

# Comprehensive stellar seismic analysis

## New method exploiting the glitches information in solar-like pulsators

M. Farnir<sup>1</sup>, M.-A. Dupret<sup>1</sup>, S. J. A. J. Salmon<sup>1</sup>, A. Noels<sup>1</sup>, and G. Buldgen<sup>2</sup>

<sup>1</sup> Institut d'Astrophysique et Géophysique de l'Université de Liège, Allée du 6 août 17, 4000 Liège, Belgium  
e-mail: martin.farnir@uliege.be

<sup>2</sup> School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

Received 8 August 2018 / Accepted 10 December 2018

### ABSTRACT

**Aims.** We develop a method that provides a comprehensive analysis of the oscillation spectra of solar-like pulsators. We define new seismic indicators that should be as uncorrelated and as precise as possible and should hold detailed information about stellar interiors. This is essential to improve the quality of the results obtained from asteroseismology as it will provide better stellar models which in turn can be used to refine inferences made in exoplanetology and galactic archaeology.

**Methods.** The presented method – WhoSGlAd – relies on Gram-Schmidt's orthogonalisation process. A Euclidean vector sub-space of functions is defined and the oscillation frequencies are projected over an orthonormal basis in a specific order. This allowed us to obtain independent coefficients that we combined to define independent seismic indicators.

**Results.** The developed method has been shown to be stable and to converge efficiently for solar-like pulsators. Thus, detailed and precise inferences can be obtained on the mass, the age, the chemical composition and the undershooting in the interior of the studied stars. However, attention has to be paid when studying the helium glitch as there seems to be a degeneracy between the influence of the helium abundance and that of the heavy elements on the glitch amplitude. As an example, we analyse the 16CygA (HD 186408) oscillation spectrum to provide an illustration of the capabilities of the method.

**Key words.** asteroseismology – stars: oscillations – stars: solar-type – stars: abundances – methods: numerical

## 1. Introduction

Since the launch of CoRoT (Baglin et al. 2009) and Kepler (Borucki et al. 2010) missions, the scientific community has access to a tremendous amount of asteroseismic data of unprecedented quality. Such data are essential to better constrain stellar structure and evolution and, in turn, improve the characterisation of exoplanets and stellar populations. However, it is essential to develop techniques that are able to retrieve stellar parameters as accurately as possible in order to benefit from the quality of the data. A very complex problem in determining stellar parameters is the model dependency of the results. The results are intrinsically dependent on the input physics such as the equation of state as well as the opacity tables used. It therefore becomes of prime importance to develop techniques that are able to test the influence of the input physics on stellar parameters or even techniques that provide results as model independent as possible.

With such precision in the data, and the precision of the future missions TESS (Ricker et al. 2014) and PLATO (Rauer et al. 2014), studying the signature of acoustic glitches becomes a natural step towards better models. This idea was originally proposed by Vorontsov (1988) and Gough (1990). They both highlighted the effect of a sharp feature in the stellar structure on the frequencies, either directly or on the second differences. Such considerations have already been the subject of several studies. As an example of the several techniques used, Mazumdar et al. (2014) illustrate four techniques either using the second frequency differences or striving to isolate the glitch oscillation directly from the frequencies. Also, most

of the current methods focus on the localisation of the helium second ionisation zone or the base of the envelope convection zone (e.g. Monteiro et al. 2000). This is a crucial first step on the way to a better understanding of stellar physics. Indeed, a characterisation of the convective envelope extension allows to get constraints on convection itself as well as on overshooting (both its efficiency and nature). Therefore, the glitch provides the necessary observational data needed to refine current convection theories. Also, the study of the helium glitch should provide information on the helium surface content in low-mass stars. This is essential as, in such stars, it cannot be derived from spectroscopic data. Therefore, a methodology taking advantage of most of the oscillation spectrum aspects is required.

Finally, it is sometimes the case that studies use correlated constraints or discard pieces of information in their studies. For example, in method C from Verma et al. (2014), the information that is not contained in the glitch (the smooth component) is not used directly as the glitch information is isolated to draw inferences. Therefore, the usual indicators (e.g. large separation, small separation ratios, ...) are computed separately from the glitch and no method is proposed to determine properly the correlation between those indicators and that of the glitch.

For those reasons, we propose a method: Whole Spectrum and Glitches Adjustment (WhoSGlAd) that takes as much of the available spectral information as possible into account. This method defines new seismic indicators in such a way that they are as independent as possible and significant from a statistical

point of view. To do so, it relies on linear algebra via Gram-Schmidt's algorithm.

The present paper is organised as follows. We first present the method in a very general and mathematical way in Sect. 2. Then Sect. 3 defines the seismic indicators and their diagnosis power will be used to study solar-like pulsators. In the following section, we demonstrate its ability to extract and analyse the glitches signal. We also show its limitations. We present a first application to the case of 16 Cygni A (HD 186408) in Sect. 5. Let us insist on the fact that we do not present here a thorough study of 16 Cygni A but we rather show an example of the ability of the method to provide constraints on stellar structure. We conclude the paper by discussing the results and detailing future perspectives.

## 2. Method

In the present section, we describe the method we developed. It aims at using as much as possible of the information available in the oscillation spectrum of a star. Therefore, both the oscillatory and smooth part of the spectrum are simultaneously analysed in a single adjustment. This avoids multiple usage of the same information to draw different inferences. The very strength of the proposed method is that the different parameters obtained will be independent of each other, meaning that their covariance matrix will be the identity matrix. This will allow to build indicators which also are independent of each other and draw statistically relevant inferences. The independence of the parameters will be ensured by using Gram-Schmidt's (Gram 1883; Schmidt 1907) algorithm. Then, the defined seismic indicators will be used as constraints to provide improved models in the framework of forward seismic modelling (see for example Miglio & Montalbán 2005 for one of the first use of Levenberg-Marquardt's algorithm to adjust a model to seismic and non-seismic observables). Finally, such models may be used as initial guesses for inverse seismic modelling. (see Roxburgh & Vorontsov (2002a,b) for the application of the inversion technique on an artificial target, which shows the feasibility of such techniques, and Buldgen et al. (2016a,b) for examples of inversions in the case of 16 Cygni A.)

### 2.1. Mathematical description

Non-radial pulsation frequencies can be mathematically defined by three integer numbers; the radial order  $n$ , the spherical degree  $l$ , and the azimuthal order  $m$  (in this paper, we do not consider the seismic probing of rotation and consider only  $m = 0$ ). The method we developed – WhoSGLAd – is based on linear algebra in a Euclidean space. The vector space we consider is the set of  $N$  observed oscillation frequencies  $\nu_i$ . The standard deviation for each frequency is written  $\sigma_i$ . Given two frequency vectors  $\mathbf{x}$  and  $\mathbf{y}$  we define their scalar product as

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2}. \quad (1)$$

Often in asteroseismology, it is useful to compare two sets of frequencies (e.g. theoretical and observed frequencies) using a merit function defined as

$$\chi^2 = \sum_{i=1}^N \frac{(\nu_{\text{obs},i} - \nu_{\text{th},i})^2}{\sigma_i^2}, \quad (2)$$

with  $\nu_{\text{th}}$  and  $\nu_{\text{obs}}$ , the theoretical and observed<sup>1</sup> frequencies. Taking advantage of the scalar product defined above and the associated norm, this simply becomes

$$\chi^2 = \|\nu_{\text{obs}} - \nu_{\text{th}}\|^2. \quad (3)$$

In the presence of a glitch, Houdek & Gough (2007) showed that the oscillatory component in frequencies due to the glitch can be isolated from the rest of the spectrum, called the smooth component. Thus, to represent observed frequencies, we define a vector sub-space that is typically a polynomial space – the smooth component – associated with an oscillating component – the glitch –. The analytical formulation of those two components will be given in the following sections. This is very similar to what has been done by Verma et al. (2014).

The method consists in the projection of the observed and theoretical frequencies over the vector sub-space. Then, we define seismic indicators from the projections. Their definitions are given in Sect. 3. To do so, it is useful to define an orthonormal basis over the vector sub-space. This is done via Gram-Schmidt's orthogonalisation process associated with the definition of the scalar product (1). For more information about this process as well as its equivalent form as a QR decomposition, the interested reader may read Appendix D.

If we write  $j$  and  $j_0$  the indices associated with the basis elements,  $\mathbf{p}_j$  the former basis elements,  $\mathbf{q}_{j_0}$  the orthonormal basis elements, and  $R_{j,j_0}^{-1}$  the transformation matrix, we have:

$$\mathbf{q}_{j_0}(n, l) = \sum_{j \leq j_0} R_{j,j_0}^{-1} \mathbf{p}_j(n, l), \quad (4)$$

where the dependence in  $n$  and  $l$  translates that the basis elements are evaluated at each observed value of the radial order  $n$  and the spherical degree  $l$ .

It is essential that the projections be done in a specific order to obtain the lowest possible value of the merit function. This will be the subject of the following subsection. Finally, we write  $a_j = \langle \nu | \mathbf{q}_j \rangle$  the projections of the frequencies over the basis elements. The fitted frequencies are then given by

$$\nu_f(n, l) = \sum_j a_j \mathbf{q}_j(n, l). \quad (5)$$

Let us add that, thanks to the orthonormalisation, the standard deviations of the coefficients  $a_j$  are  $\sigma(a_j) = 1$  and they are independent (their covariance matrix is the identity).

### 2.2. Smooth component

Now that the mathematical context is given, we may detail the vector sub-space we selected to fit the smooth component of the frequency spectrum. As the set of observed radial orders and  $\sigma_i$  are usually different for each spherical degree and the smooth component depends on  $l$ , the smooth component basis elements depend on  $l$ . For each value of  $l$ , the frequencies will be projected over the different powers considered. We also point out that, for each spherical degree, the method requires at least the same number of frequencies as of powers considered in the formulation. The polynomials are then of the general form:

$$p_{lk}(n, l') = \delta_{ll'} p_k(n), \quad (6)$$

<sup>1</sup> We denote with the subscript *obs* both the observed frequencies and the frequencies derived from a reference model – which constitutes an artificial target – and we denote by the subscript *th* the frequencies that we adjust to those observations.

where  $\delta_{ll'}$  is the *Kronecker* delta which compares two spherical degrees  $l$  and  $l'$ ,  $p_k(n)$  is a polynomial in the radial order  $n$  and  $k$  represents its ordering. We note that the previously defined  $j$  is now separated into two indices, the spherical degree  $l$  and the ordered power  $k$ . For a better understanding, the spherical degree and ordering will be explicitly written for the smooth component transformation matrix as  $R_{l,k,k_0}^{-1}$ . And the orthonormal basis elements are

$$q_{lk}(n, l') = \delta_{ll'} q_{lk}(n), \quad (7)$$

which yield:

$$a_{lk} = \langle \mathbf{v} | \mathbf{q}_{lk} \rangle = \sum_n \frac{\nu(n, l) q_{lk}(n, l)}{\sigma^2(n, l)}, \quad (8)$$

where thanks to the introduction of  $\delta_{ll'}$  in Eqs. (6) and (7), the sum over  $l'$  collapses over one fixed degree  $l$ . For the smooth component, we treat separately each spherical degree, the parameters associated to a given degree only depend on the frequencies of this degree.

According to the asymptotic theory of non-radial oscillations (Gough 1986), we have the following formulation of the expected frequencies as a function of  $n$  and  $l$ :

$$\nu(n, l) \simeq \left( n + \frac{l}{2} + \epsilon \right) \Delta, \quad (9)$$

where  $\Delta = \left( 2 \int_0^{R_*} \frac{dr}{c(r)} \right)^{-1}$  is the asymptotic large frequency separation,  $c(r)$  is the adiabatic sound speed, and  $R_*$  is the radius at the photosphere of the star.

It follows that the first two polynomials in  $n$  (taken from the right hand side of Eq. (6)) used to depict the spectrum smooth component will be:

$$p_0(n) = 1, \quad (10)$$

$$p_1(n) = n. \quad (11)$$

Then, to provide the best fit to the observed frequencies, we methodically tested several combinations of powers to find the set giving the best agreement with the observations – the observations actually referring to theoretical models taken as observed stars in a set of calibrations –, hence the lowest  $\chi^2$  value. We get:

$$p_2(n) = n^2. \quad (12)$$

At this point, it is of prime importance to note that the construction of the basis via Gram-Schmidt's process will have to be done following the ordering of the degrees because it will allow us to associate the seismic indicators to the projection of the frequencies on the successive basis elements. For example, the projection of the frequencies on the 0 order polynomial corresponds to a fit to a constant value. This estimates the mean frequency value. Moreover, we did not include other degrees as the fit of the smooth component was already very good<sup>2</sup>. Also, adding higher order polynomials to the smooth component might account for some of the glitch oscillating features. It is essential to avoid such a behaviour as the definitions of the seismic indicators, and the inferences we draw from them, will be impacted.

Furthermore, we could also include a regularisation parameter  $\lambda$  (as in Verma et al. 2014, method C) in order to prevent those behaviours. This requires a new definition of the vector

sub-spaces. The vectors are now of dimension  $2N$ . The first  $N$  components are defined as before while the components from  $N+1$  to  $2N$  are 0 for both the frequencies and the glitch polynomial, and the second derivative of the polynomial for the smooth part. The vectors are now:

$$\begin{aligned} \mathbf{v} &\rightarrow \mathbf{v}' = (\nu_{i,i \leq N}, 0_{N < i \leq 2N}) \\ \text{smooth} &\rightarrow \mathbf{q}'_j = \left( \mathbf{q}_j, \frac{\partial^2 \mathbf{q}_j}{\partial n^2} \right) \\ \text{glitch} &\rightarrow \mathbf{q}'_j = (\mathbf{q}_j, \mathbf{0}). \end{aligned} \quad (13)$$

Then, we have to define the scalar product of  $\mathbf{x}$  and  $\mathbf{y}$  as

$$\langle \mathbf{x} | \mathbf{y} \rangle = \sum_{i=1}^N \frac{x_i y_i}{\sigma_i^2} + \lambda^2 \sum_{i=N+1}^{2N} x_i y_i. \quad (14)$$

Therefore, using the definition (3) with the new scalar product gives another value of the merit function. The inclusion of the regularisation parameter allows to minimise the oscillation of the smooth component as the minimisation of the merit function will lead to a minimisation of the second derivatives of the smooth component. Let us note that, for the regularisation constant to have an influence on the results, it must at least be of the order of the inverse of the frequencies standard deviation. However, we performed some tests with and without regularisation terms and it appears that the method is very stable without it. Moreover, we observed in many cases that the results were degraded when including it (see also Appendix A for an illustrative example). Therefore, it is not necessary to include these regularisation terms to properly extract the glitch in our method. As a consequence, the results presented in this paper do not include such terms. The fact that we use fewer fitting parameters than in (Verma et al. 2014; they consider polynomials up to the fourth degree whereas we only reach the second) might explain that the regularisation constant is not necessary in our case. In addition, we note that using  $\lambda = 0$  leads to the classical  $\chi^2$  fitting.

