MNRAS **000**, 1–28 (2025) Preprint 1 July 2025 Compiled using MNRAS LATEX style file v3.0

Breaking free seismic forward modelling from the surface effect⋆

Thomas Kallinger,[1]†
[1]*Institut für Astrophysik, Universität Wien, Türkenschanzstrasse 17, 1180 Vienna, Austria*

Accepted XXX. Received YYY; in original form ZZ

ABSTRACT

Asteroseismology has revolutionized our understanding of stellar interiors, but systematic uncertainties in stellar models – particularly the near-surface effect – hinder precise comparisons between observations and theory. This study introduces `BaseFM`, a novel Bayesian framework for asteroseismic forward modeling that eliminates the need for ad hoc surface corrections. By marginalizing over frequency offsets and incorporating glitch signals from sharp structural variations, it delivers unbiased stellar parameters with remarkable accuracy. Validated against eclipsing binaries and interferometric measurements, the method achieves stellar parameters accurate to within about 1% for stars across evolutionary stages, from Sun-like stars to red giants. Applying `BaseFM` to 70 stars in open clusters NGC 6791 and NGC 6819, consistent masses and ages are derived, confirming revised seismic scaling relations. The analysis reveals that classical surface corrections systematically overestimate stellar masses by about 5%, and therefore underestimate the stellar age about 20%. The true surface effect, on the other hand, exhibits a mass and/or metallicity-dependent trend incompatible with standard power-law assumptions. Additionally, I present a new correction for mixed dipole modes in evolved stars, enhancing the utility of non-radial oscillations.

This work underscores the limitations of empirical surface corrections and provides a robust, physics-independent tool for stellar characterization. By bridging the gap between observations and models, `BaseFM` opens new avenues for probing physics in stars with outer convective layers that exhibit solar-type oscillations. The method's success paves the way for future missions like PLATO, promising unprecedented precision in asteroseismic studies.

Key words: methods: statistical – stars: oscillations – stars: fundamental parameters

1 INTRODUCTION

Asteroseismology has been a cornerstone of the space photometry revolution initiated by MOST (Walker et al. 2003), CoRoT (Baglin et al. 2006), Kepler (Borucki et al. 2010), BRITE-Constellation (Weiss et al. 2014), and TESS (Ricker et al. 2015). These missions enabled the detection of solar-like oscillations in hundreds of Sun-like stars and thousands of evolved cool giants, leading to breakthroughs such as the discovery of rapidly rotating cores in subgiants (Deheuvels et al. 2012) and red giants (Beck et al. 2012), as well as the systematic determination of fundamental stellar parameters (e.g. Kallinger et al. 2010c). The unprecedented quality and duration of photometric time series data offer the potential for precise determinations of both global stellar parameters and internal structures by comparing observed oscillation frequencies to stellar model frequencies (see Chaplin & Miglio 2013, for a review).

However, this approach is hampered by uncertainties in stellar models. A significant challenge arises from the treatment of convection, where mixing-length theory (Böhm-Vitense 1958) remains a standard approximation. While effective in deep stellar interiors, this theory becomes insufficient in the near-surface layers, resulting in systematic shifts in computed model frequencies. These discrepancies, collectively referred to as the near-surface effect, complicate direct comparisons between observations and models.

The seismic surface effect was first recognized in the Sun (e.g., Brown 1984; Christensen-Dalsgaard et al. 1988) and has since been extensively studied (e.g., Zhugzhda & Stix 1994; Schlattl et al. 1997; Stein & Nordlund 1991; Rosenthal et al. 1999; Yang & Li 2007; Houdek et al. 2017; Li et al. 2023). In the absence of a definitive solution, Kjeldsen et al. (2008) proposed an empirical surface-effect correction modeled as a frequency-dependent power law, calibrated to the differences between observed and theoretical solar radial frequencies (Christensen-Dalsgaard et al. 1996). While this method avoids matching incorrect models to observed frequencies, it overestimates the effect at low frequencies and

⋆ `BaseFM` is available at `https://github.com/tkallinger/BaseFM`
† E-mail: thomas.kallinger@univie.ac.at

assumes that the frequency shifts follow the same power law for all stars, without physical justification.

To address these limitations, Ball & Gizon (2014) proposed a surface correction that combines inverse and cubic frequency terms normalized by mode inertia, providing a theoretically motivated improvement. Although successful for the Sun, this method may not be universally applicable. For example, Gruberbauer et al. (2013) demonstrated significant variations in the surface effect across the HR diagram, and Sonoi et al. (2015) argued for constraining surface corrections through 3D hydrodynamical simulations. By patching standard 1D models with their 3D simulations, they derived scaling relations for various surface correction functions, enabling estimates based on a star's effective temperature and surface gravity without relying on observed frequencies.

Despite extensive application of standard surface corrections (e.g., Christensen-Dalsgaard et al. 2010; Tang & Gai 2011; Mathur et al. 2012; Deheuvels et al. 2012; Huber et al. 2019), the applicability to stars different from the Sun remains uncertain. Verifying these corrections requires precise independent measurements of stellar properties, which are typically unavailable for the rather faint targets observed by space missions. Rare exceptions include oscillating stars in eclipsing binary systems, where dynamic masses and radii can be precisely determined from radial velocity and photometric eclipse observations (e.g., Gaulme et al. 2016). Ball et al. (2018) investigated three such systems with red giant components and found that standard surface corrections failed to model the stars in agreement with binary-derived results, leading to systematic overestimation of masses and radii when using only seismic frequencies.

The present study introduces a new method for unbiased comparison between observed and model frequencies without explicitly correcting for the surface effect (or other systematic model deficiencies). Within a Bayesian framework, the method marginalizes over an additional frequency offset parameter, avoiding assumptions about its value. This allows for an unbiased and accurate determination of the mean stellar density but has only low discriminative power to disentangle mass and radius. This is done by correcting the observed and model frequencies for large-scale variations (like the curvature signal in red giant oscillation spectra) and comparing the residuals, which are then dominated by modulations due to sharp structural variations in the stellar interior.

The method is validated using the eclipsing binary system KIC 8410637, reproducing independently measured stellar masses and radii. It is further tested on red giants KIC 4054905, KIC 9970396, and HR 6817, as well as the Sun-like stars α Cen A and 16 Cyg A and B, achieving mass and radius estimates accurate to within $\approx 1\%$. Finally, the method is applied to 70 red giants in the open clusters NGC 6791 and NGC 6819, providing independent estimates of cluster RGB masses and ages, along with a detailed analysis of the surface effect in evolved stars. I further develop an simple approach to correct mixed dipole modes in red giants and use the cluster results to improve the non-linear seismic scaling relations from Kallinger et al. (2018).

