Gizon 2019, which provides a much more detailed explanation and description on the effects of magnetic spots on pulsation frequencies.

Finally, following the recommendation from the data editor, we have added a sentence to point toward the DOI for the Kepler/MAST data. This appears now in the acknowledgment section.

These changes appear in bold in the new version of the paper.
Furthermore, all the changes that were required for this new version of the paper (figures, data and code) were updated in the Github repository associated to that paper and put https://github.com/OthmanB/Benomar2022/tree/version_2 (to keep the first version untouched in the main).

Please find below some details about the concerns and comments from the referee:

1. In helioseismology, measurements of the splitting coefficients from the integrated intensity or Doppler shift observation are quite uncertain. This is because the frequency splittings are comparable or smaller (depending on the mode frequency) than the line width of stochastically excited oscillations. While the asymmetry in the l=2 mode splitting was detected in the BISON and GOLF solar oscillation data (e.g., Chaplin et al., 2003, MNRAS, 343, 343), the splitting a-coefficients (except a1) were not robustly determined.

We thank the referee for pointing us towards this interesting paper that we had missed before.
It is true that if one takes non-radial modes individually, the measurement is quite uncertain in integrated intensity. After comprehensive tests that were briefly summarized in Section 3 of our paper, we in fact reach that same conclusion for VIRGO/SPM data. However, what our study

shows, is that the average estimate of the a-coefficients is robust and reliable. Reading carefully Chaplin et al., 2003, MNRAS, 343, 343, we were surprised that despite a very different methodology, they reach the same conclusion as they determined that Tn22 (what they called a2 is in fact our Tn22, see our appendix) is significant at more than 4 sigma. This is a strong detection of departure from 0. This degree of significance is similar to our detection of activity for approximately the same period of observation (the significance being given for the maximum of activity starting in 1999 in both studies).
Regarding the fact that the mode blending is problematic, we agree that it is a problem and we have done many simulations to evaluate the seriousness of that problem and considered different scenario to evaluate how much information can be extracted from the power spectra before we eventually lose in reliability. As explain in our paper, this problem is in fact addressed in several papers, the most recent and the most extensive being from Kamiaka+2018, that showed that the Sun (and most G and K type stars of Kepler) is on the reliable zone regarding the determination of the average rotational splitting. As you might have noticed, our paper build upon that to extend the analysis of the reliability zone by considering not only a1, but also a2 and a4 and confirm this result of Kamiaka+2018 on this larger parameter space.

2. In the first step, the authors define the mean values of the a-coefficients by fitting a simplified model to the full oscillation power spectra. However, the power spectra and the mode fits are not shown. Therefore, it is impossible to determine the quality of the data and the fitting model. Are the frequency peaks illustrated in Fig. 1 resolved and adequately fitted? This needs to be demonstrated. Finding a model's best set of parameters does not mean that the model provides a good fit to the data.

Indeed, having the best parameters does not necessarily mean a good fit. However, among all classes of fitting algorithms, the MCMC sampling algorithm is one that is insensitive to initial guesses, likelihood multimodality in large parameter spaces. In other words, the accuracy of the fit is in principle only limited by the accuracy of the power spectrum model: In case of poor information content in the power spectrum, the precision becomes simply large. The assumption on the fitting model being well documented and relatively standard and in use since Appourchaux+2008 (on the CoRoT star HD49933, except that they do not use a-coefficient, but just the average a1), we did not think that it was necessary to show the best fit due to the already-large number of figures.
However, as the referee asked, we have added an appendix that shows all the fit along with the residuals. In the case of the Sun during its maximum of activity (Figure 20), an asymmetry in mode power for the l=2 is evident and is the consequence of the frequency shifts of m-components. The effect is not visible at the minimum of activity (Figure 21). For 16 Cyg A and B, the shift being less significant, it is barely visible by eye, which explains its lower significance than for the Sun.

3. One significant concern in the validity of the fitted model is that the even a-coefficients should be scaled with the mode inertia (e.g., Dziembowski et al., 2000, ApJ, 537, 1026), which is a function of frequency. It seems that this scaling was not taken into account in the model. Although the measurements of the a2 and a4 coefficients may not be available for the total solar irradiance observations, the authors could compare their results with

the results of disk-resolved observations published in Dziembowski et al.(2000) and other helioseismology papers.