Finally, as we have three polynomials for each value of  $l$ , we built, for the smooth component, a vector sub-space of dimension  $3 \times l$  which equals 12 if we have four values for  $l$  (e.g. 0, 1, 2, 3). As hinted earlier, more than three observed frequencies of each degree are necessary to apply the developed method.

### 2.3. Glitch

The formulations used by Verma et al. (2014) and Houdek & Gough (2007) allow us to fit properly the helium and convection zone glitches but they are highly non linear with respect to the free parameters. Below is the expression from Verma et al. (2014), which we adapted:

$$\begin{aligned} \delta \nu_{g, \text{Verma}} &= \mathcal{A}_{\text{He}} \nu e^{-c_2 \nu^2} \sin(4\pi \tau_{\text{He}} \nu + \phi_{\text{He}}) \\ &+ \frac{\mathcal{A}_{\text{CZ}}}{\nu^2} \sin(4\pi \tau_{\text{CZ}} \nu + \phi_{\text{CZ}}), \end{aligned} \quad (15)$$

where the first term takes the helium glitch into account and the second, the convection zone glitch. The quantities  $\tau_{\text{He}}$ ,  $\phi_{\text{He}}$ , and  $\mathcal{A}_{\text{He}}$  represent respectively the acoustic depth of the second ionisation zone of helium, the phase of the helium glitch, and its amplitude. The same goes for the quantities  $\tau_{\text{CZ}}$ ,  $\phi_{\text{CZ}}$ , and  $\mathcal{A}_{\text{CZ}}$  in the case of the base of the envelope convective zone. Finally,  $c_2$  is the rate of decrease in amplitude of the helium glitch with the squared frequency.

<sup>2</sup> Appendix E shows that adding new elements to the set of basis functions is indeed not relevant.

Moreover, Eq. (15) is implicit since the frequency appears on the right-hand side. To adjust at best the frequencies, it is therefore necessary to use non-linear least square fitting algorithms (e.g. Levenberg-Marquardt's method, genetic algorithms, ...) which can be unstable and are very sensitive to the initial guesses on the optimal parameters.

For this reason, we decided to adopt the following *linearised* functions, expressed as a function of  $\tilde{n} = \left(n + \frac{l}{2}\right)$  for the helium glitch:

$$p_{\text{He}Ck}(\tilde{n}) = \cos(4\pi T_{\text{He}}\tilde{n})\tilde{n}^{-k}, \quad (16)$$

$$p_{\text{He}Sk}(\tilde{n}) = \sin(4\pi T_{\text{He}}\tilde{n})\tilde{n}^{-k}, \quad (17)$$

with  $k = (4, 5)$ ; and, for the convection zone glitch:

$$p_{CC}(\tilde{n}) = \cos(4\pi T_{CZ}\tilde{n})\tilde{n}^{-2}, \quad (18)$$

$$p_{CS}(\tilde{n}) = \sin(4\pi T_{CZ}\tilde{n})\tilde{n}^{-2}. \quad (19)$$

To obtain the above expressions we replaced the value of the frequency  $\nu$  by its first order approximation from the asymptotic formulation (9):  $\tilde{n}\Delta$ . We also defined  $T_{\text{He}} = \tau_{\text{He}}\Delta$  and  $T_{CZ} = \tau_{CZ}\Delta$ .

Moreover, we approximated the exponential decrease in frequency by the combination of two polynomials in  $\tilde{n}\Delta$ . The degrees  $-4$  and  $-5$  have been chosen to reproduce at best the decrease of the glitch amplitude towards high frequencies (which was described using a gaussian by Houdek & Gough 2007). To do so, we compared the polynomial formulation with the exponential one (Eq. (15)). Let us add that we adjusted the helium glitch using several sets of degrees and the method proved to be very stable and the results remained good. Both the fitted coefficients and defined indicators values were quasi unaffected by the choice of degrees. The glitch then writes

$$\delta\nu_{\text{He}}(\tilde{n}) = \sum_{k=5}^4 [c_{\text{He},k}p_{\text{He}Ck}(\tilde{n}) + s_{\text{He},k}p_{\text{He}Sk}(\tilde{n})], \quad (20)$$

$$\delta\nu_{CZ}(\tilde{n}) = [c_{CZ}p_{CC}(\tilde{n}) + s_{CZ}p_{CS}(\tilde{n})]. \quad (21)$$

Furthermore, such a formulation allows us to move on with Gram-Schmidt's process and generate orthonormal vectors to append the smooth component basis. As the glitch is generated in the superficial layers, it should not depend on  $l$ . Therefore, we defined the coefficients  $c$  and  $s$  to be independent of  $l$ . By doing so, the vector sub-space associated with the glitch is only of dimension 6 and this sub-space is used to complete the orthonormal basis over which frequencies are projected. We are thus able to write the glitch contribution to the frequencies as

$$\delta\nu_g(\tilde{n}) = \sum_{k=5}^4 [C_{\text{He},k}q_{\text{He}Ck}(\tilde{n}, T_{\text{He}}) + S_{\text{He},k}q_{\text{He}Sk}(\tilde{n}, T_{\text{He}})] + C_{CZ}q_{CC}(\tilde{n}, T_{CZ}) + S_{CZ}q_{CS}(\tilde{n}, T_{CZ}). \quad (22)$$

At this point, one should be aware to the fact that the basis functions depend on the acoustic depths of the second ionisation zone of helium and the base of the envelope convective zone, respectively – through  $T_{\text{He}}$  and  $T_{CZ}$  –. We observe that the functions depend non-linearly on the values of  $T_{\text{He}}$  and  $T_{CZ}$ . To preserve the linearity of the method, it is necessary to provide values for  $\tau_{\text{He}}$  and  $\tau_{CZ}$  and leave them unchanged to generate the basis and project the observed frequencies over it. In the case of a theoretical model, this estimation is done using the definition of the acoustic depth, that is:

$$\tau_{\text{He}/CZ} = \int_{R_*}^{r_{\text{He}/CZ}} \frac{dr}{c(r)}, \quad (23)$$

where  $R_*$  is the radius at the photosphere,  $r_{\text{He}/CZ}$  represents the radius of the helium second ionisation zone or of the base of the envelope convection zone (in practice we take the corresponding local maximum between the two local minima of  $\Gamma_1$  due to the partial ionisation of He and H for  $r_{\text{He}}$  and the last point below the surface for which  $\nabla < \nabla_{\text{rad}}$  for  $r_{CZ}$ ).

For observed data, we first generate the optimal model that does not take the glitches into account. We then fit this model for the glitches and retrieve the model values of the acoustic depths as estimators of the optimal values. Therefore, we do not provide any new means to estimate the involved acoustic depths.

Eventually, we may optimise over the values of  $\tau_{\text{He}}$  and  $\tau_{CZ}$  to get the best results. This is done through the use of Brent (1973)'s minimisation algorithm. However, it makes the problem non linear again. Furthermore, the optimised estimations of  $\tau_{\text{He}}$  and  $\tau_{CZ}$  always remain very close to the theoretical value and do not decrease significantly the  $\chi^2$  value. Also, we observed that by using an initial value of  $\tau_{\text{He}}$  different from that at the  $\Gamma_1$  maximum (e.g. at the minimum of the helium second ionisation zone) and adjusting it, we found back the value at the maximum. Moreover, we tried to find the value of  $Y_f$  giving the best agreement with the helium amplitude (defined in Sect. 3.3.1) observed in the case of 16 Cyg A at fixed values of  $\tau_{\text{He}}$  corresponding either to the second local minimum or the local maximum of the  $\Gamma_1$  profile. We noted that the difference between both values of  $Y_f$  is smaller than the standard deviation. In addition, freeing  $\tau_{\text{He}}$  does not impact the surface helium abundance retrieved in a significant way. This stems from the fact that the acoustic depth value changes at most of 10% and has little influence on the measured helium glitch amplitude. Therefore, the influence on the calculated  $Y_f$  is negligible as well. Thus we finally decided to give up that last non-linear minimisation for theoretical models.

By keeping a fully linear implementation of the spectrum fitting we guarantee the stability of the algorithm, the independence<sup>3</sup> of the parameters obtained via the projection on the orthonormal basis as well as small computation times. This is essential for it has to be included into a non-linear routine that searches for a stellar model accounting at best for the seismic and non-seismic observables.

### 3. Seismic indicators

The main advantage of the developed method is that it provides – via Gram-Schmidt's process – fitted coefficients which are independent of each other. It therefore allows us to derive seismic indicators as uncorrelated as possible. We will define the ones that we explored in the current section. To characterise those indicators, we computed their evolution along the grid of models presented in the next subsection and using the set of modes observed for 16 Cyg A.

#### 3.1. Models

The grid of models we used was computed using CLES (Scuflaire et al. 2008a) combined with LOSC (Scuflaire et al. 2008b) stellar evolution and oscillation codes. The models used the FreeEOS software (Cassisi et al. 2003) to generate the equation of state table, the reaction rates prescribed by Adelberger et al. (2011), the metal mixture of AGSS09 (Asplund et al. 2009), and the OPAL opacity table (Iglesias & Rogers 1996)

<sup>3</sup> Provided that the measurements of the frequencies are independent, which is not always the case.

combined with that of [Ferguson et al. \(2005\)](#) at low temperatures. Moreover, the mixing inside convective regions was computed according to the mixing length theory ([Cox & Giuli 1968](#)) and using the value  $\alpha_{\text{MLT}} = l/H_p = 1.82$  (where  $l$  is the mixing length and  $H_p$  the pressure scale height) that we obtained via a solar calibration. Microscopic diffusion was taken into account in the computation by using [Thoul et al. \(1994\)](#)'s routine. For each model, the temperature at the photosphere and the conditions above the photosphere are determined by using an Eddington  $T(\tau)$  relationship. The models have masses ranging from  $0.90 M_\odot$  to  $1.30 M_\odot$  by steps of  $0.01 M_\odot$  and are in the main sequence phase. Moreover, each model has an initial composition of  $Y_0 = 0.25$  and  $Z_0 = 0.016$  to remain close to the solar case. Finally, unless specified otherwise, the observed frequencies have been corrected for the surface effects using [Kjeldsen et al. \(2008\)](#)'s prescription of which the coefficients  $a$  and  $b$  have been calibrated by [Sonoii et al. \(2015\)](#).

### 3.2. Smooth component indicators

#### 3.2.1. Large separation

A commonly used indicator is the large separation which holds a local (i.e. based on the individual frequencies) and an asymptotic definition. To construct an estimator of the large separation, we will take inspiration in the asymptotic definition. In the asymptotic regime ( $n \gg l$ ), Eq. (9) is satisfied. We notice that, in this formulation,  $\Delta$  represents the slope in  $n$  of the straight line fitting at best the frequencies. Moreover, to fit the spectrum smooth component, we project the frequencies over the basis in a specific order given by the sequence of degrees (0, 1, 2). This means that keeping only the expression of order 0 will give an adjustment of the frequencies by a constant term, thus estimating the mean value. Furthermore, if we now keep the expression of first order, we adjust the frequencies to a straight line of which the slope, if we rely on Eq. (9), is  $\Delta$ . This is the most common way to define the mean large separation in seismic analyses. However, we must note that we have different basis vectors depending on the spherical degree considered. This means that, for each value of  $l$ , we will have a different regression to a straight line, therefore a different estimate of the large separation  $\Delta_l$ . Using expressions (4), (5), (7), and (8), we isolate this slope to write:

$$\Delta_l = a_{l,1} R_{l,1,1}^{-1}. \quad (24)$$

We may finally average these indicators over  $l$  to estimate at best the large separation. Knowing that the standard deviation of  $a_{l,1}$  is 1,  $(R_{l,1,1}^{-1})^2$  is the variance of  $\Delta_l$ . The weighted mean of the large separations thus yields

$$\Delta = \frac{\sum_l a_{l,1} / R_{l,1,1}^{-1}}{\sum_l 1 / (R_{l,1,1}^{-1})^2}. \quad (25)$$

Finally, we expect from [Ulrich \(1986\)](#) that  $\Delta$  should be an estimator of the mean stellar density.

#### 3.2.2. Normalised small separation

Two commonly used indicators are the small separations  $d_{01}(n)$  and  $d_{02}(n)$  of which the definitions are:

$$d_{01}(n) = (\nu(n-1, 1) - 2\nu(n, 0) + \nu(n, 1)) / 2, \quad (26)$$

$$d_{02}(n) = (\nu(n, 0) - \nu(n-1, 2)). \quad (27)$$

They allow a measurement of the spacing between the observations and the asymptotic relation (9). However, they happen to be sensitive to the surface effects. Therefore, [Roxburgh & Vorontsov \(2003\)](#) suggested to divide these expressions by the large separation in order to minimise such effects. Indeed, they showed that those ratios are almost independent of the structure of the outer layers of the star.