2 ASTEROSEISMIC FORWARD MODELLING

One of the key objectives in asteroseismic model fitting is to identify a stellar model that accurately represents a given star. The most effective way to test existing methods and develop new ones is to focus on stars with precise seismic data and independently measured fundamental parameters, such as mass and radius. Despite the wealth of data from missions like Kepler and TESS, such targets remain exceptionally rare. The most promising candidates are eclipsing binary systems, particularly those containing at least one star with a well-defined frequency spectrum.

2.1 The test case KIC 8410637

To illustrate asteroseismic forward modeling, I focus on KIC 8410637, a detached eclipsing binary system with a red giant component (Hekker et al. 2010). Frandsen et al. (2013) conducted an initial dynamical analysis, revealing that the system comprises a red giant branch (RGB) star and a main-sequence (MS) star with masses of $1.56 \pm 0.03 M_\odot$ and $1.32 \pm 0.02 M_\odot$, respectively, in a 408.3-day orbit.

Subsequently, Themeßl et al. (2018) refined these parameters using new radial velocity data and improved spectral disentangling, determining the RGB component's mass and radius to be $1.47 \pm 0.02 M_\odot$ and $10.60 \pm 0.05 R_\odot$, respectively. Additionally, they found the system to have a near-solar chemical composition.

The four-year-long Kepler data provide a clear red-giant oscillation spectrum, making this star an excellent target for testing a range of asteroseismic applications, from seismic scaling relations (Gaulme et al. 2016; Kallinger et al. 2018) to detailed modeling (Ball et al. 2018; Buldgen et al. 2019; Li et al. 2018). The peak bagging of the Kepler data is described in Appendix A, with the resulting mode frequencies listed in Table A1.

2.2 Stellar models

Stellar evolutionary models are computed with Modules for Experiments in Stellar Astrophysics[1] (`MESA`, r12778; Paxton et al. 2019, and references therein) and cover masses from 0.7 to 2.0 $M_\odot$ in steps of 0.05 $M_\odot$ from the Zero-Age-Main-Sequence to after the core-He burning phase. To improve performance of the subsequent analysis the full grid is split into a main-sequence, sub-giant, RGB, and post-RGB part, which are available on Github[2]. For the solar abundances, the overall abundance mixture of Grevesse & Sauval (1998) is implemented and the models are computed for an initial chemical composition of (Y, Z) = (0.280, 0.023)[3]. The mixing length parameter $\alpha_{\rm MLT}$ is set to 2.2, which is in agreement with the findings of Li et al. (2018). During evolution, overshooting is turned off and RGB mass loss is implemented according to Reimers (1975), using a scaling factor $\eta = 0.2$.

Since the default timestep settings of `MESA` are not meant to produce sufficiently dense grids, especially on the

[1] https://docs.mesastar.org/en/latest/
[2] https://github.com/tkallinger/BaseFM/tree/main/grids.
[3] Y is set according to $Y \simeq 0.249 + 1.33Z$ (e.g. Li et al. 2018)

Table 1. Basic parameters of the used stellar model grids. Mass ranges are in solar units with a step size of 0.05. The last column gives the degrees of modes computed for the different grid fractions: main-sequence and sub-giant (MS-SG), red giant branch (RGB), and post-RGB (pRGB).

	(Y,Z)	M	$\alpha_{\rm MLT}$	# of models	l
G1	(0.26,0.010)	[0.7;2]	2.2	MS-SG: 55762	0,1,2
				RGB: 45266	0,1
				pRGB: 28231	0
G2	(0.28,0.023)	[0.7;2]	2.2	MS-SG: 48922	0,1,2
				RGB: 44581	0,1
				pRGB: 25308	0
G3	(0.31,0.035)	[0.8;2]	2.1	MS-SG: 46360	0,1,2
				RGB: 40846	0,1
				pRGB: 21648	0

lower giant branch, the timestep control[4] is adjusted so that the model's mean density changes on average by only about 0.2%, causing typical changes in the eigenfrequencies of about 30nHz. Adiabatic mode frequencies for the about 118 800 models are then computed with the `GYRE`[5] (v5.2) stellar oscillation code (Townsend & Teitler 2013) in the range $[0.2, 1]$ times the acoustic cutoff frequency $\nu_{\rm ac}$, using an outer mechanical boundary condition as described by Christensen-Dalsgaard (2008). Grid parameters are summarised in Table 1.

In addition to the eigenfrequencies, several seismic parameters are computed for each model and stored in the grid files. Among these is the asymptotic large frequency separation, $\Delta\nu_{\rm as}$, which can be derived from the asymptotic expression for p-mode oscillations (Tassoul 1980) as:

$$\Delta\nu_{\rm as} = \left(2\int_0^R \frac{dr}{c_s}\right)^{-1}, \tag{1}$$

where c_s is the sound speed throughout the star, and R is its radius. However, $\Delta\nu_{\rm as}$ is valid only for very high frequencies, which are typically unobservable, and it generally overestimates the large frequency separation of the observed modes.

To address this limitation, Ong & Basu (2019a,b) proposed a more practical parameter: a modified high-order estimator for the theoretical asymptotic separation that accounts for the evolution of the acoustic turning points within the mode cavities. This estimator introduces an explicit frequency dependence and is defined as:

$$\Delta\nu_{\rm as}(\nu) = \left(2\int_0^R \frac{dr}{c_s}\frac{1}{\sqrt{1-\nu_{\rm ac}^2/\nu^2}}\right)^{-1}, \tag{2}$$

where the frequency dependence is incorporated through the term involving the acoustic cutoff frequency, $\nu_{\rm ac}$. In practice, this parameter is evaluated at frequencies near $\nu_{\rm max}$. To compute $\Delta\nu_{\rm as}(\nu_{\rm max})$, the standard output from `MESA` must be extended to include the (default) Eddington-grey atmosphere, along with the central and surface points, while retaining double points for accurate integration. This adjustment ensures precise computation of the modified asymptotic separation in stellar models.

The pure g-mode period spacing, $\Delta\Pi_1$, of dipole modes in red giants provides valuable insights into the properties of the stellar core. While it is often not directly observable, it can be reconstructed from the observed mixed p-g modes (e.g., Mosser et al. 2014), allowing for a meaningful comparison with theoretical models.

The theoretical period spacing is determined directly from the model structure and is given by (e.g., Dziembowski 1977):

$$\Delta\Pi_1 = \sqrt{2}\pi^2\left(\int_{r_1}^{r_2}\mathcal{N}\frac{dr}{r}\right)^{-1}, \tag{3}$$

where $\mathcal{N}$ is the Brunt-Väisälä frequency, integrated over the g-mode cavity defined by the radii r_1 and r_2. This formulation provides a direct link between the model's internal structure and the observed seismic properties, enabling detailed studies of stellar interiors.