We understand the concern and it would indeed be sensical to correct from the inertia if one was measuring the individual a-coefficients, in function of the mode degree and radial order. However, the following reasons make us think that this is unnecessary with respect to the quality of the data we are dealing with, and the subsequent chosen methodology,

- The figure 1 of this response to the referee shows that the inertia l-dependence is negligible for l=1,2,3.
- Our fit being global (constant a-coefficients and inclination over the whole spectrum), the modes with highest HNR are vastly dominating the estimates of their best values. These highest HNR modes are around the numax, which is where inertia has variations of ~30 – 50 %.
- We attempted fitting the terms <a2(nu)>_l and <a4(nu)>_l using a linear function (ie of the type k1.nu + k2 averaged over l=1,2,3) to evaluate if a measure of the frequency-dependence of a-coefficients can be determined on VIRGO/SPM data during the 1999-2002 activity maximum. We found that the uncertainties are too large to see any significant slope with frequencies for a2 and a4, see Figure 2 of this document. This was briefly summarized in Section 3.3. And this is in fact consistent with Chaplin et al., 2003, MNRAS, 343, 343 (see their figure 7). Kepler solar-like stars having a HNR of at best a third of these solar-data, we conjecture that it is impossible to determine any frequency-dependence of the coefficients, which justifies the choice of a fitting a constant a coefficient over the whole frequency range.
- In fact, we tried to measure the slope of the a-coefficients in 16 Cyg A and B, but this severely affected the robustness of the fit (multimodality in the linear fit parameters of the a-coefficients).
- As shown with the extensive simulation that are presented in the paper, robustness, fair accuracy, and precision is ensured if one accept to focus only on the average of the a-coefficients. Robustness is severely degraded in conditions like Kepler observations if one attempts to measure more subtle effects.

Accounting for the inertia would only be possible if we were not fitting the individual a-coefficients. As the last point emphasizes, this is unfortunately difficult with integrated photometric data of the Sun, and thus very unlikely for other solar-like stars, which necessarily have a lower height-to-noise ratio.

In Appendix D1, we have added a paragraph that explains those reasons for not using inertia.

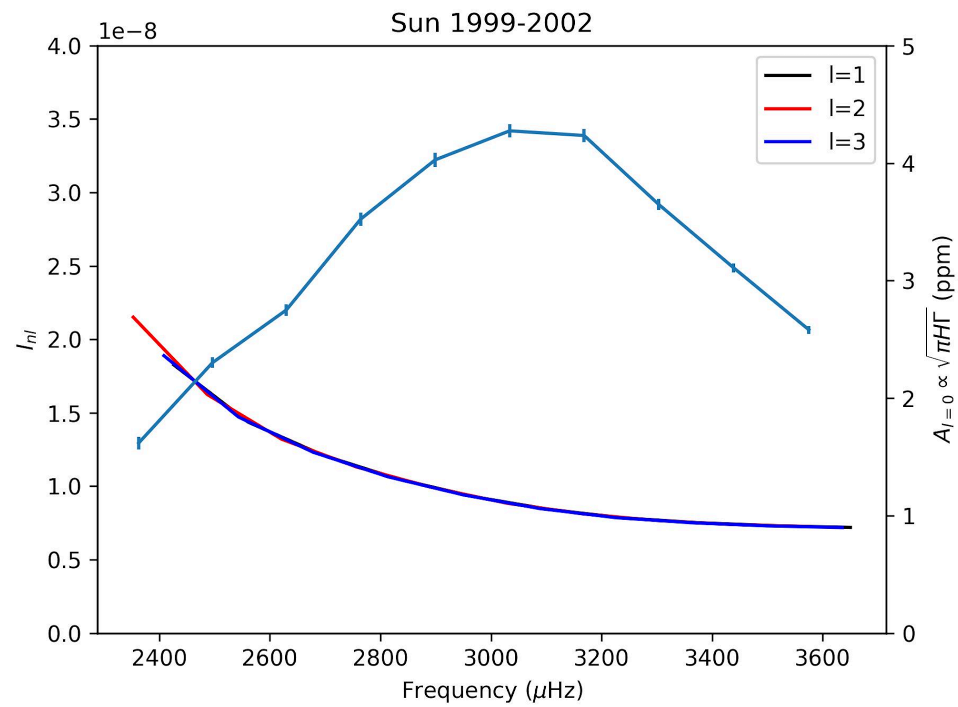

Figure 1: Inertia of low-order non-radial modes and amplitudes of radial modes in the case of the Sun. Modes with the highest power are those that will contribute the most to the determination of the global parameters (a-coefficients and stellar inclination)