We thus introduce estimators of these ratios. Such ratios represent the spacing between ridges of spherical degrees 0 and 1 for Eq. (26) and degrees 0 and 2 for Eq. (27) in the échelle diagram ([Grec et al. 1983](#)). In a more general way, we approximate the mean difference between the ridges of spherical degrees 0 and  $l$  by comparing the mean values of the frequencies for those degrees. That is  $(\bar{\nu}_0 - \bar{\nu}_l) / \Delta_0$ . Assuming expression (9) to be exact, this difference is  $\bar{n}_0 + \epsilon_0 - (\bar{n}_l + \epsilon_l + l/2)$ . Then, we added  $-\bar{n}_0 + \bar{n}_l + l/2$  to the expression to make its value come close to  $\epsilon_0 - \epsilon_l$ . We then obtained the following expression:

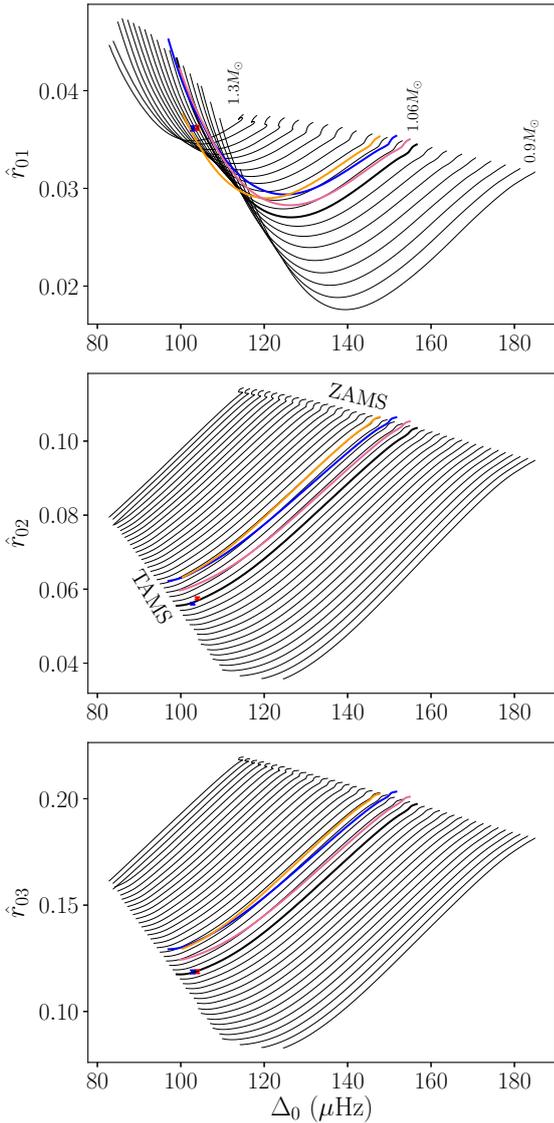
$$\hat{r}_{0l} = \frac{\bar{\nu}_0 - \bar{\nu}_l}{\Delta_0} + \bar{n}_l - \bar{n}_0 + \frac{l}{2}, \quad (28)$$

where  $\bar{\nu}_l$  and  $\bar{n}_l$  are respectively the weighted mean values of  $\nu(n, l)$  and of  $n$  for the spherical degree  $l$  in accordance with the definition of the scalar product. In addition, the mean value  $\bar{\nu}_l$  equals  $a_{l,0} R_{l,0,0}^{-1}$  as it is the fitting of the frequencies of degree  $l$  to a constant value. Finally, we stress that the above expression is slightly different from Eqs. (26) and (27) as they represent the local spacing between ridges in the échelle diagram and Eq. (28) corresponds to the mean spacing. Figure 1 shows the evolution of those indicators along the grid presented in Sect. 3.1<sup>4</sup> for the set of modes observed in 16 Cygni A. The  $x$ -axis is the large separation of spherical degree 0 as we defined above and the  $y$ -axis is the considered indicator. We also display the observed values for 16 Cygni A (HD 186408) using the frequencies determined by [Davies et al. \(2015\)](#). In blue is the observed value and, in red, the value corrected for the surface effects according to [Kjeldsen et al. \(2008\)](#)'s prescription. We note that these indicators are almost insensitive to surface effects excepted for the case of  $\hat{r}_{02}$ , which value is changed by about  $1\sigma$ . We only show the standard deviation for the estimators of the small separation as the one for the large separation is too small to be visible on the plot. Indeed, we computed a standard deviation for  $\Delta_0$  of  $\sigma(\Delta_0) = 5 \times 10^{-3} \mu\text{Hz}$ .

To provide a comparison, we computed the evolution of the ‘‘usual’’ indicators along the same tracks as in Fig. 1 and display it in Fig. B.1. We observe that the new indicators exhibit the same behaviour as the usual ones and provide smaller standard deviations, therefore, tighter constraints.

We note on Fig. 1 that  $\hat{r}_{02}$  is a very good indicator of the core conditions and should hold information about the evolutionary stage on the main sequence as its evolution is almost monotonic. It is therefore very similar to the small separation that has been shown to carry information on the evolution ([Christensen-Dalsgaard 1988](#)) as it is sensitive to the sound speed gradient which in turn is sensitive to the chemical composition changes. On the other hand,  $\hat{r}_{01}$  is not a good indicator of the evolution but carries additional information. For example, [de Meulenaer et al. \(2010\)](#) showed that, for stars with masses and metallicities close to that of  $\alpha$  Centauri A (HD 128620), it should provide an upper limit on the amount of convective-core overshooting. We also draw attention to the fact that  $\hat{r}_{01}$  shows a turn off for evolved stars. This means that we have to be cautious when fitting models to the observations as, for specific sets of input physics, there

<sup>4</sup> However, in the case of the indicator  $\hat{r}_{01}$  we used a  $0.02 M_\odot$  step for a better visibility.



**Fig. 1.** Seismic HR diagram defined with the new indicators  $\hat{r}_{0l}$  and computed along the grid presented in Sect. 3.1. The masses increase from right to left. The blue marker shows the observed value for 16 Cyg A while the red one shows the value corrected for the surface effects following Kjeldsen et al. (2008)’s prescription. The thick line represents the track for  $1.06 M_{\odot}$ . The blue line has been computed for  $Y_0 = 0.27$ , the orange one for  $\alpha_{\text{MLT}} = 1.5$ , and the pink one for  $(Z/X)_0 = 0.018$ . All the coloured lines have been computed for  $1.06 M_{\odot}$ .

exists inaccessible regions. This should allow to constrain the input physics.

Finally,  $\hat{r}_{03}$  does not provide new information. In addition, we may observe from the comparison between the observations and the theoretical tracks that the expected masses retrieved from the different indicators are in agreement. Indeed, from  $\hat{r}_{01}$ , we should expect masses between  $1.06 M_{\odot}$  and  $1.11 M_{\odot}$ . Then, from  $\hat{r}_{02}$ , the expected values are in between  $1.06 M_{\odot}$  and  $1.07 M_{\odot}$ . Again,  $\hat{r}_{03}$  does not add some information as the values range from  $1.05 M_{\odot}$  to  $1.06 M_{\odot}$ . As a consequence, we would expect the mass of 16 Cygni A to be around  $1.06 M_{\odot}$ . This value has been highlighted on the figure by using a thick line. However, we must observe that this does not provide a precise estimate of 16 Cyg A mass as we only tested a specific chemical composition –  $Y_0 = 0.25$ ,  $Z_0 = 0.016$  – as well as given choice

of  $\alpha_{\text{MLT}} = 1.82$ . A proper adjustment is needed to draw conclusions. Those values are only given to illustrate the compatibility between the different indicators. We also show the influence of the composition and  $\alpha_{\text{MLT}}$  on the coloured tracks. The blue line is for a higher initial helium abundance. Then, the orange line depicts the influence of a lower value of  $\alpha_{\text{MLT}}$ . Finally, the pink line shows how a higher metallicity modifies the results. We observe in all cases that the inferred mass should be lower.

### 3.2.3. $\Delta_{0l}$ indicators

As it has been shown in several studies, the combination of the small separation ratios  $r_{01}$  and  $r_{10}$  first introduced by Roxburgh & Vorontsov (2003) into  $r_{010}$  allows to provide inferences about the stellar central mixed region extension (Popielski & Dziembowski 2005; Deheuvels et al. 2010; Silva Aguirre et al. 2011). Indeed, the mean value and slope of this indicator occupy very specific regions in the parameter space according to the extent of the central mixed region. This should therefore provide constraint on the amount of overshooting necessary to reproduce observations. As an example, an extensive study of several *Kepler* targets has been realised by Deheuvels et al. (2016) who have been able to provide constraints on the overshooting parameter  $\alpha_{\text{ov}}$  for eight of those targets. In the framework of this paper, the indicator  $\hat{r}_{01}$  represents an estimator of the mean value of  $r_{010}$ . We may therefore build an indicator for its slope as follows:

$$\Delta_{0l} = \frac{\Delta_l}{\Delta_0} - 1, \quad (29)$$

with  $\Delta_l$  the large separation for modes of spherical degree  $l$  defined by Eq. (24). It is straightforward to show that, in the asymptotic regime,  $\Delta_{01}$  indeed represents the slope of the frequency ratio  $r_{01}$ .

As for the small separation indicators, we computed the evolution of such indicators along the grid of models presented in Sect. 3.1 to demonstrate their regularity and validity. This is shown in Fig. 2. We also display the observed value for the case of 16 Cygni A. This value is corrected for the surface effect using Kjeldsen et al. (2008)’s prescription. As in Fig. 1, the error on  $\Delta_0$  is too small to be visible. Moreover, we show the influence of a change in the composition and of  $\alpha_{\text{MLT}}$  with the coloured tracks. The colours are the same as in Fig. 1.

Again, it is possible to get an estimate of the mass value for the given composition and physics. From  $\Delta_{01}$ , we expect masses ranging from  $1.06 M_{\odot}$  to  $1.07 M_{\odot}$ . Then, from  $\Delta_{02}$ , we expect that they lie between  $1.05 M_{\odot}$  and  $1.07 M_{\odot}$ . Finally, from  $\Delta_{03}$ , the mass should be between  $1.00 M_{\odot}$  and  $1.05 M_{\odot}$ . This time, we observe a slight incompatibility between the first two indicators and the last one. Let us add that we highlighted the value of  $1.06 M_{\odot}$  as in Fig. 1. Moreover, we note that, as opposed to the small separation ratio indicators, the relative behaviours of the coloured tracks are different for the three indicators. This could allow to discriminate various choices in the physics of the models considered as they represent different values of  $\alpha_{\text{MLT}}$ ,  $Y_0$ , and  $Z/X_0$  and also to solve the slight mass discrepancy observed.

Finally, as detailed above, the simultaneous use of both  $\hat{r}_{01}$  and  $\Delta_{01}$  allows to provide estimations of the extent of stellar mixed cores. To illustrate this, we plotted main sequence evolutionary tracks for a  $1.2 M_{\odot}$ ,  $Y_0 = 0.25$  and  $X_0 = 0.734$  with several overshooting parameter values ranging from 0.005 to

0.3. The overshooting parameter gives the extent of the mixed core above the Schwarzschild limit through  $d = \alpha_{\text{ov}} \min(H_p, h)$  where  $H_p$  is the pressure scale height and  $h$  the thickness of the convection zone. This is shown in Fig. 3. The ZAMS is at the converging point of the tracks and the TAMS is at their end. We observe that the various tracks occupy very specific regions in the  $(\Delta_{01}, \hat{r}_{01})$  diagram. Moreover, we note the striking resemblance of Fig. 3 and Fig. 3 of Deheuvels et al. (2016). This should allow us to constrain the amount of overshooting. However, Deheuvels et al. (2016) noted that for this diagnostic tool to be efficient, the mean large separation of the target should not exceed  $\sim 110 \mu\text{Hz}$  and mixed modes should not be present in the oscillation spectrum.

### 3.2.4. $\hat{\epsilon}$

To provide an estimate of  $\epsilon$ , we define the following vector subspace, where frequencies are described as

$$v(n, l) = \left(n + \frac{l}{2} + \epsilon\right) \hat{\Delta} = \left(n + \frac{l}{2}\right) \hat{\Delta} + K, \quad (30)$$

where  $\hat{\Delta}$  and  $K$  are free parameters.

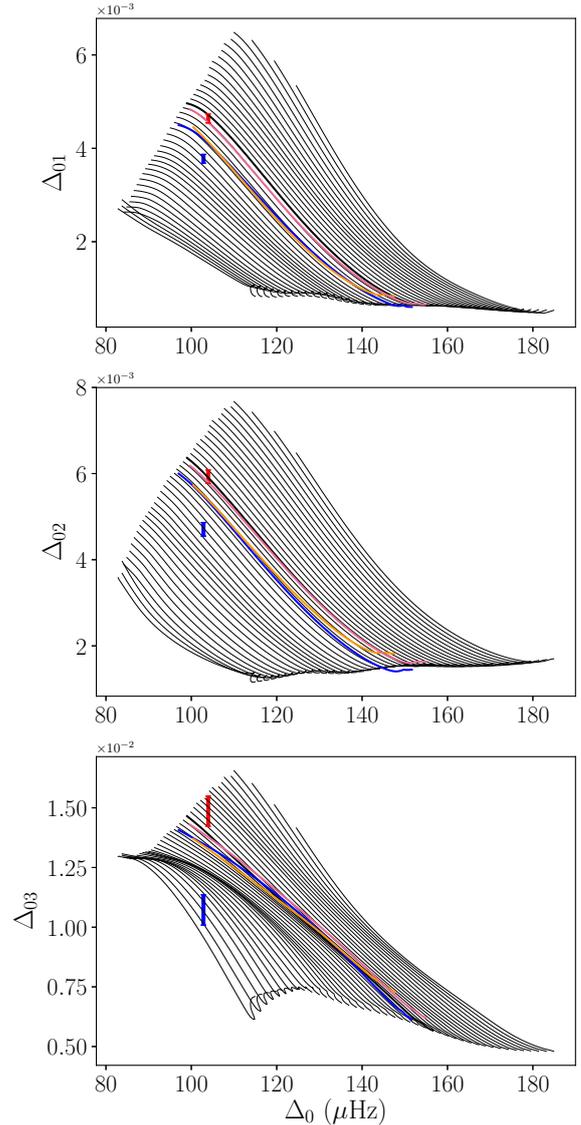
Then, we define an orthonormal basis over this sub-space:  $\tilde{q}_0$  and  $\tilde{q}_1$ . Finally, by projection of the frequencies over this basis and identification of the several coefficients with the asymptotic formulation, we can retrieve an expression for  $\hat{\epsilon}$ , the estimator of  $\epsilon$ . We note that this projection also provides an expression to estimate the large separation that is different from Eq. (25).

Figure 4 shows the evolution of the indicator  $\hat{\epsilon}$  along the grid presented in Sect. 3.1 but with a step of  $0.02 M_\odot$ . We may observe that this indicator is almost insensitive to the mass for the early stages of the main sequence. The influence of the mass only becomes visible when the stars become older. Moreover, the red marker in Fig. 4 shows the observed value for 16 Cygni A. We note a disagreement between theoretical and observed data. This disagreement can however be tackled by correcting the observed frequencies for the surface effects. This is what the blue and green markers represent. For the blue one, we have computed a correction to the surface effects following Kjeldsen et al. (2008)'s prescription. Then, for the green one, we have computed the correction prescribed by Sonoi et al. (2015). Therefore, it seems reasonable that the indicator we defined could be of some use to constrain the surface effects. It could provide a complementary method to that of Roxburgh (2016). Indeed the method presented uses differences between observed and model  $\epsilon$  values – under the hypothesis that both the model and observed star have the same inner structure – to isolate only the surface contribution to the measured frequencies. This allows to account for surface effects without the need of empirical corrections. On the other hand, the present indicator should allow to discriminate several surface effects corrections without the need of any physical assumption.