It is important to highlight that the specific model parameters (e.g., Y, Z, $\alpha_{\rm MLT}$, η) and the applied model physics (e.g., overshooting, diffusion, rotational mixing) are not critical in the current context. The constructed model grids are intended solely to demonstrate the capabilities of the methodology, rather than to determine the most accurate parameters for the examples presented.

2.3 Classical χ^2 approach

Asteroseismic model fitting involves matching a set of N observed frequencies ($\nu_{\rm obs}$) to those calculated from stellar models ($\nu_{\rm mod}$). Assuming that both observed and model frequencies have normally distributed uncertainties, the total variance can be expressed as $\sigma^2 = \sigma_{\rm obs}^2 + \sigma_{\rm mod}^2$. Using this, the χ^2-statistic is defined as:

$$\chi^2 = \frac{1}{N}\sum_{\rm i}^{N}\frac{(\nu_{\rm i,obs}-\nu_{\rm i,mod})^2}{\sigma_{\rm i}^2}, \tag{4}$$

which must be minimized to identify the best-fitting model (e.g., Guenther & Brown 2004). By searching across a large grid of stellar models, it is possible to derive insights into the stellar interior and obtain precise fundamental parameters.

However, this straightforward approach is susceptible to several systematic uncertainties, such as those arising from surface effects, stellar activity, magnetic fields, rotational effects, mode misidentification, and finite grid resolution. These uncertainties can prevent the most appropriate model from being the one that minimizes Eq. 4. Consequently, the formalism may yield incorrect or nonsensical results for both the best-fit model and the derived uncertainties. Moreover, such issues can propagate into the interpretation of the underlying physics used to construct the models, leading to flawed conclusions.

For stars with convective envelopes, it is well-established that model frequencies systematically deviate from observations due to shortcomings in modeling the outer stellar layers. This discrepancy, known as the (near) surface effect, can be mitigated either by analyzing ratios of frequency differences (Roxburgh 2005) or by applying empirical corrections to the model frequencies.

[4] In practice the `max_years_for_timestep` parameter in the `MESA` inlist is manually adjusted to a value of $10^{7-1.13M}$, where M is the initial mass of the evolutionary track.
[5] https://gyre.readthedocs.io/en/stable/index.html

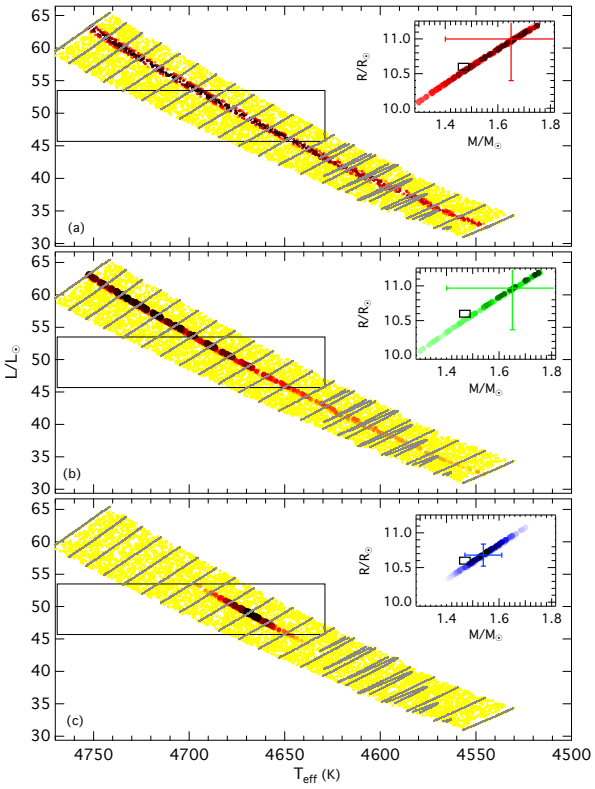

Figure 1. HR diagram showing the uncertainty box of KIC 8410637 and a subset of the stellar model grid used to fit the observed frequencies. Grey dots indicate the original grid and colored dots are interpolated values with the color scale indicating the "goodness of the fit" (the darker the better). Panel (a) gives the result for a classical χ^2 fit using the K08 surface correction. Panel (b) and (c) show posterior model probabilities from the surface-correction independent Bayesian model-fitting based on Eq. 11 and 16, respectively. Inserts show the χ^2 values (truncated above a value of four) and probabilities in the $M-R$ plane and the corresponding best-fit values and uncertainties. The black error box shows the dynamic solution from Themeßl et al. (2018).

One widely used empirical correction was proposed by Kjeldsen et al. (2008) (hereafter K08), who described the difference Δ between observed and model solar frequencies using a power law:

$$\Delta = \nu_{\rm obs} - \nu_{\rm mod} = a \left(\frac{\nu_{\rm mod}}{\nu_{\rm ref}} \right)^b , \tag{5}$$

where $\nu_{\rm ref}$ is a reference frequency (e.g., $\nu_{\rm max}$), and a and b are parameters to be determined. For the Sun, they found $b \approx 4.9$, a value that has been adopted for other stars by many authors. However, it is unlikely that this solar-calibrated correction is universally applicable to all stars (e.g., Gruberbauer et al. 2013; Sonoi et al. 2015).

Several modifications to this approach have been proposed. For instance, Ball & Gizon (2014, 2017) (hereafter B14) introduced a cubic correction ($b \hat{=} 3$) normalized by the mode inertia I (relative to the mode inertia at $\nu_{\rm ref}$):

$$\Delta = \nu_{\rm obs} - \nu_{\rm mod} = a \left(\frac{\nu_{\rm mod}}{\nu_{\rm ref}} \right)^b / I. \tag{6}$$

This modification provides a better fit to solar observations. Although the B14 correction has some theoretical justification (Gough 1990; Goldreich et al. 1991), it remains highly model-dependent and should therefore be applied cautiously (see Sec. 2.7).

The χ^2 statistic for the observed frequencies of KIC 8410637 is calculated using the K08 and B14 surface corrections, with $\nu_{\rm max}$ as the reference frequency for the model frequencies. The search is constrained to models satisfying $0.95 < \Delta\nu_{\rm c,mod}/\Delta\nu_{\rm c,obs} < 1.05$. To enhance the grid resolution, random pairs of mass and radius values are selected within the range of the searched models, and bilinear interpolation is performed on the model parameters and frequencies using the four nearest models on adjacent evolutionary tracks (see Appendix B).