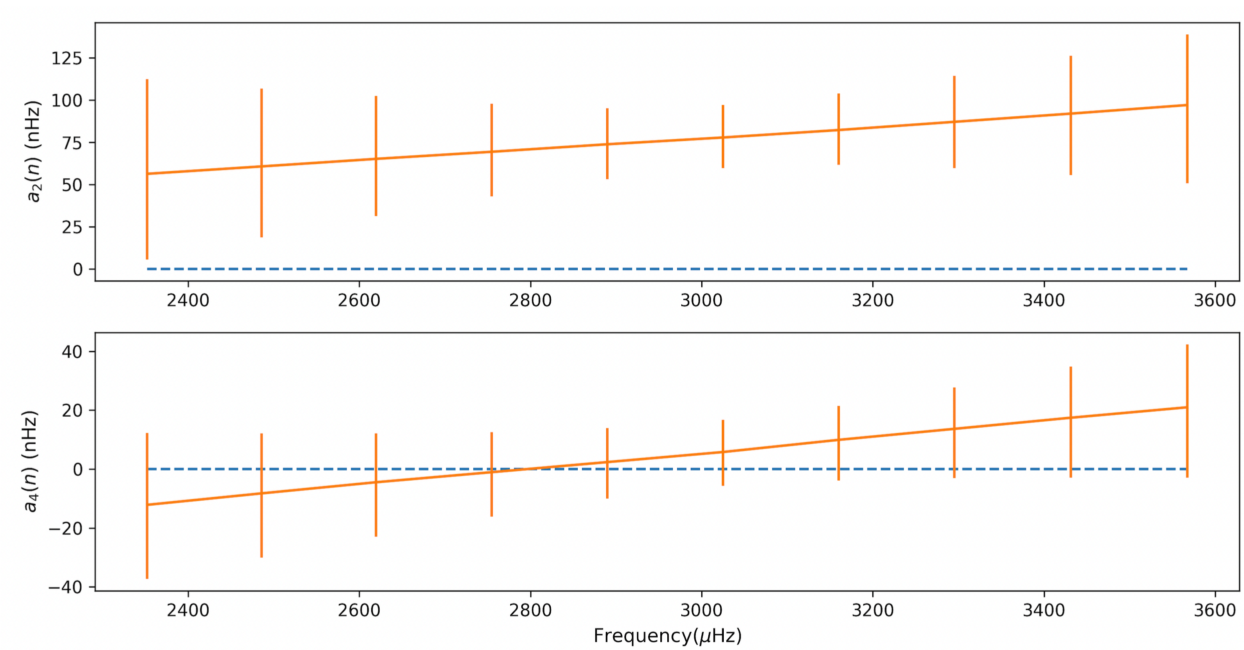

Figure 2. <a2(nu)>_l and <a4(nu)>_l coefficients for the VIRGO SPM data between 1999-2002. Uncertainties are large so that all a-coefficients are consistent with a constant fit.

4. In the second step, the authors fit a step-wise function of latitude to the a-coefficients to determine the location of the activity belts. While these fits provide estimates of the discontinuities of this function with high precision, such a model is unrealistic, and inconsistent with the actual distribution of solar activity shown in Fig.2. Therefore, there is limited astrophysical value in such estimates.

We added a triangular and gaussian shape. The triangular shape is indeed a good approximation of the activity zone, as illustrated in the new Figure 2. Now all shapes (gate, triangle, gaussian) gives similar results, although with larger uncertainty for the triangular and gaussian case. This leads to an update of many of the figures of the paper that show the PDF along with an update of the Table 3. We emphasize that this does not in fact affect the significance of the activity detection, which, as explain in the new version of the paper is mostly the same since all these models have the same number of parameters and very similar priors.

5. The argument that a more realistic Gaussian shape model did not provide a good fit is hardly acceptable.

We agree with the referee, and we were in fact very surprised by this. However, while implementing the new triangular activity zone, we found that there was a typographic error in the integration range in the Gaussian case. Once corrected, the fit became consistent with the gate function. We have modified accordingly the paper to reflect this. In particular, the section that was discussing the large inaccuracies of the Gaussian activity zone (Section 3.2) was removed.

6. In addition, it is well-known that the even splitting coefficients are directly related to the Legendre-polynomial expansion of the asphericity (e.g., Dziembowski et al., 2000; Antia et al., 2000). This means that Step 2 of the author's analysis is redundant.

It is indeed true that the Legendre-polynomials do carry information about the shape of the star. However, they do not necessarily provide an interpretation of the stellar distortion. By interpreting the a-coefficients using step 2, we add assumptions to allow the interpretation in statistical terms. Thus, we think that this is not redundant. For example, the shape of the active region is introduced. Furthermore, we introduced the hypothesis that the magnetic field is the source of the stellar distortion.

7. The latitudinal structure, determined from the a2 and a4 coefficients, is represented by two Legendre polynomials, P2 and P4. Thus, the angular resolution provided by these coefficients is insufficient for accurately determining the active solar latitudes, as shown in the cited papers.

The angular resolution of P2 and P4 is indeed low. However, as this paper shows (and Gizon 2002 also showed), the measure of P2 and/or P4 do provide a means of determining whether an activity is present in the pole, in mid-latitude or at the equator. We admit that this is not exciting when considering the Sun. However, there are several studies suggesting that some sun-like stars have an activity on the pole. This either from the spot observation (see e.g., Strassmeier 2002) or from a theoretical point of view (e.g., Yadav et al. 2014), we think that the method presented here can provide a powerful tool to diagnose active latitudes for more Kepler stars (another paper is in progress regarding this) and in the future with PLATO.