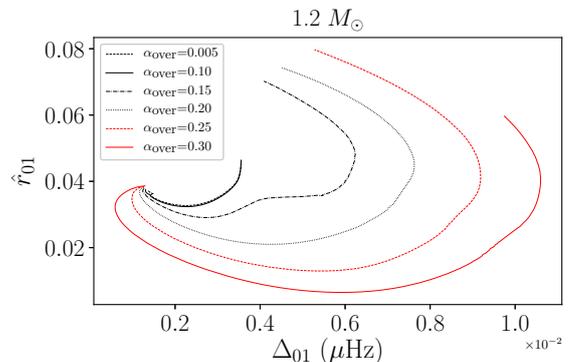
## 3.3. Glitch indicators

### 3.3.1. Helium amplitude

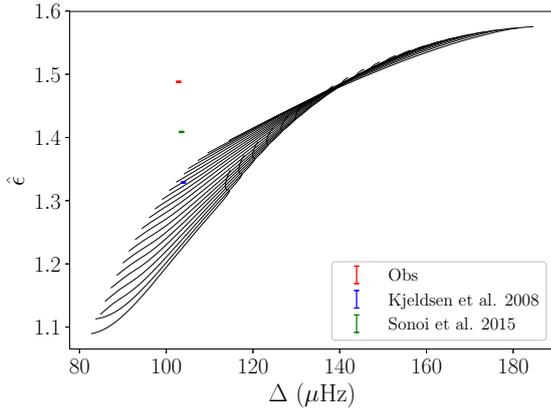
With the aim of retrieving the photospheric helium abundance, we built an indicator of the helium glitch amplitude. Verma et al. (2014) obtain their indicator via an integration of the glitch amplitude over the spectrum. We prefer taking advantage of the



**Fig. 2.** Seismic HR diagram defined with the new indicators  $\Delta_{0l}$  and computed along the grid presented in Sect. 3.1. The masses increase from right to left. The red marker shows the observed value corrected for the surface effects following Kjeldsen et al. (2008)'s prescription. The thick line represents the track for  $1.06 M_\odot$ . The blue line has been computed for  $Y_0 = 0.27$ , the orange one for  $\alpha_{\text{MLT}} = 1.5$ , and the pink one for  $(Z/X)_0 = 0.018$ . All the coloured lines have been computed for  $1.06 M_\odot$ .



**Fig. 3.** MS evolutionary tracks in the  $(\Delta_{01}, \hat{r}_{01})$  plane for models of  $1.2 M_\odot$  with  $Y_0 = 0.25$  and  $X_0 = 0.734$  for several values of the overshooting parameter  $\alpha_{\text{ov}}$  shown in the legend.



**Fig. 4.** Evolution of  $\hat{\epsilon}$  along the grid presented in Sect. 3.1. The masses increase from top to bottom. The step is here of  $0.02 M_{\odot}$ . The red, green, and blue markers respectively represent the observation for 16 Cygni A, the observation for 16 Cygni A corrected for the surface effects using Kjeldsen et al. (2008)’s prescription and the one corrected using Sonoii et al. (2015)’s prescription.

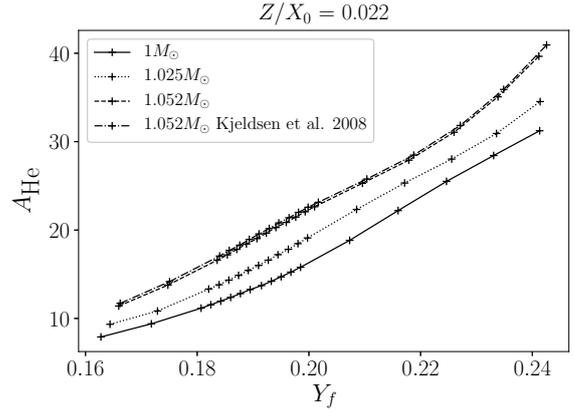
scalar product and define the indicator as the norm of the helium glitch component, thus generating the following expression:

$$A_{\text{He}} = \sqrt{C_{\text{He},5}^2 + S_{\text{He},5}^2 + C_{\text{He},4}^2 + S_{\text{He},4}^2} \quad (31)$$

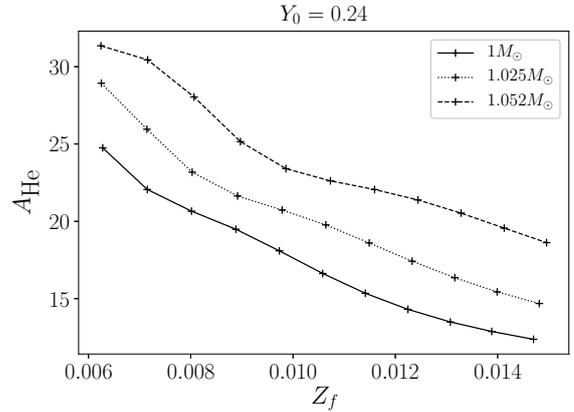
Thanks to the orthonormalisation, it is independent of the other indicators and its standard deviation equals 1. We calculated the evolution of this indicator with respect to the surface helium mass fraction  $Y_f$  and the surface mass fraction of metals  $Z_f$ . This is shown in Figs. 5 and 6. Figure 5 has been computed for stars with a fixed  $(Z/X)_0$  ratio of 0.022 for several surface helium mass fractions, displayed on the abscissa, and for the three values of the mass shown in the legend. To have a reference, we imposed the models to have a fixed value of the large separation – the one observed for 16 Cyg A. In a similar way, Fig. 6 has been computed for a fixed value of  $Y_0 = 0.24$  and for several masses shown in the legend. The helium glitch amplitude has then been computed for the values of  $Z_f$  displayed on the abscissa. We insist on the fact that these tracks do not represent the evolution of the helium amplitude with the surface composition along the evolution of a given model. Instead, each point corresponds to a given stellar model that fits the observed 16 Cygni A large separation for a given surface composition. This means that those models were not selected from the grid presented in Sect. 3.1.

We observe in Fig. 5 an increasing trend in the helium glitch amplitude with the helium mass fraction as well as with the mass. This has to be expected as a larger quantity of helium inside the star would lead to a more important depression of the first adiabatic index  $\Gamma_1$  at the second ionisation zone of helium, and, therefore, a glitch of greater amplitude. Moreover, we also show the influence of the surface effects by computing the amplitude evolution for a  $1.052 M_{\odot}$  star for which the surface effects have been taken into account via Kjeldsen et al. (2008)’s prescription (dot-dashed line). It is apparent that they have little influence on the amplitude as the values remain in the  $1\sigma$  error bars of the uncorrected models. This was expected as the glitch is of greater amplitude in the low frequencies regime while the surface effects corrections are greater in the high frequencies regime.

Furthermore, as shown in Fig. 6, the glitch amplitude and the metallicity are anti-correlated. This corroborates Basu et al.



**Fig. 5.** Evolution of the helium glitch amplitude  $A_{\text{He}}$  with the surface helium abundance  $Y_f$ . Each track corresponds to a given mass, written in the legend. The dot-dashed line represents the amplitude for a  $1.052 M_{\odot}$  model of which the frequencies have been corrected for surface effects as in Kjeldsen et al. (2008). Every model has an initial heavy elements abundance of  $(Z/X)_0 = 0.022$ . Each point has been computed with the same large separation to remain at the same evolutionary stage.



**Fig. 6.** Evolution of the helium glitch amplitude  $A_{\text{He}}$  with the surface heavy elements abundance  $Z_f$ . Each track corresponds to a given mass labeled in the legend.

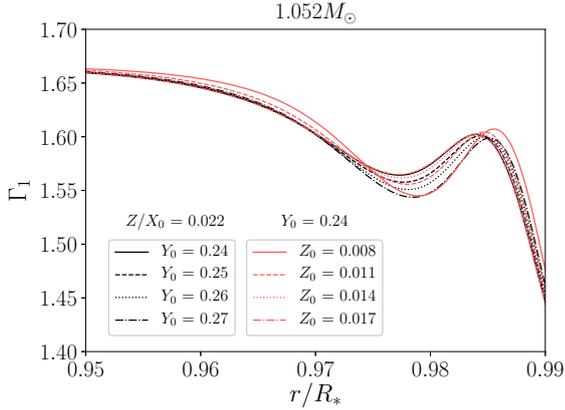
(2004)’s observations. Therefore, we are facing a degeneracy and the glitch amplitude alone will not be sufficient to estimate properly the surface helium mass fraction. Thus, the smooth component indicators defined above will be of great help. This clearly shows that the  $A_{\text{He}}-Y_f$  relation is model dependent which should never be forgotten.

Figure 7 illustrates this degeneracy. It represents the profile of  $\Gamma_1$  as a function of the reduced radius in the superficial layers of stars of a fixed large separation but with several chemical compositions. We immediately notice that both an increase of the surface helium abundance and a decrease of the surface heavy elements abundance lead to a minimum that increases in magnitude. Therefore, the helium glitch amplitude becomes greater as well. We provide an interpretation of this phenomenon in Sect. 4.2.

### 3.3.2. Convection zone amplitude

The definition of the envelope convective zone glitch amplitude we provide is very similar to that of the helium glitch and is the following:

$$A_{\text{CZ}} = \sqrt{C_{\text{CZ}}^2 + S_{\text{CZ}}^2}. \quad (32)$$



**Fig. 7.** Evolution of  $\Gamma_1$  as a function of the reduced radius in the superficial layers of stars of fixed large separation. Every star has a mass of  $1.052 M_\odot$ . Two families of curves are displayed. The black ones have a fixed value of the ratio between the initial hydrogen and heavy elements abundances of  $(Z/X)_0 = 0.022$  and variable initial helium abundances. The red ones have a fixed initial helium abundance of  $Y_0 = 0.24$  and a variable initial heavy elements abundance. The different values are displayed on the figure.

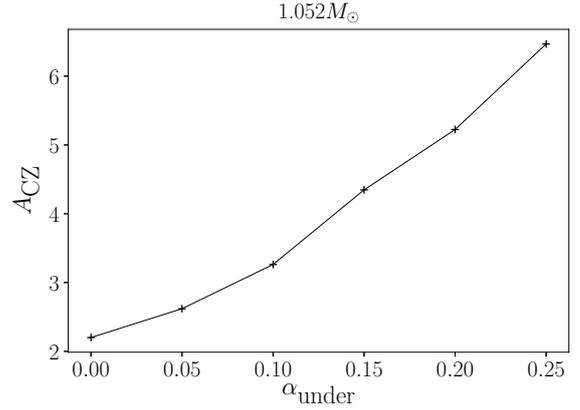
We expect this indicator to be a proxy of the sharpness of the transition between the envelope convective zone and the radiative zone. Again, thanks to the orthonormalisation, it is independent of the other indicators and its standard deviation is equal to 1. We present in Fig. 8 its evolution with the importance of the undershooting, characterised by the coefficient  $\alpha_{\text{under}}$ . This coefficient determines the size  $d$  of the undershooting region at the bottom of the convective envelope. This size is given by  $d = \alpha_{\text{under}} \min(H_p, h)$  where  $H_p$  is the pressure scale height and  $h$  the thickness of the convection zone. In the undershoot region, the temperature gradient is set to the adiabatic one and the mixing is assumed to be instantaneous.

We expect an increase of  $A_{\text{CZ}}$  with  $\alpha_{\text{under}}$  as the introduction of undershooting in a stellar model will create a discontinuity of the temperature gradient. The temperature gradient jump increases with the value of  $\alpha_{\text{under}}$ . Thus, the glitch amplitude increases as well. This is what we show in Fig. 8 where the computed models have a mass of  $1.052 M_\odot$ , an initial hydrogen mass fraction of  $X_0 = 0.744$ , and an initial metal mass fraction of  $Z_0 = 0.016$ . As for the case of the helium amplitude evolution, we kept a fixed large separation – which is that of 16 Cyg A – for each model. Moreover, we noted that each computed model was at the same evolutionary stage (constant central hydrogen mass fraction). Therefore, the observed effect is not evolutionary but rather the effect of the temperature gradient discontinuity as expected. Finally, we may add that, when setting the temperature gradient to the radiative one in the undershoot region, we do not observe any significant trend in the amplitude with  $\alpha_{\text{under}}$ . We may thus conclude that we observe the effect of the temperature gradient and not of the chemical composition.

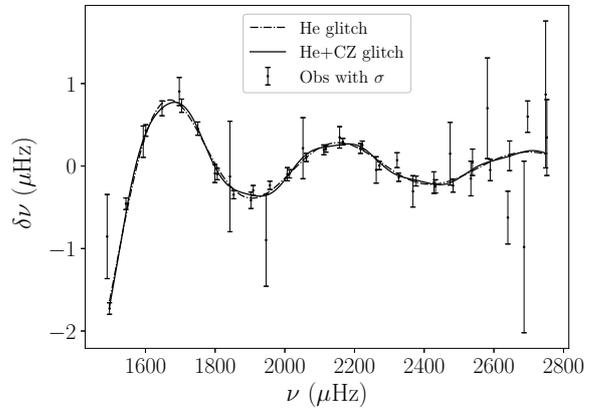
## 4. Method characterisation

### 4.1. Capabilities

For the observed data, we have chosen the frequencies computed by Davies et al. (2015) for the component A of the binary system 16 Cygni (HD 186408). We come back to the particular case of 16 Cyg A in Sect. 5. Figure 9 shows the difference



**Fig. 8.** Evolution of the convection zone glitch amplitude  $A_{\text{CZ}}$  with  $\alpha_{\text{under}}$  for a star of mass  $1.052 M_\odot$ ,  $X_0 = 0.744$ ,  $Z_0 = 0.016$  and at a constant value of the large separation.



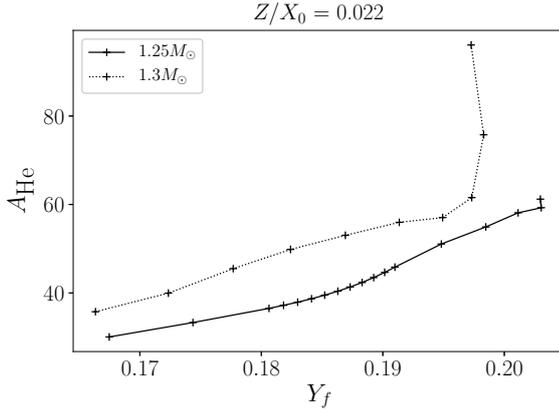
**Fig. 9.** Fitted glitch to 16CygA (HD 186408) data (Davies et al. 2015). Only the  $l = 0$  fitted curves are displayed.

between the observed frequencies and the smooth component of the fitted frequencies, the fitted helium glitch alone and the fitted helium and convection zone glitches. Only the  $l = 0$  fitted curves are displayed. We observe that the fit is good and that the helium glitch has been properly isolated. However, the convection zone glitch is of very low amplitude, compared to the helium glitch, and has a negligible contribution. This is visible in the negligible improvement of the  $\chi^2$  value from the results without including the convective zone glitch (about 10% variation). These results are similar to those of Verma et al. (2014).

### 4.2. Limitations

The presented method has been developed for the study of solar-like pulsators and to provide a comprehensive analysis of their oscillation spectra. However, it is not yet adapted to study evolved stars which exhibit mixed modes.