The resulting χ^2 distribution using the K08 correction is shown in Figure 1a, revealing a minimum at $\chi^2 = 0.32$. The corresponding best-fit model has a mass of $1.65 M_\odot$ and a radius of $11.0 R_\odot$. While precise uncertainties are challenging to determine, they are approximated as the range spanned by models with $\chi^2 < 1$, yielding uncertainties of approximately $\pm 0.25 M_\odot$ and $\pm 0.6 R_\odot$.

The results obtained with the B14 correction are similar (see Table 2), consistent with the findings of Ball et al. (2018), who concluded that best-fit model parameters derived solely from frequency constraints are generally overestimated compared to dynamical parameters.

In summary, the classical approach of matching observed frequencies to model frequencies, combined with standard surface corrections, is fraught with challenges and often results in misleading outcomes. To address these issues, a more robust and unbiased treatment of systematic effects is necessary, as outlined in the following sections.

2.4 Bayesian seismic Forward Modelling – BaseFM

To consistently incorporate prior information about the investigated star and to leverage the benefits of Bayesian statistics, it is advantageous to perform model fitting within a probabilistic framework. Kallinger et al. (2010b) formulated asteroseismic model fitting using Bayes' theorem, where the posterior probability of a model ($\mathcal{M}$) predicting some data ($\mathcal{D}$) given prior information ($\mathcal{I}$) is defined as:

$$p(\mathcal{M}|\mathcal{D},\mathcal{I}) = \frac{p(\mathcal{M}|\mathcal{I}) \cdot p(\mathcal{D}|\mathcal{M},\mathcal{I})}{p(\mathcal{D}|\mathcal{I})}, \tag{7}$$

where $p(\mathcal{M}|\mathcal{I})$ represents the prior probability of a specific model. This prior can encode additional information, such as constraints on the star's fundamental parameters. If no prior information is available, a uniform prior of $1/N_{\rm mod}$ is used, where $N_{\rm mod}$ is the total number of models in the grid. The likelihood function,

$$p(\mathcal{D}|\mathcal{M},\mathcal{I}) \equiv \mathcal{L} = \prod_{i=1}^{N} \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(\nu_{i,\rm obs} - \nu_{i,\rm mod})^2}{2\sigma_i^2} \right), \tag{8}$$

combines the probabilities that the model frequencies $\nu_{\rm mod}$ match the observed frequencies $\nu_{\rm obs}$. This is based on the assumption that the observational uncertainties are normally distributed and represents the conjunction of individual probabilities.

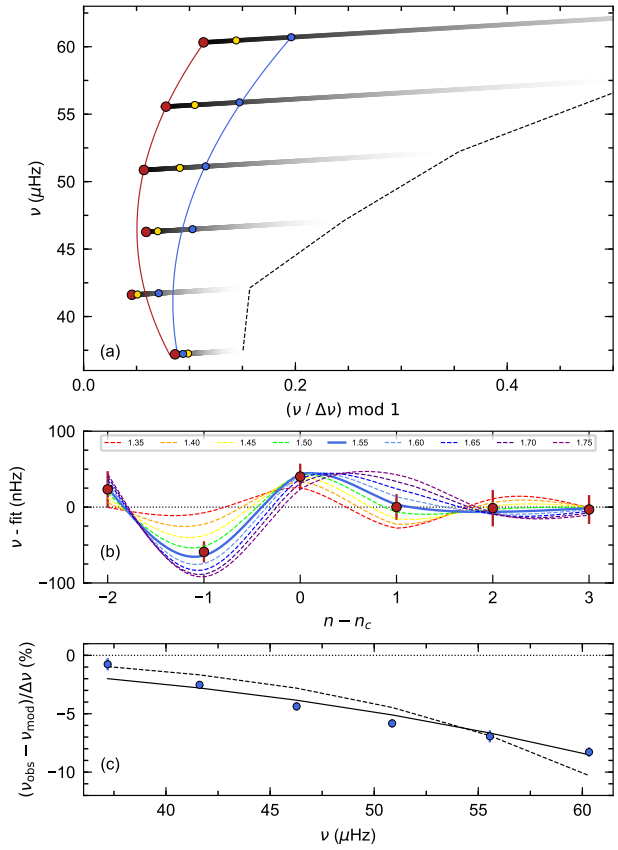

Figure 2. Echelle diagram of KIC 8410637. Panel (a) shows the observed (red symbols) and best-fit model frequencies, where yellow and blue symbols correspond to fits based on Eq. 11 and 16, respectively. Red and blue lines give second-order polynomial fits to the frequencies (i.e., curvature). The grey-shaded trails indicate the prior probability density for varying $\Delta_{\rm i}$, with black and white corresponding to a probability of one and zero, respectively. The dashed line gives $\Delta_{\rm i,max}$. Panel (b) gives the observed frequenices, for curvature corrected, and a sequence of models (colored lines), where instead of the actual model frequencies spline fits are shown for better visibility. Panel (c) shows the (relative) surface effect of the best-fit model frequencies. Full and dashed lines indicate a power-law fit and the original K08 surface correction, respectively.

The denominator in Eq. 7 is a normalization factor for the model probabilities, expressed as:

$$p(\mathcal{D}|\mathcal{I}) = \sum_{i=1}^{N_{\rm mod}} p(\mathcal{M}_i|\mathcal{I}) \cdot p(\mathcal{D}|\mathcal{M}_i, \mathcal{I}). \tag{9}$$

By maximizing Eq. 7, one can identify the best match between the observed and model frequencies. Furthermore, the best-fit model parameters and their uncertainties can be determined by constructing the marginal distributions of the corresponding parameters.

Although the Bayesian approach enables a realistic determination of uncertainties and a statistically sound comparison of different solutions, it still suffers, in its current form, from some of the same limitations as the classical χ^2 method. One of the major strengths of the Bayesian framework, however, is the ability to marginalize over model parameters.

Any parameter θ required to describe a model can be marginalized by applying the sum rule of probability theory (Jaynes 2003), which, in the case of continuous parameters, transforms into an integral. By integrating the full posterior over the parameter range of θ, the marginal posterior is obtained:

$$p(\mathcal{M}|\mathcal{D}, \mathcal{I}) = \int p(\mathcal{M}|\mathcal{D}, \mathcal{I}, \theta)\, p(\theta)\, d\theta. \tag{10}$$

This marginal posterior retains the overall effects of including θ in the model but becomes independent of any specific value of θ. In other words, the parameter θ is effectively removed from the detailed analysis, simplifying the interpretation while preserving the contributions of θ to the model's behavior.