In addition, from masses around  $1.25 M_\odot$  and above as well as for the highest values of the helium abundances considered (from  $Y_f \sim 0.195$  and above), the evolution of the helium glitch amplitude with the surface helium mass fraction is less monotonic and inferences become unreliable. Figure 10 shows this limitation. Indeed, for  $1.25 M_\odot$  we observe a very sharp increase of the helium amplitude as we defined in Sect. 3.3.1 for a quasi constant surface helium mass fraction. This corresponds to a

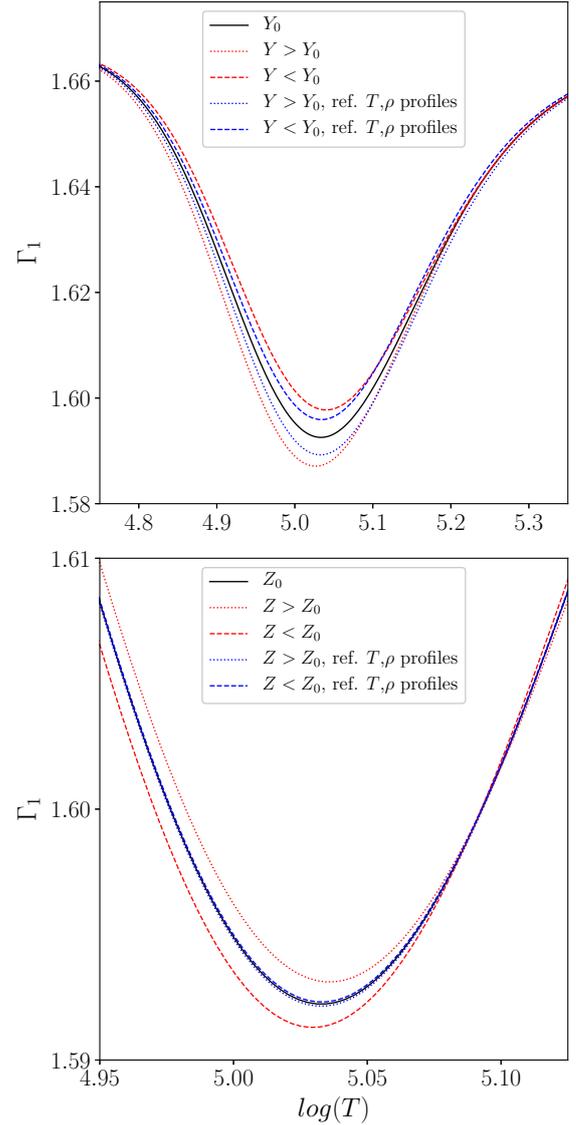


**Fig. 10.** Evolution of the helium glitch amplitude  $A_{\text{He}}$  with the surface helium abundance  $Y_f$  for high masses.

decrease of the surface  $Z/X$  ratio which is shown in Sect. 3.3.1 to lead to a higher amplitude.

Also, we observe for  $1.3 M_{\odot}$  that the last point ( $Y_f \sim 0.197$  and  $A_{\text{He}} \sim 96$ ) moves backwards. This is due to microscopic diffusion. We observe that the size of the convective envelope decreases with the initial helium abundance. This leads to a more efficient gravitational settling as the diffusion velocities are greater close to the surface. Therefore, the surface abundance becomes smaller than for the previous point that had a lower initial helium abundance. We observe the same decrease in the surface metallicity which explains the increase of the amplitude even though the surface helium mass fraction remained constant.

We also noted in Sect. 3.3.1 that the surface abundance of metals has a relevant influence on the helium glitch amplitude and separating its contribution from that of the helium is not an easy task. To investigate such a behaviour, we developed a toy model for the first adiabatic index  $\Gamma_1$  which is thought to be the main contributor to the helium glitch amplitude (Gough 1990; Houdek & Gough 2007). In this model, we trace back the influence of the helium and metals abundances on the dip of  $\Gamma_1$  in the helium second ionisation zone. More information about the construction of this model is given in Appendix C. Using it, we were able to test the influence of the chemical composition decoupled from its evolutionary effect on the temperature and density profiles. To do so, we artificially modified the chemical composition profile of a reference model (black curve in Fig. 11), without changing its temperature and density profiles, to match the surface abundance, in either metals or helium, of a second reference model. This way, we were able to isolate the contribution of the chemical composition alone. Figure 11 shows the comparison between the helium second ionisation zone toy models with modified chemical composition profiles only (blue curves) and toy models with the same composition but for which its effect on temperature and density have been taken into account (red curves). We observe in the top panel that the effect of the helium abundance dominates over the effect of temperature. In the bottom panel, we observe that the metal abundance alone does not modify by a significant amount the dip in the helium second ionisation zone. However, when taking its influence on the temperature and density profiles into account, the effect becomes significant. This allows to better understand the degeneracy in composition on the helium amplitude. One effect, that of the helium abundance is direct on the shape of the  $\Gamma_1$  profile in the helium second ionisation zone while the second, that of the metal abundances, is indirect as it influences the



**Fig. 11.** Comparison with a reference model (black curve) of the toy model  $\Gamma_1$  profiles in cases for which the temperature and density profiles have been decoupled from the composition profiles (blue curves) and coupled cases (red curves). The dashed lines have a common decreased abundance of the considered element and the dotted lines have an increased abundance. *Top panel:* effect of the helium abundance; *bottom panel:* effect of the metals abundance.

temperature and density profiles which in turn modify the  $\Gamma_1$  profile.

Finally, we observe that, in a  $\ln P - \ln T$  diagram, the curves for the different models in the temperature region of the  $\Gamma_1$  dip are parallel to each other but at various height. A higher  $\ln P - \ln T$  curve corresponds to a shallower  $\Gamma_1$  dip. We also observe this behaviour with a fixed composition and a variable mass. Higher masses models have lower curves and deeper depressions in  $\Gamma_1$ . Therefore, the toy model allows to understand the influence of the mass on the helium glitch amplitude as well.

## 5. Illustration with 16 Cygni A (HD 186408) observations

16 Cyg A (HD 186408) is one of the brightest stars in the *Kepler* field of view. It belongs to a binary system of solar analogs, both exhibiting solar-like pulsations. The quality and length of the

**Table 1.** Observed seismic indicators.

Indicator	Value		$\sigma$
	Kjeldsen	Sonoi	
$\Delta$ ( $\mu\text{Hz}$ )	104.088	103.611	0.005
$A_{\text{He}}$	30.4	30.1	1.0
$A_{\text{CZ}}$	2.2	1.5	1.0
$\bar{\epsilon}$	1.3288	1.4086	0.0009
$\hat{r}_{01}$	0.0362	0.0362	0.0002
$\hat{r}_{02}$	0.0575	0.0561	0.0003
$\hat{r}_{03}$	0.1187	0.1184	0.0008
$\Delta_{01}$	$4.6 \times 10^{-3}$	$3.8 \times 10^{-3}$	$0.1 \times 10^{-3}$
$\Delta_{02}$	$5.9 \times 10^{-3}$	$4.8 \times 10^{-3}$	$0.1 \times 10^{-3}$
$\Delta_{03}$	$14.9 \times 10^{-3}$	$10.6 \times 10^{-3}$	$0.6 \times 10^{-3}$

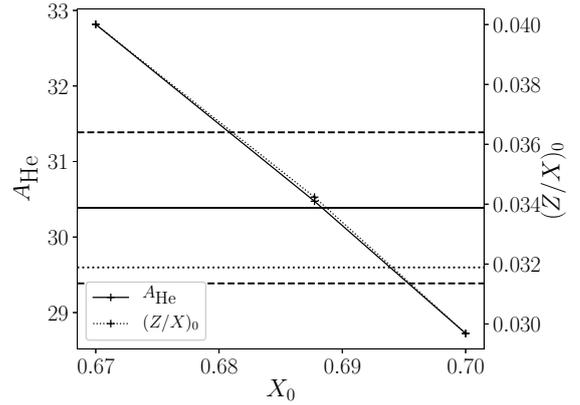
**Table 2.** Adjusted stellar parameters.

Quantity	Value		$\sigma$	
	Kjeldsen	Sonoi		
	$n^0, n, n^2$	$n^{-1}$		
$M(M_{\odot})$	1.06	1.06	1.06	0.02
$R(R_{\odot})$	1.218	1.219	1.223	0.001
age (Gyr)	6.8	6.9	7.1	0.1
$X_0$	0.684	0.697	0.685	0.010
$(Z/X)_0$	0.035	0.031	0.036	0.002
$Y_{\text{f}}$	0.242	0.232	0.240	0.028
[Fe/H]	0.188	0.131	0.199	0.03

collected time series make it the ideal subject to test the method. It should be noted that we do not provide a detailed study of 16 Cygni A. Indeed, we only present here the capability of the method to provide structural constraints. A detailed study will be the object of a future paper.

### 5.1. Methodology

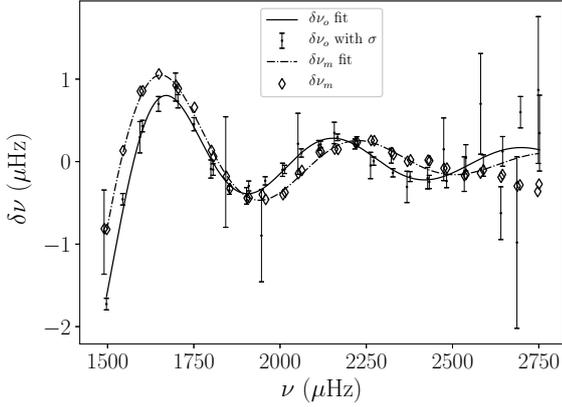
To obtain constraints on 16 Cyg A, we compute the value of the seismic indicators using the frequencies determined by Davies et al. (2015). We have corrected the surface effects for the observed frequencies by using the power law prescribed by Kjeldsen et al. (2008) and the  $a$  and  $b$  coefficients fitted by Sonoi et al. (2015) as a function of  $T_{\text{eff}}$  and  $g$ . The authors have done this coefficient adjustment by comparing the adiabatic frequencies of patched models based on 3D simulations and that of unpatched standard 1D models. We then fit the observed values of  $\Delta$ ,  $\hat{r}_{01}$ ,  $\hat{r}_{02}$ , and  $A_{\text{He}}$  with the age, mass, initial mass fraction of hydrogen, and the initial heavy elements over hydrogen abundance ratio as free parameters. To do so, we select an initial guess value of  $X_0$  and derive the best fit values of  $\Delta$ ,  $\hat{r}_{01}$ , and  $\hat{r}_{02}$ . This results in a set of values for the mass, the age, and  $(Z/X)_0$ . Then, the adjustment is done by applying the secant method to find the value of  $X_0$  in best agreement with the target  $A_{\text{He}}$ . At each step of the secant algorithm – at each value of  $X_0$  –, Levenberg-Marquardt’s algorithm (L-M) computes the optimal set of stellar parameters giving the best fit of  $\Delta$ ,  $\hat{r}_{01}$ , and  $\hat{r}_{02}$ . Finally, when we have a good estimate for  $A_{\text{He}}$ , we use L-M’s algorithm one last time to derive the complete set of values for the mass, the age,  $X_0$ , and  $(Z/X)_0$  fitting the observed parameters. We used the secant method in order to diminish the computational time needed to converge towards the solution.



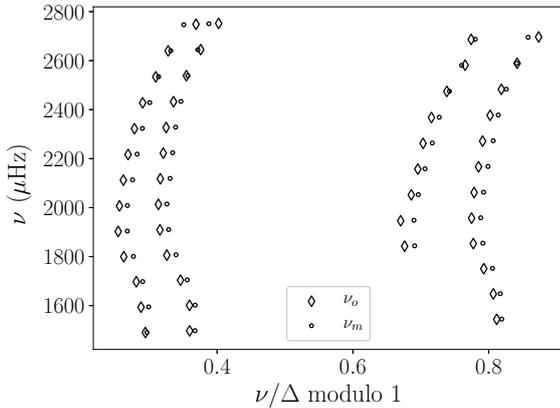
**Fig. 12.** Helium amplitude (solid line) and  $(Z/X)_0$  ratio (dotted line) versus the initial hydrogen abundance. The horizontal solid line represents the target value for the amplitude – corrected for the surface effects following Kjeldsen et al. (2008) – while the dashed lines represent the  $1\sigma$  interval. The horizontal dotted line is the value of the amplitude without surface effects correction.

### 5.2. Results

The values of the relevant seismic indicators as well as the associated standard deviations computed from the observed frequencies are given in Table 1. The column labeled Kjeldsen refers to the observed indicators corrected for the surface effects using Sonoi et al. (2015)’s coefficient fitted to Kjeldsen et al. (2008)’s prescription as explained in Sect. 3.1. The Sonoi column corresponds to the correction of surface effects using a Lorentzian profile as in Sonoi et al. (2015). We also derived the seismic indicators while adding smooth basis elements with  $n^{-1}$ . As the frequencies are projected in a specific order to build the seismic indicators, only the glitches amplitudes are affected and we have:  $A_{\text{He}} = 27.6$  and  $A_{\text{CZ}} = 3.3$ . Using the methodology described above, we managed to derive the stellar parameters given in Table 2 for the different sets of observed seismic indicators. We note that changing the treatment of surface effects or adding the  $n^{-1}$  basis elements have an impact on the fitted parameters comparable to that of the frequencies uncertainties. Thus, we will only discuss here the parameters adjusted to the data corrected for the surface effects using Kjeldsen et al. (2008)’s prescription and without adding  $n^{-1}$  basis elements. The parameters we derived do not constitute a detailed characterisation of the target 16 Cyg A. Rather, they illustrate the ability of the method to provide constraints on a solar analog. Indeed, only one single set of input physics was tested. Therefore, the standard deviations tend to be underestimated as they are the ones intrinsic to the method. The abundances used for the computations were the solar ones determined by Asplund et al. (2009). Let us add that we obtain a surface helium abundance of  $Y_{\text{f}} = 0.242 \pm 0.028$  which lies in the interval obtained by Verma et al. (2014),  $Y_{\text{f},V} \in [0.231, 0.251]$ . This is comforting us in the idea that the developed method is efficient in isolating the glitches and drawing inferences from their signatures. In Fig. 12, we show the evolution of the helium glitch amplitude resulting from the three parameters adjustment of  $\Delta$ ,  $\hat{r}_{01}$ , and  $\hat{r}_{02}$  as a function of  $X_0$ . We also show the evolution of the value of  $(Z/X)_0$ . The very linear trend justifies that we used the secant method to provide successive estimates of  $A_{\text{He}}$  in order to lessen the computational charge. In addition, we illustrate both the observed value for  $A_{\text{He}}$  and the corrected value under Kjeldsen et al. (2008)’s prescription. We note that the corrected value is of about  $0.79\sigma$  lower than the uncorrected one. Using Sonoi et al. (2015)’s prescription only leads to a  $0.55\sigma$



**Fig. 13.** Comparison between the observed helium glitch  $\delta\nu_o$  (solid line) and the one resulting from the best fit model  $\delta\nu_m$  (dot-dashed line), for  $l = 0$ . We also display the observed glitch as a function of the frequencies (errorbars) as well as the best model glitch associated with the theoretical frequencies (diamond).



**Fig. 14.** Comparison between the observed frequencies (diamonds) and the best model frequencies (circles) in an échelle diagram.

variation of the measured amplitude. This demonstrates that the surface effects have a small influence on the amplitude we derive using the method. Also, we retrieved a value of 970.97 s for  $\tau_{\text{He}}$  and 3042.32 s for  $\tau_{\text{CZ}}$ . As they were fixed to model values and, as a consequence, were neither parameters nor constraints of the adjustment, we do not provide uncertainties. However, we may compare their values to the ones adjusted by Verma et al. (2014). They obtained  $\tau_{\text{He}} \in [868, 944]$  s and  $\tau_{\text{CZ}} \in [2992, 3234]$  s. We observe that  $\tau_{\text{CZ}}$  lies in the interval calculated by Verma et al. (2014) while the value of  $\tau_{\text{He}}$  is slightly above the upper limit. However, as we show in Sect. 2.3 this does not impact the inferences drawn from the helium glitch amplitude in a significant way. To illustrate this statement, we freed the value of  $\tau_{\text{He}}$  (but not that of  $\tau_{\text{CZ}}$  as the convection zone glitch is of negligible amplitude compared to that of the helium glitch). The relative change between the optimised  $\tau_{\text{He}}$  and its estimator is of only 6%. Moreover, the observed value of the helium glitch amplitude remains unchanged compared to its standard deviation (we observe a change of  $0.06\sigma$ ). As expected, the best fit  $Y_f$  also remains untouched. At this point, one should be reminded that we focus on the glitch amplitude to draw our inferences. And, as we showed in Sect. 2.3, the exact location of the glitch does not have a significant impact on its amplitude as well as on the derived surface helium abundance. Finally, Fig. 13 shows the fit-

ted helium glitches<sup>5</sup> for the observed data and the 4 parameters best fit model. It is visible that the observed and fitted glitches are close in amplitude and period. This demonstrates the ability of the method to both isolate the glitch and the related parameters and to provide a model reproducing at best those parameters. We also show in Fig. 14 the comparison in an échelle diagram of the observed frequencies and those of the best fit model. We observe in this last figure that the smooth component of the spectrum (not visible in the previous figure) is properly adjusted. Thus, by fitting the set of indicators  $\Delta_0$ ,  $\hat{r}_{01}$ ,  $\hat{r}_{02}$ , and  $A_{\text{He}}$  we obtain a good representation of the observed frequencies. We also show in Appendix F the adjustment to frequencies corrected for surface effects using a Lorentzian profile as in Sonoi et al. (2015). We observe in Fig. F.1 that the agreement between the model and observed glitches is better than in Fig. 13. However, we also note in the corresponding échelle diagram (Fig. F.2) that there is an offset between observed and model ridges. This corresponds to different values of  $\hat{\epsilon}$ . Figure 4 also illustrates this behaviour as it is clearly visible that only Kjeldsen et al. (2008)’s prescription allow to reproduce the observed value for  $\hat{\epsilon}$ .

## 6. Discussion and conclusions

### 6.1. Principle

In the present paper, we provide a new method that uses as much of the available seismic information as possible. Indeed, we take advantage of the information contained in both the glitches and the smooth component (usually discarded in glitches analyses but then used separately for forward seismic modelling, see Verma et al. 2014 for example) of the spectrum to define indicators that are as independent as possible of each other. To do so, we take advantage of Gram-Schmidt’s algorithm to create an orthonormal basis over which we project the frequencies. The obtained coefficients are therefore independent of each other and consist in a linear combination of the frequencies. Thus, using the appropriate combination of those allows us to define uncorrelated indicators. Such indicators are constructed in order to reproduce the behaviour of “usual” indicators such as the large frequency separation. Up to this day and to our knowledge, no method, has been proposed to provide proper correlations between the smooth and glitch indicators as they are built separately. Therefore, our method provides the asteroseismologists with new diagnosis means.

### 6.2. Advantages

As the method only relies on linear algebra, it is very stable and the computation times are negligible – of the order of a fraction of a second –. Thus, it could easily be implemented in stellar model fitting algorithms, which are non-linear, without impacting the total computational time. This can be done with any algorithm as the method only focuses on the definition of new seismic indicators and does not rely on the physics of the model itself. Let us add that the defined indicators should be used in combination with non-seismic constraints through a single merit function while searching for a stellar model in order to obtain proper covariances between the inferred quantities – which is not often the case in seismic analyses. The illustration for 16 Cyg A

<sup>5</sup> We do not display the envelope convective zone glitch as its amplitude is negligible compared to that of the helium glitch, as mentioned in Sect. 4.

truly demonstrates the possibility to use the new indicators to provide further constraints on stellar structure in the framework of forward seismic modelling.

In addition, the usual indicators often hold a local definition and may use correlated information while the new ones are built in a way that the information used is averaged over the whole frequency range and that it is not redundant (each observed frequency is used only once for each indicator). This results in smaller standard deviations and smoother behaviours. Appendix B shows the evolution of the classical indicators  $r_{01}(n)$  and  $r_{02}(n)$  as defined by Roxburgh & Vorontsov (2003). We indeed observe that the new definitions of the indicators give smaller error bars while preserving the expected trends, providing tighter constraints. Finally, the method has the advantage that it can be implemented even in cases where some modes are missing. It is not the case of the “usual” methods as they need successive modes to define some of their indicators. For example, the classical local definitions of the large separation and second differences require at least two consecutive frequencies. Therefore, whenever some modes are missing, pieces of information might be discarded.

### 6.3. Information carried by the indicators

We have defined indicators that provide estimates for the classical indicators that are the large separation,  $\Delta$  and  $\Delta_l$ , and the small separation ratios,  $\hat{r}_{0l}$ . As expected,  $\Delta$  provides an estimation of the stellar mean density (Ulrich 1986). Also, we defined the indicators  $\Delta_{0l}$  which combine the large separations associated with the spherical degrees 0 and  $l$  which are known to give an estimate of the resonant cavity of the  $l$  degree modes (Monteiro 2002). Those indicators provide an estimator of the slope of  $r_{010}$  which, combined with  $\hat{r}_{01}$  can be used to constrain the overshooting parameter as shown in Fig. 3.

Moreover, we observe in Fig. 1 that the indicator  $\hat{r}_{02}$ , defined to estimate the small separation ratio between the spherical degrees 0 and 2, is a proper indicator of the evolution of the star, as expected from Christensen-Dalsgaard (1988). Also, we observe that the indicator  $\hat{r}_{01}$  presents a degeneracy, as a turn-off occurs in the  $\Delta_0$ - $\hat{r}_{01}$  plane, and there is an inaccessible region. This is a very interesting observation as it should provide tight constraints on the stellar structure. Indeed, we show that a change in composition or in the value of  $\alpha_{\text{MLT}}$  allows to modify that region. It will therefore be necessary to use such parameters to reproduce observed values.

Furthermore, Fig. 4 shows that we might get constraints on the surface effects from the indicator  $\hat{\epsilon}$  we defined. It should allow to discriminate from several empirical formulations meant to account for the surface effects. We have tested both formulations from Kjeldsen et al. (2008) and Sonoï et al. (2015) as an illustration of the diagnosis power of the indicator  $\hat{\epsilon}$ . However, we are aware of the existence of the formulation from Ball & Gizon (2014) and it should also be inspected in further studies. Besides, we showed that the helium amplitude indicator is almost unaffected by the surface effects as it has to be expected (see for example Fig. 12). Indeed, empirical surface effects corrections are important for the high frequencies compared to  $\nu_{\text{max}}$  while the glitch is of great amplitude only for the low frequencies. Also, the helium glitch amplitude should allow to draw inferences on the surface helium abundance as shown by Fig. 5. However, attention has to be paid as it is also anti-correlated with the metallicity (see Fig. 6). We have demonstrated via a toy model for the first adiabatic index that both effects on the amplitude stem from the position of the adiabat which in turn

determines the amplitude of the second local  $\Gamma_1$  minimum due to helium partial ionisation.

In addition, we observe that the convective zone glitch amplitude has a significantly lower amplitude than that of the helium glitch and is correlated with the amount of undershooting at the base of the envelope convection zone (in agreement with Verma et al. 2014).

Finally, let us add that it is possible to define other indicators than those presented in this paper. This should therefore be carefully studied to take advantage of as much of the available information as possible.

### 6.4. Limitations

However, we show that the method is only fit to draw inferences about solar-like stars, that is low-mass stars on the main-sequence. Indeed, Fig. 10 illustrates that from masses around  $1.3 M_{\odot}$  and above, diffusion plays an important role and the relation between helium surface abundance and the helium glitch amplitude is not monotonic anymore. Therefore, the method will have to be adapted for massive and evolved stars.

Moreover, one could argue that using model values as estimators of the acoustic depths of the glitches is a major drawback of the method. However, the proposed method does not focus on those quantities. Rather, we focus on the information that the amplitude of the glitch (not the period of the signal) carries. This means that we only need proper, but not exact, estimators for the acoustic depths in order to draw inferences. Indeed, we showed in Sect. 2.3, in the case of the helium glitch, that a small excursion from the estimated value (either by manually setting another value, that of the second minimum in the first adiabatic index for the helium glitch, or by optimising over its value) does not lead to a significant change in the measured amplitude. Therefore, the inferences drawn remained unchanged. However, one could still regard this as a flaw of the method as we do not provide a new way of retrieving the acoustic depths of glitches. This could be explored in future studies by, for example, finding the global optimum for the acoustic depths of the glitches and then using the measured value as a constraint for the best fit model. Nevertheless, this would make the calculations more time consuming and annihilate the benefit of the orthonormalisation that is the independence of the fitted parameters.

Also, the indicators should be used to complement non-seismic data as we have shown that, for example, there exists a degeneracy between the helium glitch amplitude, the helium surface abundance and the metallicity<sup>6</sup> (see Figs. 5 and 6). Therefore, the helium amplitude alone is not sufficient and additional information are needed to lift such a degeneracy – these information may be contained in other indicators (seismic and non-seismic).

### 6.5. Future perspective

The next step will be the detailed study with the new method of several stars from the *Kepler* legacy sample (Lund et al. 2017). This sample consists of 66 main sequence stars for which at least one year of continuous observations has been made. Having such long time series provides the necessary precision to study glitches. Let us add that this future study is intended to be independent of what has been done in Verma et al. (2017).

<sup>6</sup> Which we have shown in Sect. 4.2 to be an indirect effect as it is due to the influence on density and temperature profiles for the metallicity.

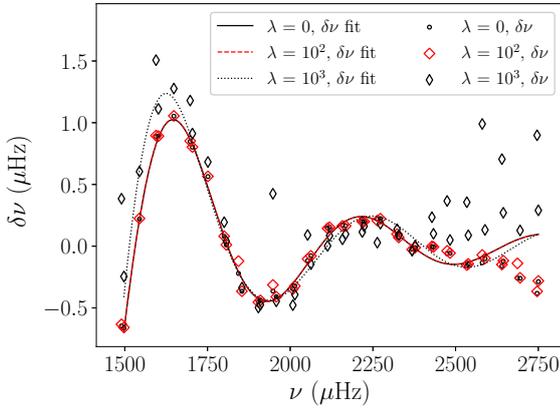
Also, another important step will be the improvement of the method by enabling the study of glitches present in red subgiants spectra. The peculiarity of such pulsators is that they exhibit mixed modes. This will consist in a new challenge as their analytical formulation is not a simple task. However, this is a necessary step to improve our knowledge of the evolution of a star such as the Sun. Moreover, this will allow a better description of such stars and will provide a deeper understanding of their properties via, for example, the characterisation of the mixing processes.

*Acknowledgements.* The authors would like to thank the referee for their very constructive remarks that improved the paper in a significant way. M.F. is supported by the FRIA (Fond pour la Recherche en Industrie et Agriculture) – FNRS PhD grant. S.J.A.J.S. is funded by ARC grant for Concerted Research Actions, financed by the Wallonia-Brussels Federation. GB acknowledges support from the ERC Consolidator Grant funding scheme (project ASTEROCHRONOMETRY, G.A. no. 772293).

## References

- Adelberger, E. G., García, A., Robertson, R. G. H., et al. 2011, *Rev. Mod. Phys.*, **83**, 195
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Baglin, A., Auvergne, M., Barge, P., et al. 2009, in *Transiting Planets*, eds. F. Pont, D. Sasselov, & M. J. Holman, *IAU Symp.*, **253**, 71
- Ball, W. H., & Gizon, L. 2014, *A&A*, **568**, A123
- Basu, S., Mazumdar, A., Antia, H. M., & Demarque, P. 2004, *MNRAS*, **350**, 277
- Bedding, T. R., Mosser, B., Huber, D., et al. 2011, *Nature*, **471**, 608
- Böhm-Vitense, E. 1958, *ZAp*, **46**, 108
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, in *American Astronomical Society Meeting Abstracts*, BAAS, **42**, 215
- Brent, R. P. 1973, *Algorithms for Minimization without Derivatives* (Englewood Cliffs, NJ: Prentice-Hall)
- Buldgen, G., Reese, D. R., & Dupret, M. A. 2016a, *A&A*, **585**, A109
- Buldgen, G., Salmon, S. J. A. J., Reese, D. R., & Dupret, M. A. 2016b, *A&A*, **596**, A73
- Cassisi, S., Salaris, M., & Irwin, A. W. 2003, *ApJ*, **588**, 862
- Christensen-Dalsgaard, J. 1988, in *Advances in Helio- and Asteroseismology*, eds. J. Christensen-Dalsgaard, & S. Frandsen, *IAU Symp.*, **123**, 295
- Cox, J. P., & Giuli, R. T. 1968, *Principles of Stellar Structure* (New York: Gordon and Breach)
- Davies, G. R., Chaplin, W. J., Farr, W. M., et al. 2015, *MNRAS*, **446**, 2959
- Deheuvels, S., Michel, E., Goupil, M. J., et al. 2010, *A&A*, **514**, A31
- Deheuvels, S., Brandão, I., Silva Aguirre, V., et al. 2016, *A&A*, **589**, A93
- de Meulenaer, P., Carrier, F., Miglio, A., et al. 2010, *A&A*, **523**, A54
- Ferguson, J. W., Alexander, D. R., Allard, F., et al. 2005, *ApJ*, **623**, 585
- Gough, D. O. 1986, *Highlights Astron.*, **7**, 283
- Gough, D. O. 1990, in *Progress of Seismology of the Sun and Stars*, eds. Y. Osaki, & H. Shibahashi (Berlin: Springer Verlag), *Lect. Notes Phys.*, **367**, 283
- Gram, J. P. 1883, *J. Reine Angew. Math.*, **94**, 41
- Grec, G., Fossat, E., & Pomerantz, M. A. 1983, *Sol. Phys.*, **82**, 55
- Houdek, G., & Gough, D. O. 2007, *MNRAS*, **375**, 861
- Iglesias, C. A., & Rogers, F. J. 1996, *ApJ*, **464**, 943
- Kjeldsen, H., Bedding, T. R., & Christensen-Dalsgaard, J. 2008, *ApJ*, **683**, L175
- Lund, M. N., Silva Aguirre, V., Davies, G. R., et al. 2017, *ApJ*, **835**, 172
- Manchon, L., Belkacem, K., Samadi, R., et al. 2018, *A&A*, **620**, A107
- Mazumdar, A., Monteiro, M. J. P. F. G., Ballot, J., et al. 2014, *ApJ*, **782**, 18
- Miglio, A., & Montalbán, J. 2005, *A&A*, **441**, 615
- Monteiro, M. J. P. F. G. 2002, in *Observed HR Diagrams and Stellar Evolution*, eds. T. Lejeune, & J. Fernandes, *ASP Conf. Ser.*, **274**, 77
- Monteiro, M. J. P. F. G., Christensen-Dalsgaard, J., & Thompson, M. J. 2000, *MNRAS*, **316**, 165
- Popielski, B. L., & Dziembowski, W. A. 2005, *Acta Astron.*, **55**, 177
- Ramírez, I., Meléndez, J., & Asplund, M. 2009, *A&A*, **508**, L17
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Exp. Astron.*, **38**, 249
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2014, in *Space Telescopes and Instrumentation 2014: Optical, Infrared, and Millimeter Wave*, Proc. SPIE, **9143**, 914320
- Roxburgh, I. W. 2016, *A&A*, **585**, A63
- Roxburgh, I. W., & Vorontsov, S. V. 2002a, in *Stellar Structure and Habitable Planet Finding*, eds. B. Battrick, F. Favata, I. W. Roxburgh, & D. Galadi, *ESA Spec. Pub.*, **485**, 337
- Roxburgh, I. W., & Vorontsov, S. V. 2002b, in *Stellar Structure and Habitable Planet Finding*, eds. B. Battrick, F. Favata, I. W. Roxburgh, & D. Galadi, *ESA Spec. Pub.*, **485**, 341
- Roxburgh, I. W., & Vorontsov, S. V. 2003, *A&A*, **411**, 215
- Saha, M. N. D. 1920, *Lond. Edinb. Phil. Mag. J. Sci.*, **40**, 472
- Schmidt, E. 1907, *Math. Ann.*, **63**, 433
- Scuflaire, R., Théado, S., Montalbán, J., et al. 2008a, *Ap&SS*, **316**, 83
- Scuflaire, R., Montalbán, J., Théado, S., et al. 2008b, *Ap&SS*, **316**, 149
- Silva Aguirre, V., Ballot, J., Serenelli, A. M., & Weiss, A. 2011, *A&A*, **529**, A63
- Sonoi, T., Samadi, R., Belkacem, K., et al. 2015, *A&A*, **583**, A112
- Thoul, A. A., Bahcall, J. N., & Loeb, A. 1994, *ApJ*, **421**, 828
- Ulrich, R. K. 1986, *ApJ*, **306**, L37
- Verma, K., Faria, J. P., Antia, H. M., et al. 2014, *ApJ*, **790**, 138
- Verma, K., Raodeo, K., Antia, H. M., et al. 2017, *ApJ*, **837**, 47
- Vorontsov, S. V. 1988, in *Advances in Helio- and Asteroseismology*, eds. J. Christensen-Dalsgaard, & S. Frandsen, *IAU Symp.*, **123**, 151