Gruberbauer et al. (2012) demonstrated that this Bayesian approach can be effectively applied to account for unknown systematic errors in seismic model fitting. They showed that if $\Delta_{\rm i}$ represents the absolute value of the systematic error for an individual model frequency, and $\gamma = \pm 1$ specifies whether the model frequency is expected to be systematically lower or higher than the observed frequency, Eq. 8 can be modified as:

$$\mathcal{L} = \prod_{\rm i}^{N_{\rm obs}} \int_{\Delta_{\rm i,min}}^{\Delta_{\rm i,max}} p(\Delta_{\rm i}|\mathcal{I}) \times \frac{1}{\sigma_{\rm i}\sqrt{2\pi}} \exp\left(\frac{-(\nu_{\rm i,obs} - \nu_{\rm i,mod} - \gamma\Delta_{\rm i})^2}{2\sigma_{\rm i}^2}\right) d\Delta_{\rm i}. \tag{11}$$

To marginalize (i.e., to remove) the influence of the surface effect on the frequency comparison, $\Delta_{\rm i}$ is treated as an unknown but consistently incorporated parameter within the probabilistic framework, provided its lower and upper bounds can be reasonably estimated. The prior probability, $p(\Delta|\mathcal{I})$, encodes prior knowledge about the expected behavior of the systematic error, with the condition that:

$$\int p(\Delta_{\rm i}|\mathcal{I})\, d\Delta_{\rm i} = 1. \tag{12}$$

Gruberbauer et al. (2012) further argued that any prior used for the systematic error parameter $\Delta_{\rm i} \in [0, \Delta_{\rm i,max}]$ should be explicitly defined by its parameter limits and free from additional parameters that could influence the evidence. They proposed a simple beta prior:

$$p(\Delta_{\rm i}|\mathcal{I}) = \frac{2(\Delta_{\rm i,max} - \Delta_{\rm i})}{\Delta_{\rm i,max}^2}. \tag{13}$$

This prior defines a linearly decreasing and properly normalized probability density that reaches zero at $\Delta_{\rm i,max}$, ensuring a well-constrained treatment of the systematic error.

The final step is to define the parameter range for $\Delta_{\rm i}$. Tests have shown that this is not particularly critical, provided the actual $\Delta_{\rm i}$ is included within the chosen range. For radial modes, one could, in principle, set $\Delta_{\rm i,max} = \Delta\nu$. However, since the surface effect is known to increase at higher radial orders, it is reasonable to allow $\Delta_{\rm i,max}$ to grow with radial order. Following Gruberbauer et al. (2012), a power law is adopted:

$$\Delta_{\rm i,max} = a \cdot \Delta\nu \left(\frac{\nu_{\rm i,obs}}{\nu_{\rm max}}\right)^b, \tag{14}$$

where the parameters are fixed to $a = 0.3$ and $b = 4.9$. The results are, however, quite robust with respect to the exact values of a and b. Even altering the functional form (e.g., switching to a Lorentzian) only marginally affects the outcome.

The formalism described above is applied to the observed frequencies of KIC 8410637, where instead of a conventional grid search a nested sampling approach is used (see Appendix B). The resulting posterior model probabilities are shown in Figure 1b. Similar to the χ^2 approach, models with approximately equal mean density form a ridge of locally high probability that intersects the uncertainty box of KIC 8410637. However, the actual best-fit values of $M = 1.63 \pm 0.2 M_\odot$ and $R = 10.9 \pm 0.5 R_\odot$ are again larger than the dynamical parameters.

In other words, the surface-correction-independent approach correctly identifies the star's mean density (i.e., the scale of the mode pattern) but fails to accurately match the mass and radius (i.e., the absolute frequencies). This outcome is not unexpected, as the method favors model frequencies with small surface effects (via Eq. 13), which are not necessarily the most appropriate frequencies (see Figure 2). Additionally, the discriminative power along the ridge of equal mean density is limited, resulting in relatively large uncertainties for M and R. This limitation arises because the marginalization of Δ_i assigns similar weights to a broad range of models. Gruberbauer et al. (2013) addressed this issue by incorporating prior information about the star's fundamental parameters into the calculation of posterior probabilities.

However, I choose to avoid this approach, as the model's $M - L - T_\mathrm{eff}$ relationship is heavily influenced by assumptions about α_MLT and the primordial chemical composition, making it highly model-dependent. Instead, my goal is to constrain the models using only the observed frequencies.

2.5 Bayesian forward modelling including the glitch signal

While the average frequency difference between consecutive modes of the same spherical degree (i.e., the large frequency separation) serves as a measure of the mean density of a star (e.g., Tassoul 1980), periodic deviations from the regular mode pattern are indicative of acoustic glitches caused by sharp structural variations in the stellar interior (e.g., Gough 1990). Any sharp variation in the sound speed profile introduces an oscillatory component into the frequencies, known as the glitch signal. The period of this modulation is related to the sound travel time from the structural feature to the stellar surface, while its amplitude depends on the width of the glitch. The amplitude also decreases with increasing frequency because, at higher frequencies, the wavelength of a mode becomes comparable to or smaller than the extent of the glitch.

In red giants, the primary source of observable discontinuities is the second helium ionization zone. Miglio et al. (2010) found that the relative acoustic radius of this zone varies with mass in their models. From an observational perspective, Vrard et al. (2015) demonstrated that the phase of the modulation due to the second helium ionization is also mass-dependent. Consequently, it is reasonable to infer that the glitch signal in the observed frequencies provides additional information that can help more effectively disentangle

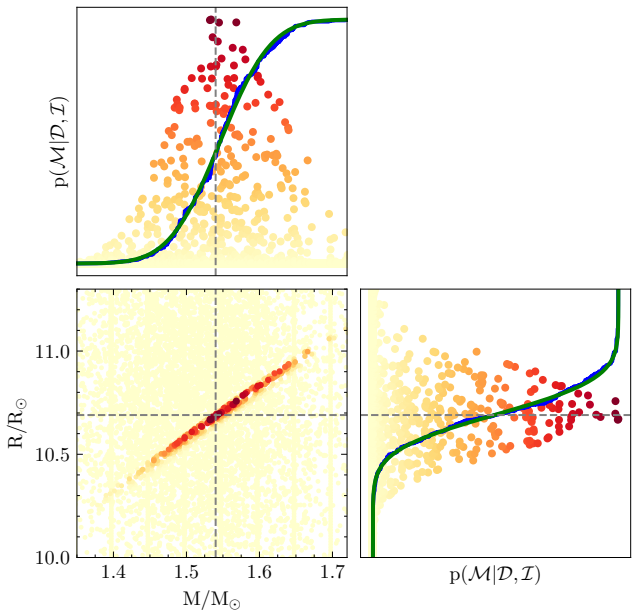

Figure 3. Posterior (colored dots) and cummulative posterior (blue lines) probability distributions for the mass and radius of the Bayesian model fit considering the glitch signal to the observed frequencies of KIC 8410637. Green lines indicate error functions fitted to the cumulative posterior probabilities. Dashed grey lines mark the best-fit values (i.e., where the error functions equal 0.5).

mass and radius compared to previous analyses. For further details, see Appendix C.

In addition to the glitch signal, the frequencies in a red-giant oscillation spectrum are modulated by the so-called curvature. This curvature can be described by the second-order development of asymptotic theory (e.g., Mosser et al. 2011) and is likely caused by hydrogen ionization just below the stellar surface.

To separate the glitch signal from the curvature in the observed and model frequencies (which also include a modulation from the surface effect), I follow Kallinger et al. (2018, hereafter K18) and subtract a second-order polynomial fit from the original frequencies:

$$\nu'(n) = \nu(n) - \left[\nu_\mathrm{c} + \Delta\nu_\mathrm{cor}\left(n - n_\mathrm{c} + \frac{\alpha}{2}(n-n_\mathrm{c})^2\right)\right], \quad (15)$$

where ν_c and n_c are the frequency and radial order of the central radial mode, α is the curvature parameter, and $\Delta\nu_\mathrm{cor}$ is the curvature-corrected large frequency separation. In practice, it is not necessary to know the absolute radial orders of the individual modes. Instead, their relative order $(n - n_\mathrm{c})$ with respect to ν_c is sufficient for the analysis.

One can now add a term to Eq. 11:

$$\begin{aligned}\mathcal{L} = \prod_\mathrm{i}^{N_\mathrm{obs}} & \int_{\Delta_\mathrm{i,min}}^{\Delta_\mathrm{i,max}} p(\Delta_\mathrm{i}|\mathcal{I}) \\ & \times \frac{1}{\sigma_\mathrm{i}\sqrt{2\pi}} \exp\left(\frac{-(\nu_\mathrm{i,obs} - \nu_\mathrm{i,mod} - \gamma\Delta_\mathrm{i})^2}{2\sigma_\mathrm{i}^2}\right) d\Delta_\mathrm{i} \\ & \times \frac{1}{\sigma'_\mathrm{i}\sqrt{2\pi}} \exp\left(\frac{-(\nu'_\mathrm{i,obs} - \nu'_\mathrm{i,mod})^2}{2\sigma'^2_\mathrm{i}}\right), \end{aligned} \quad (16)$$

which compares the glitch contributions of the observed and

Table 2. Mass and radius for KIC 8410637 using different methods to constrain a best fit between the observed and model frequencies. The parameters a (in units of $\Delta\nu$) and b are determined from power-law fits (Eq. 5 and 6) to the difference between the observed and best-fit model frequencies with † indicating fixed exponents. Values in parentheses are for a grid with the frequencies computed from `MESA` models without the Eddington-grey atmosphere (see Sec. 2.7).

	$M/M_\odot$	$R/R_\odot$	a_{Eq5}	b_{Eq5}	a_{Eq6}	b_{Eq6}
dynamic solution[(a)]	$\mathbf{1.47 \pm 0.02}$	$\mathbf{10.60 \pm 0.05}$	-	-	-	-
dynamic solution[(b)]	1.56 ± 0.03	10.74 ± 0.11	-	-	-	-
nonlinear scaling[(c)]	1.53 ± 0.07	10.84 ± 0.18	-	-	-	-
χ^2 with K08 correction	1.65 ± 0.25	11.00 ± 0.60	-0.013	$4.9^\dagger$	-	-
Bayesian fit with K08 correction	1.67 ± 0.07	11.00 ± 0.15	-0.015	$4.9^\dagger$	-	-
χ^2 with B14 correction	1.75 ± 0.20	11.30 ± 0.50	-	-	-0.011	$3.0^\dagger$
Bayesian fit with B14 correction	1.77 ± 0.10	11.30 ± 0.20	-	-	-0.017	$3.0^\dagger$
	$(1.71 \pm 0.10$	11.10 ± 0.25	-	-	-0.015	$3.0^\dagger)$
Bayesian fit based on Eq. 11	1.65 ± 0.25	10.97 ± 0.60	-0.022 ± 0.004	2.6 ± 0.9	-0.021 ± 0.004	-1.3 ± 0.9
	$(1.60 \pm 0.30$	10.80 ± 0.60	-0.016 ± 0.004	1.8 ± 1.6	-0.016 ± 0.004	$1.1 \pm 1.6)$
Bayesian fit based on Eq. 16	$\mathbf{1.55 \pm 0.05}$	$\mathbf{10.71 \pm 0.12}$	-0.053 ± 0.003	2.5 ± 0.3	-0.049 ± 0.005	-1.4 ± 0.7
	$(1.53 \pm 0.07$	10.66 ± 0.17	-0.044 ± 0.004	2.8 ± 0.5	-0.049 ± 0.003	$2.1 \pm 0.3)$

[(a)] Themeßl et al. (2018) – [(b)] Frandsen et al. (2013) – [(c)] Kallinger et al. (2018)

model frequencies, where the uncertainties σ' result from Eq. 15.

In practice (and if $\gamma = -1$), the integral (I) in Eq. 11 and 16 is solved analytically as:

$$I = \frac{2}{B^2\sqrt{2\pi}}\left((A+B)\sqrt{\frac{\pi}{2}}\left[\mathrm{erf}\left(\frac{A+B}{\sigma\sqrt{2}}\right) - \mathrm{erf}\left(\frac{A}{\sigma\sqrt{2}}\right)\right]\right. \tag{17}$$

$$\left. + \sigma\left[e^{\frac{-A^2}{2\sigma^2}} - e^{\frac{-(A+B)^2}{2\sigma^2}}\right]\right), \tag{18}$$

with $A = \nu_{\mathrm{obs}} - \nu_{\mathrm{mod}}$ and $B = \Delta \cdot \Delta_{\mathrm{max}}$ (with Δ_{max} from Eq. 14), which significantly speeds up the likelihood computation compared to numerical integration.

The posterior model probabilities are calculated using Eq. 16, and the results are shown in Figure 1c. The best-fit model is now much better constrained and closely matches the dynamical parameters of KIC 8410637. The corresponding model parameters and their 1σ uncertainties are determined by fitting an error function to the cumulative posterior probability distributions (see Figure 3), yielding $M = 1.55 \pm 0.05 M_\odot$ and $R = 10.71 \pm 0.12 R_\odot$. The best-fit model frequencies are displayed in Figure 2a.