## Appendix A: Effect of the regularisation constant

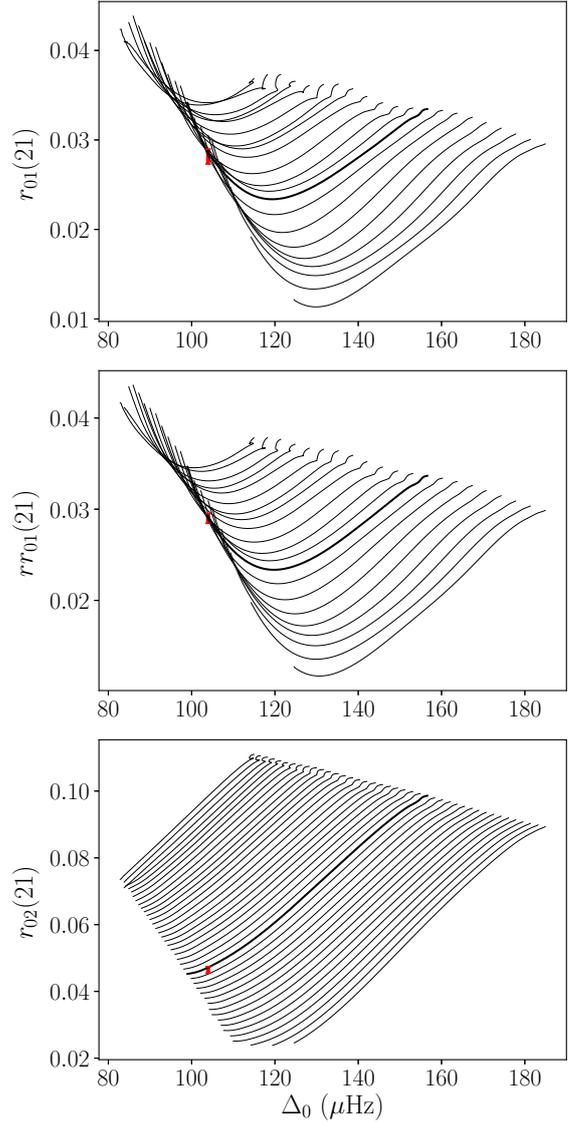


**Fig. A.1.** Comparison between the glitch adjustment for several values of the regularisation constant  $\lambda$  for 16 Cyg A best model determined in Sect. 5.2. The values are given in the legend. The lines correspond to the fitted glitch and the markers to the observed glitch. Only the fitted curves for  $l = 0$  are displayed.

Figure A.1 shows the influence of the regularisation constant  $\lambda$  on the quality of the fitted glitch. We fitted the glitch of the best model obtained in Sect. 5.2 using the values of  $\lambda$  shown in the legend. We observe that, when using a value of  $10^2$  which is already significant as discussed in Sect. 2.2, the fit (red dashed curve) remains close to the one obtained without including regularisation terms (solid black curve). There is only a slightly higher dispersion in the model frequencies subtracted from the smooth part of the spectrum. To show the degradation caused by a regularisation constant that dominates the adjustment, we used  $\lambda = 10^3$ . We observe both a discrepancy between results without and with (dotted black curve) regularisation terms and a higher dispersion. This translates in a higher value of the merit function (from  $\chi^2 \sim 1$  without regularisation to  $\chi^2 \sim 500$  with  $\lambda = 10^3$ ).

## Appendix B: Comparison with usual indicators

We present in Fig. B.1 the evolution of the “usual” indicators  $r_{01}(n)$  and  $r_{02}(n)$  (Roxburgh & Vorontsov 2003) used in asteroseismology in order to compare them with the new indicators. They are evaluated at the value of  $n = 21$  which corresponds to the measured value of  $n_{\max}$  for 16 Cyg A ( $n_{\max}$  being the value of  $n$  at  $l = 0$  of closest frequency to the  $\nu_{\max}$  value). First let us draw the attention to the fact that the usual indicators hold a local value, as Eqs. (26) and (27) show, while the new indicators are averaged over the complete set of available modes. It results from this averaging a lower standard deviation. Then, we notice that the behaviour of both the usual indicators and the new ones follow similar trends. This means that the definition is consistent with what has been done up to now and that it should hold the same information. Therefore, it should be able to provide similar diagnostics, but with higher precisions.



**Fig. B.1.** Evolution of the normalised small separation as defined by Roxburgh & Vorontsov (2003) and evaluated at  $n = 21$  along the grid presented in Sect. 3.1. The red marker shows the observed value corrected for the surface effects following Kjeldsen et al. (2008)’s prescription. The thick line represents the track for  $1.06 M_{\odot}$ . *Top panel:* three points normalised small separation between spherical degrees 0 and 1; *middle panel:* five points normalised small separation between spherical degrees 0 and 1; *bottom panel:* normalised small separation between spherical degrees 0 and 2.

## Appendix C: $\Gamma_1$ toy model

To build a toy model for  $\Gamma_1$  that replicates at best its behaviour in the helium second ionisation zone we use the following hypotheses:

- In the helium second ionisation zone, hydrogen is fully ionised;
- Metals are in their atomic form;

– We consider a perfect gas.

We then define:

– The once ionised helium number of particles per unit volume  $\text{He}^+$  and twice ionised helium number of particles per unit volume  $\text{He}^{++}$  such that:

$$\text{He}^+ + \text{He}^{++} = \frac{Y\rho}{4m_u}, \quad (\text{C.1})$$

where  $m_u$  is the atomic mass unit;

– The helium ionisation fraction:

$$x_{\text{He}} = \frac{\text{He}^{++}}{\text{He}^+ + \text{He}^{++}}, \quad (\text{C.2})$$

which equals 0 if the helium is ionised only once and 1 if it is fully ionised;

– The number of electrons per unit volume:

$$e = \left( X + (1 + x_{\text{He}}) \frac{Y}{4} \right) \frac{\rho}{m_u}; \quad (\text{C.3})$$

– The total number of particles:

$$n = \left( 2X + (2 + x_{\text{He}}) \frac{Y}{4} + \sum_i \frac{Z_i}{A_i} \right) \frac{\rho}{m_u}, \quad (\text{C.4})$$

where  $Z_i$  is the mass fraction of the metal labeled  $i$  and  $A_i$  its mass number.

Using Saha's equation (Saha 1920):

$$S_a = \frac{\text{He}^{++}e}{\text{He}^+} = \frac{g}{h^3} (2\pi m_e k_B T)^{\frac{3}{2}} e^{-\frac{\chi}{k_B T}}, \quad (\text{C.5})$$

where  $g$  is the statistical weight of helium at its fundamental state,  $h$  is Planck's constant,  $k_B$  is Boltzmann's constant,  $m_e$  the electron mass, and  $\chi$  the helium second ionisation energy, we may obtain a second order equation for  $x_{\text{He}}$ :

$$\frac{Y\rho}{4m_u} x_{\text{He}}^2 + \left[ \left( X + \frac{Y}{4} \right) \frac{\rho}{m_u} + S_a \right] x_{\text{He}} - S_a = 0, \quad (\text{C.6})$$

which we solve to obtain the evolution of  $x_{\text{He}}$  with temperature and density. We may also derive this expression with respect to temperature at constant density. This yields

$$\left. \frac{\partial x_{\text{He}}}{\partial T} \right|_{\rho} = \frac{(1 - x_{\text{He}}) \left( \frac{3}{2} + \frac{\chi}{k_B T} \right) \frac{m_u S_a}{\rho T}}{X + \left( \frac{1}{2} + x_{\text{He}} \right) \frac{Y}{4} + \frac{m_u}{\rho} S_a}. \quad (\text{C.7})$$

Then we have the following expressions:

$$P_T = \left. \frac{\partial \ln P}{\partial \ln T} \right|_s = 1 + \frac{eD}{n} \left( \frac{3}{2} + \frac{\chi}{k_B T} \right), \quad (\text{C.8})$$

$$P_{\rho} = \left. \frac{\partial \ln P}{\partial \ln \rho} \right|_s = \left[ (2 - D)X + (2 + x_{\text{He}} - (1 + x_{\text{He}})D) \frac{Y}{4} + \sum_i \frac{Z_i}{A_i} \right] \frac{\rho}{m_u n}, \quad (\text{C.9})$$

$$c_v = \frac{3}{2} \frac{k_B P_T}{\mu m_u} + \frac{Y\chi}{4m_u} \left. \frac{\partial x_{\text{He}}}{\partial T} \right|_{\rho}, \quad (\text{C.10})$$

where  $\mu$  is the mean molecular weight and

$$D = \left( \frac{e}{x_{\text{He}}(1 - x_{\text{He}})} \frac{4m_u}{Y\rho} + 1 \right)^{-1}. \quad (\text{C.11})$$

Finally, we may insert their values in the relation linking  $\Gamma_1$ ,

$P_T$ ,  $P_{\rho}$ , and  $c_v$ :

$$\Gamma_1 = P_{\rho} + P_T^2 \frac{P}{c_v \rho T}. \quad (\text{C.12})$$

## Appendix D: Gram-Schmidt's process and QR decomposition

As a reminder, Gram-Schmidt's algorithm consists in the construction of orthonormal basis elements from a set of non-orthonormal basis elements. Let us consider the element of index  $j_0 : \mathbf{p}_{j_0}$ . Let us also assume that we have already built the set of orthonormal basis functions up to index  $j_0 - 1$ , that is the set  $(\mathbf{q}_1, \dots, \mathbf{q}_{j_0-1})$ . To build element  $\mathbf{q}_{j_0}$ , we first subtract to  $\mathbf{p}_{j_0}$  its projection over the successive previous basis elements. We thus have

$$\mathbf{u}_{j_0} = \mathbf{p}_{j_0} - \sum_{j=1}^{j_0-1} \langle \mathbf{p}_{j_0} | \mathbf{q}_j \rangle \mathbf{q}_j, \quad (\text{D.1})$$

where  $\mathbf{u}_{j_0}$  is the basis element orthogonal to the set  $(\mathbf{q}_1, \dots, \mathbf{q}_{j_0-1})$ . Finally, it is normalised to obtain

$$\mathbf{q}_{j_0} = \frac{\mathbf{u}_{j_0}}{\|\mathbf{u}_{j_0}\|}. \quad (\text{D.2})$$

It is also possible to express this process as a QR decomposition. To do so we call  $\mathcal{P}_l$  the matrix of initial polynomials and  $\mathcal{Q}_l$  the matrix of orthonormal polynomials for a given spherical degree:

$$\mathcal{P}_l = \begin{pmatrix} p_0(n_{\min}) & \cdots & p_2(n_{\min}) \\ \vdots & & \vdots \\ p_0(n_{\max}) & \cdots & p_2(n_{\max}) \end{pmatrix} = (\mathbf{p}_0 \cdots \mathbf{p}_2)_l, \quad (\text{D.3})$$

$$\mathcal{Q}_l = \begin{pmatrix} q_0(n_{\min}) & \cdots & q_2(n_{\min}) \\ \vdots & & \vdots \\ q_0(n_{\max}) & \cdots & q_2(n_{\max}) \end{pmatrix} = (\mathbf{q}_0 \cdots \mathbf{q}_2)_l, \quad (\text{D.4})$$

where  $n_{\min}$  is the lowest observed radial order and  $n_{\max}$  the highest one for the spherical degree considered.

We may then express Gram-Schmidt's procedure in a matrix form as a QR decomposition as follows:

$$\mathcal{Q}_l = \mathcal{P}_l R_l^{-1}, \quad (\text{D.5})$$

where, using our definition of the scalar product

$$R_l = \begin{pmatrix} \langle \mathbf{q}_0 | \mathbf{p}_0 \rangle_l & \cdots & \langle \mathbf{q}_0 | \mathbf{p}_2 \rangle_l \\ & \ddots & \vdots \\ 0 & & \langle \mathbf{q}_2 | \mathbf{p}_2 \rangle_l \end{pmatrix}, \quad (\text{D.6})$$

is an upper triangular matrix.