The discriminative power of the glitch signal is demonstrated in Figure 2b, where the glitch signal of various mass models along the ridge of equal mean density is compared. Clearly, models with masses around $1.5 M_\odot$ reproduce the observations far better than higher-mass models, which are favored by the χ^2 approach and the `BaseFM` fit without considering the glitch signal. Finally, Figure 2c shows the surface effect of the best-fit model, parameterized by fitting Eqs. 5 and 6 to the differences between the observed and model frequencies. The resulting parameters are listed in Table 2.

If the final best-fit model is considered a more accurate representation of KIC 8410637 than those obtained from the classical χ^2 approach or the `BaseFM` fit without considering the glitch signal, its surface effect is significantly stronger than the estimates derived using the K08 and B14 corrections. This is evident in the fractional surface correction at ν_{max} (i.e., the parameter a), which is approximately three times larger than expected. Furthermore, the frequency-dependent increase (i.e., the parameter b) is incompatible with standard surface corrections.

The results presented here are in quantitative agreement with those of Ball et al. (2018), who also found that seismic modeling using standard surface corrections fails to reproduce the dynamical mass and radius of KIC 8410637 when relying solely on the observed frequencies. In contrast, Buldgen et al. (2019) determined a seismic mass and radius using the B14 correction that are consistent with the dynamical values.

However, a closer inspection reveals that they included the radial mode at approximately $32.8\,\mu$Hz in their forward modeling (also used by Ball et al. 2018), which is found to be formally insignificant in the present peak-bagging analysis (see Sec. 2.1). To investigate the influence of this mode on the analysis, I repeated the model fitting with the mode included. The results show that the mode has a marginal impact on the `BaseFM` outcome but significantly "improves" the results from classical approaches by bringing the seismic parameters closer to the dynamical ones.

For example, including this mode changes the mass estimate from approximately $1.67 M_\odot$ to $1.55 M_\odot$ (a shift of nearly 8%) when the K08 correction is applied, and the frequencies are matched within the Bayesian framework defined by Eqs. 7 to 9. This level of sensitivity to a single mode is quite surprising but does not explain why Ball et al. (2018) and Buldgen et al. (2019) obtained substantially different best-fit model parameters (e.g., around $1.75 M_\odot$ and $1.56 M_\odot$, respectively) despite using nearly the same set of observed frequencies and a similar method for matching them to models.

In summary, the formalism described above provides a straightforward approach to identifying a best-fit model using *only* the observed frequencies, without requiring more than very *basic* assumptions about the scale and functional form of the surface effect. This approach yields realistic results along with well-constrained uncertainties. By eliminating the need for empirical surface corrections, it represents a significant advancement in asteroseismic model fitting.

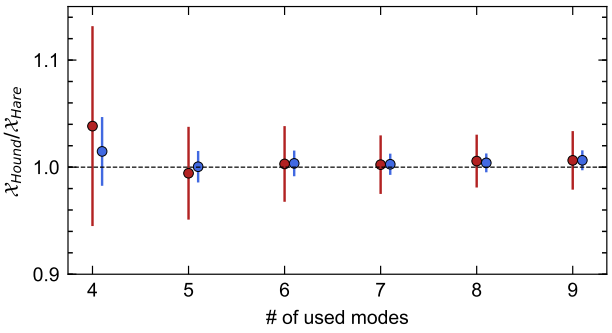

Figure 4. Fractional difference in mass (red) and radius (blue – slightly shifted for better visibility) between the input (hare) and best-fit model parameter (hound) as a function of the number of modes used in the `BaseFM` fit.

2.6 "Hare and hounds" test

To evaluate the reliability and performance of `BaseFM`, a simple "hare and hounds" test is conducted. The radial mode frequencies of a $1.5M_\odot$ and $10R_\odot$ model are used as input. To simulate "observed" frequencies, surface corrections based on Eq. 5 (with $a = -0.04$ and $b = 4.9$) are applied, along with normally distributed uncertainties matching those of the observed frequencies of KIC 8410637. These frequencies are then fitted using `BaseFM`, with the number of considered modes incrementally increased from four (the minimum required for a meaningful curvature correction) to nine (the typical maximum number of observed modes in this region of the HR diagram).

The results are illustrated in Figure 4, showing that `BaseFM` requires only four frequencies to reproduce the correct mass and radius of the input model, albeit with relatively large uncertainties and a slight systematic offset. The fits improve significantly as the number of used modes increases, with uncertainties stabilizing at approximately ±3% in mass and ±1% in radius, and with negligible systematic offset.

Repeating the test with different input frequencies (i.e., modifying the surface effect by changing a and b in Eq. 5) yields only marginal changes to the results, demonstrating the robustness of the method.

2.7 Model-dependency of the Ball & Gizon (2014) surface correction

In addition to identifying a best-fit model and its parameters, `BaseFM` enables the measurement of the actual surface effect in the model frequencies. To compare this with the predictions from K08 and B14, Eqs. 5 and 6, respectively, are fitted to the differences between the observed and best-fit model frequencies. The resulting scaling factor a and exponent b are listed in Table 2.

Both fits fail to reproduce the surface corrections proposed by K08 and B14, indicating that these corrections are not capable of identifying the correct model for KIC 8410637. Specifically, the K08 correction underestimates the scale of the surface effect (i.e., a is too small) and overestimates its steepness (b is too high). In reality, b should be

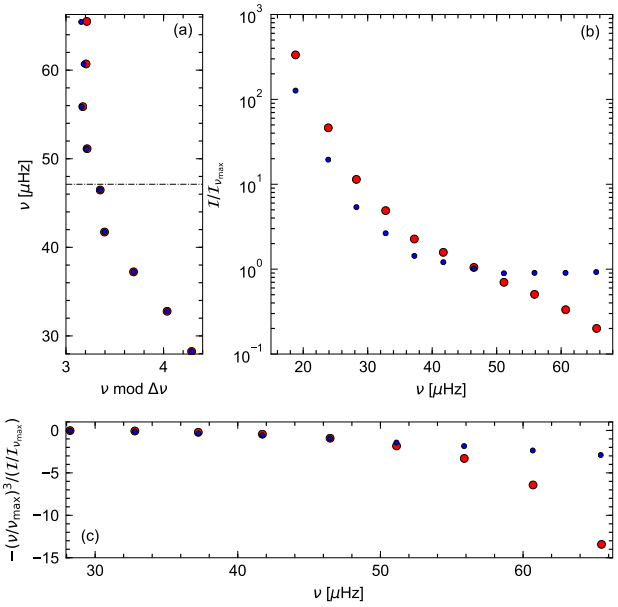

Figure 5. Frequencies and the corresponding mode interia of the best-fit models of KIC 8410637. Panel (a) shows the frequencies in an echelle diagram, where the red and blue symbols correspond to frequenices from the original model grid and those computed for the same model but without an Eddington-grey atmosphere, respectively. The dash-dotted lines indicates the observed ν_{max}. Panel (b) gives the mode inertia normalised by the inertia at ν_{max}. Panel (c) shows the unscaled (i.e. without the scaling factor a) surface correction according to Eq. 6.

closer to three, whereas K08 originally suggested a value of 4.9, based on the solar calibration.