We may now generalise by considering several spherical degrees. For the smooth part, the matrices become block matrices which are the combination of the matrices for the several spherical degrees considered. This block disposition illustrates the fact that we have different basis functions for the different spherical degrees – this is not the case for the glitch part –. This is represented by a *Kronecker delta* in Eq. (6). In these matrices, each row corresponds to a given mode and each column corresponds to a given element of the basis. Finally, we append the glitch part to those block matrices. They take the following form (we remind that the independent variable for the glitch basis is  $\bar{n} = n + l/2$ ):

$$\mathbb{P} = \left( \begin{array}{cc|ccc} \mathcal{P}_0 & 0 & \mathbf{p}_{\text{He},C,5}(n) & \cdots & \mathbf{p}_{C,S}(n) \\ & \ddots & \vdots & & \vdots \\ 0 & & \mathbf{p}_{\text{He},C,5}(n+3/2) & \cdots & \mathbf{p}_{C,S}(n+3/2) \end{array} \right), \quad (\text{D.7})$$

$$\mathbf{Q} = \left( \begin{array}{cc|ccc} \mathbf{Q}_0 & 0 & \mathbf{q}_{\text{He},C,5}(n,l) & \cdots & \mathbf{q}_{C,S}(n,l) \\ & \ddots & \vdots & & \vdots \\ 0 & \mathbf{Q}_3 & \mathbf{q}_{\text{He},C,5}(n,l) & \cdots & \mathbf{q}_{C,S}(n,l) \end{array} \right), \quad (\text{D.8})$$

And the  $\mathbf{R}$  matrix becomes

$$\mathbf{R} = \left( \begin{array}{cc|ccc} R_0 & 0 & \langle \mathbf{q}_k | \mathbf{p}_{\text{He},C,5} \rangle_{l=0} & \cdots & \langle \mathbf{q}_k | \mathbf{p}_{C,S} \rangle_{l=0} \\ & \ddots & \vdots & & \vdots \\ 0 & R_3 & \langle \mathbf{q}_k | \mathbf{p}_{\text{He},C,5} \rangle_{l=3} & \cdots & \langle \mathbf{q}_k | \mathbf{p}_{C,S} \rangle_{l=3} \\ & 0 & \langle \mathbf{q}_{\text{He},C,5} | \mathbf{p}_{\text{He},C,5} \rangle & \cdots & \langle \mathbf{q}_{\text{He},C,5} | \mathbf{p}_{C,S} \rangle \\ & & & \ddots & \vdots \\ & & 0 & & \langle \mathbf{q}_{C,S} | \mathbf{p}_{C,S} \rangle \end{array} \right), \quad (\text{D.9})$$

where  $\langle \mathbf{q}_k | \mathbf{p}_g \rangle_{l=0}$  is a column vector of which the rows are the successive scalar products of the basis elements with the glitch function – denoted with the  $g$  index – such that  $\langle \mathbf{q}_k | \mathbf{p}_g \rangle_{l=0} = (\langle \mathbf{q}_0 | \mathbf{p}_g \rangle_{l=0} \cdots \langle \mathbf{q}_2 | \mathbf{p}_g \rangle_{l=0})^T$ . We may again write the QR decomposition as

$$\mathbf{Q} = \mathbf{P}\mathbf{R}^{-1}. \quad (\text{D.10})$$

## Appendix E: A numerical example

We generated a set of frequencies for spherical degrees from 0 to 2 (listed in Table E.1) and applied our method. We show in Figs. E.1 and E.2 the successive adjustments of the basis functions. This shows the validity of using the set of functions described in Sect. 2 and allows oneself to compare their results with ours. As a reminder, we project the frequencies over the basis in the specific order explained in Sect. 2. Therefore, taken in the correct order, these plots provide intermediary results for the glitch adjustment. We also show in Table E.2 the values of the fitted coefficients. In this example, we also fitted  $n^{-1}$  polynomials – inspired by the second order form of the asymptotic expansion – to show that it is not necessary, nor relevant, to add supplementary basis elements to our method. Indeed, the fitted coefficient values become comparable to the standard deviation that is equal to 1, through the orthonormalisation. Such values and plots were obtained as follows:

1. Considering the set of standard deviations from Table E.1, we use Gram-Schmidt procedure (Eq. (4)) associated with the definition of the scalar product (Eq. (1)) to produce the ordered orthonormal basis functions for each value of the spherical degree  $l$ . Thus, we successively project the former basis elements  $p_j(n, l)$  (i.e. the ordered set of polynomials  $n^0, n^1, n^2$  and the glitch functions for each spherical degree) on the already defined orthonormal basis elements  $q_{j_0}(n, l)$ . Then, we normalise those projections. This provides us with the orthonormal basis elements  $q_{j_0}(n, l)$ , as well as the transformation matrix  $R_{j,j_0}^{-1}$ .
2. For the smooth part and one spherical degree at a time, the frequencies from Table E.1 are projected on the orthonor-

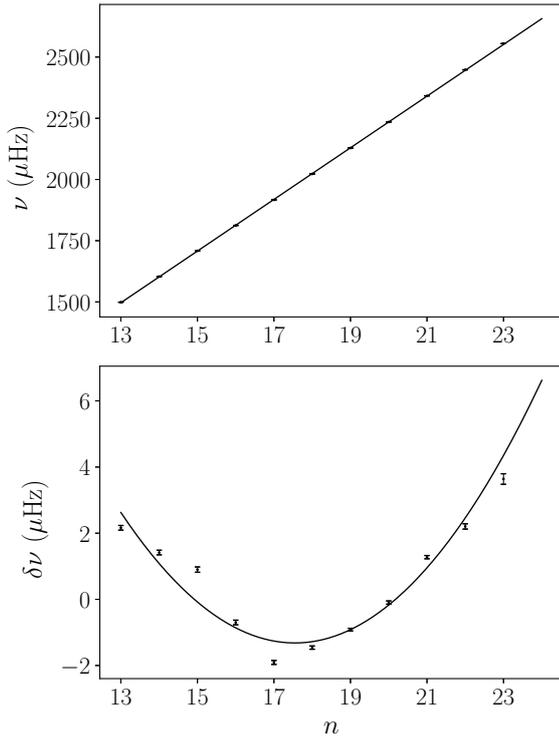
**Table E.1.** Example set of frequencies.

$l$	$n$	$\nu(\mu\text{Hz})$	$\sigma(\mu\text{Hz})$
0	13	1498.89	0.07
0	14	1603.60	0.07
0	15	1708.55	0.08
0	16	1812.40	0.07
0	17	1916.65	0.06
0	18	2022.56	0.05
0	19	2128.56	0.04
0	20	2234.84	0.05
0	21	2341.67	0.05
0	22	2448.06	0.08
0	23	2554.95	0.16
1	13	1546.42	0.07
1	14	1651.36	0.09
1	15	1755.56	0.08
1	16	1860.40	0.05
1	17	1965.44	0.05
1	18	2071.47	0.05
1	19	2178.50	0.04
1	20	2284.98	0.05
1	21	2391.77	0.06
1	22	2499.11	0.08
1	23	2606.15	0.13
2	13	1596.42	0.19
2	14	1701.68	0.17
2	15	1805.69	0.11
2	16	1910.30	0.10
2	17	2016.47	0.08
2	18	2122.70	0.06
2	19	2229.41	0.06
2	20	2336.56	0.09
2	21	2443.24	0.13
2	22	2550.54	0.21

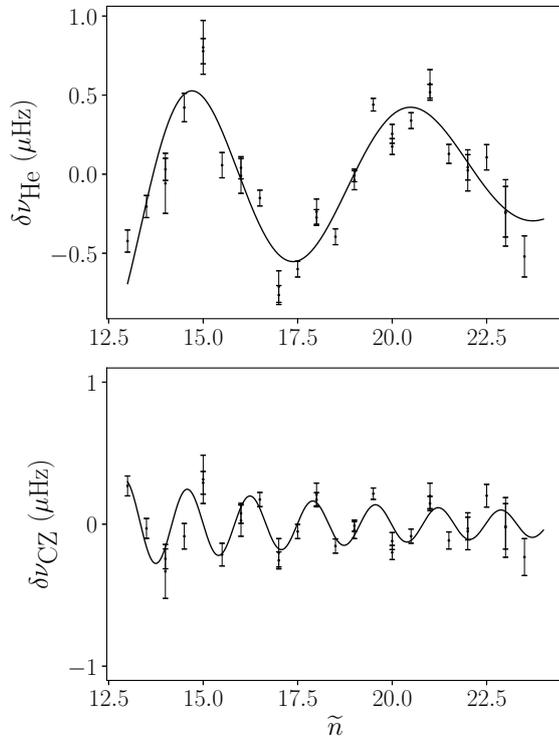
**Table E.2.** Fitted parameters to the frequencies of Table E.1.

	$l = 0$	$l = 1$	$l = 2$
$a_{l0}$	112948.0	117296.1	72112.2
$a_{l1}$	14952.0	14674.1	6882.0
$a_{l2}$	74.0	79.4	35.4
$a_{l-1}$	-3.5	9.0	6.3
$A_{\text{He}}$		27.5	
$A_{\text{CZ}}$		9.4	

mal basis elements following the proper order to produce the fitted frequencies,  $\nu_f(n, l)$ , according to Eq. (5) where the fitted coefficients are given by Eq. (8). Then, for the glitch part, the frequencies are projected simultaneously for every spherical degree on the glitch basis elements following the same procedure. This is due to the fact that the glitch coefficients should not depend on  $l$ . This produces the coefficients from Table E.2.

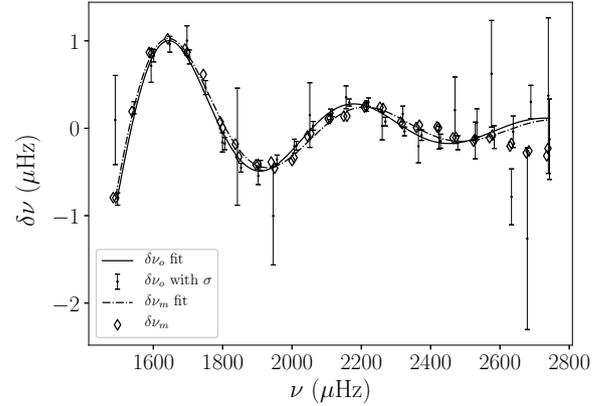


**Fig. E.1.** Comparison between the successive adjustments and the observed radial modes frequencies listed in Table E.1. *Upper panel:* observed frequencies compared to the first order adjustment; *lower panel:* residual of the first order adjustment compared to the second order fit to those residuals.

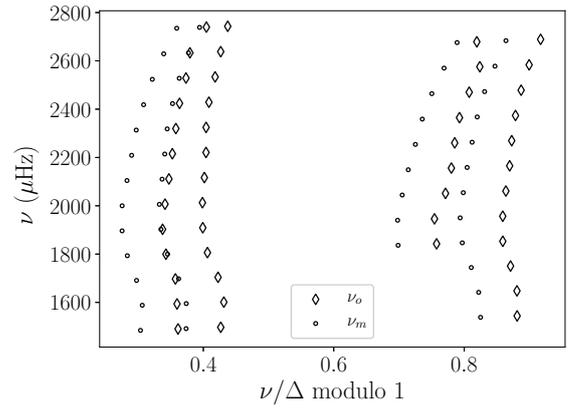


**Fig. E.2.** Separated glitches adjustments to the frequencies in Table E.1 of both glitches shown for the spherical degree  $l = 0$ . *Upper panel:* helium glitch; *lower panel:* convection zone glitch.

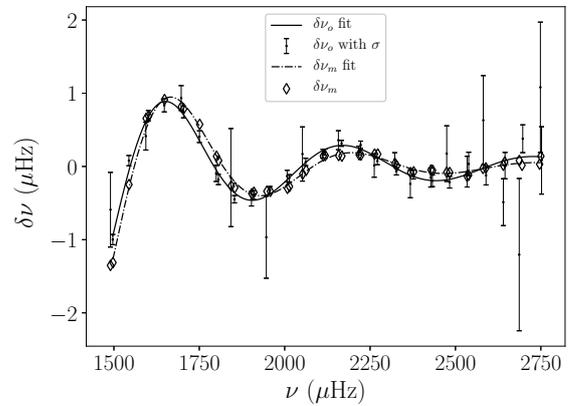
## Appendix F: Supplementary adjustments



**Fig. F.1.** Comparison between the observed helium glitch  $\delta\nu_o$  (solid line) and the one resulting from the best fit model  $\delta\nu_m$  (dot-dashed line) for  $l = 0$ . We also display the observed glitch as a function of the frequencies (errorbars) as well as the best model glitch associated with the theoretical frequencies (diamond). The observed frequencies have been corrected for surface effects using [Sonoii et al. \(2015\)](#)'s prescription.



**Fig. F.2.** Comparison between the observed frequencies (diamonds) and the best model frequencies (circles) in an échelle diagram. The observed frequencies have been corrected for surface effects using [Sonoii et al. \(2015\)](#)'s prescription.



**Fig. F.3.** Comparison between the observed helium glitch  $\delta\nu_o$  (solid line) and the one resulting from the best fit model  $\delta\nu_m$  (dot-dashed line) for  $l = 0$ . We also display the observed glitch as a function of the frequencies (errorbars) as well as the best model glitch associated with the theoretical frequencies (diamond). We include polynomials in  $n^{-1}$  to the basis functions.

We show in the present section the best fit model frequencies adjusted to the observed 16 Cygni A frequencies for several cases. First we show the adjustment to the frequencies corrected for surface effects using a Lorentzian profile as in [Sonoï et al. \(2015\)](#). Figure [F.1](#) shows that both glitches are in good agreement. However, the frequencies are systematically shifted, as Fig. [F.2](#) illustrates, as a consequence of the difference between the observed and theoretical  $\hat{\epsilon}$  values. Figure [4](#) illustrates such a discrepancy. This shows that, even though [Sonoï et al. \(2015\)](#) showed that for high frequency regimes [Kjeldsen et al. \(2008\)](#)'s prescription is not able to reproduce frequency differences between patched and unpatched model, it is

the only tested empirical correction that allowed us to reproduce the observed value for  $\hat{\epsilon}$ . In a further study, it would be appropriate to try out a scaled formulation of [Ball & Gizon \(2014\)](#)'s correction such as presented in [Manchon et al. \(2018\)](#).

We also provide in Fig. [F.3](#) the best fit model including terms in  $n^{-1}$  in the basis functions for the smooth part of the spectrum. By eye, the glitch adjustment seems better than in Fig. [13](#). However, as shown in Table [E.2](#), the improvement is not significant as the fitted parameters values are comparable to their standard deviation. Moreover, Table [2](#) demonstrates that the effect of including such terms in the adjustment is comparable to a variation of  $1\sigma$  in the frequencies.