The results for the B14 correction are even more peculiar. Here, a negative exponent ($b \simeq -1.1$) is obtained, which deviates significantly from the originally proposed cubic dependence of the surface correction. To investigate this large discrepancy, I examined the way the model frequencies are calculated.

In order to compute a higher-order estimator for the theoretical asymptotic frequency separation (Ong & Basu 2019a,b), the standard `MESA` output (see Sec. 2.2) needs to be modified before calculating the frequencies with `GYRE`. While this modification has minimal impact on the frequencies themselves (less than 0.5% of $\Delta\nu$ in the case of KIC 8410637), it significantly alters the mode inertia (see Figure 5). Specifically, the mode inertia decreases much more rapidly with frequency than in "standard" models (i.e., frequencies computed for `MESA` models with the default output). For the modified models, the mode inertia for the highest observable mode is about 25 times smaller than for the lowest observable mode, compared to a factor of approximately 3 in the standard models.

As a result, the frequency corrections using the B14 cubic surface correction are significantly larger than those for the standard models (see Figure 5). This discrepancy necessitates compensation by a smaller (or even negative) exponent in Eq. 6. One could argue that this behavior arises from the specific method used to calculate the model frequencies, and the B14 correction might still work for standard models.

To test this, I constructed a model grid based on the standard `MESA` output and repeated the model fitting. While this improved the situation, the exponent obtained when fitting Eq. 6 to reproduce the actual surface effect of KIC 8410637 was 2.1 ± 0.3, still significantly smaller than the originally suggested value of three.

These results clearly demonstrate that directly applying surface corrections developed for the Sun to evolved stars is highly problematic, introducing unknown systematics into the results (see also Ball et al. 2018).

3 SOME EXAMPLES

3.1 Red giant in the SB2 system KIC 4054905

Benbakoura et al. (2021) identified 16 new eclipsing binaries observed by *Kepler* that host a red giant. Among them is the SB2 system KIC 4054905, for which they solved the binary orbit and determined dynamical masses and radii for both components. Given the relatively low measurement accuracy, Brogaard et al. (2022) conducted a follow-up study to better characterize KIC 4054905. They determined the system consists of two stars with nearly identical masses (approximately $0.95 M_\odot$) in a 274.7-day eccentric orbit. The primary star is a red giant with a radius of $8.36 \pm 0.03 R_\odot$, while the secondary is a subgiant with a radius of $3.09 \pm 0.01 R_\odot$. Additionally, they measured the primary's effective temperature and metallicity as $T_{\rm eff} = 4850 \pm 70$K and $[{\rm Fe/H}] = -0.60 \pm 0.02$, respectively, which is consistent with values of 4740 ± 100 and -0.46 ± 0.04 from the APOKASC catalog (Pinsonneault et al. 2018). From a seismic perspective, the *Kepler* photometry reveals a clear oscillation spectrum only for the red giant component.

Based on an analysis of the mixed-mode spectrum, Benbakoura et al. (2021) classified the red giant as a core helium-burning star, suggesting it could be a rare system that underwent mass transfer during the primary's evolution through the tip of the RGB (e.g., Verbunt & Phinney 1995). However, this classification is inherently ambiguous. The determined asymptotic period spacing ($\Delta\Pi_1 \simeq 160$s) would place the star in the secondary clump given its relatively high $\Delta\nu$ of about $5.4\,\mu$Hz. However, the star's mass is far too low for it to be a secondary clump star (e.g., Mosser et al. 2014). While one might argue that this could result from the star's low metallicity, a more plausible explanation is that the star is still ascending the giant branch. This conclusion is supported by classification methods from Elsworth et al. (2017) and Kallinger et al. (2012), which clearly place the star in the RGB regime.

A closer inspection of the observed dipole modes shows they are consistent with $\Delta\Pi_1 = 59 \pm 2.5$s, a value obtained by fitting Eq. 13 from Mosser et al. (2015) to the observed period spacings (see Sec. 5). This is illustrated in Figure 6, where the observed dipole modes of KIC 4054905 are compared with the expected mode patterns for $\Delta\Pi_1 = 59$s (typical for an RGB star with $\Delta\nu = 5.4\,\mu$Hz) and 160s.

The low metallicity of KIC 4054905 corresponds to a metal mass fraction of approximately Z = 0.006. To facilitate further analyses and applications, I computed a new metal-poor model grid consisting of about 129,000 models with an initial chemical composition of (Y, Z) = (0.26, 0.010) and $\alpha_{\rm MLT} = 2.2$. The grid spans a mass range from 0.7 to $2.0 M_\odot$in steps of $0.05 M_\odot$, covering the evolution from the ZAMS to the post-core helium-burning phase. Details of the grid parameters are listed in Table 1, with the modeling procedures described in Sec. 2.2.

Table 3. Stellar mass and radius of KIC 4054905 from various sources. `BaseFM` fits are based on the model grids from Table 1. The last column gives the resulting fit probability in percent when applying the $\Delta\Pi_1$ prior.

	$M/M_\odot$	$R/R_\odot$	p
dynamic solution[(a)]	0.95 ± 0.01	8.36 ± 0.03	
nonlinear scaling[(b)]	0.95 ± 0.05	8.37 ± 0.15	
G1	0.97 ± 0.03	8.24 ± 0.09	88
G2	0.94 ± 0.03	8.12 ± 0.08	12
G3	0.88 ± 0.04	7.94 ± 0.13	$\ll 1$

[(a)] Brogaard et al. (2022) – [(b)] Kallinger et al. (2018)

Using the seven consecutive radial modes from Kallinger (2019), `BaseFM` identifies a best-fit model with a mass and radius that agree with the dynamical values to within approximately 2% (see Table 3). Additionally, the model-independent corrected seismic scaling relations of Kallinger et al. (2018) yield consistent results, using $\nu_{\rm max} = 48.8 \pm 0.2\,\mu$Hz and $\Delta\nu_{\rm cor} = 5.39 \pm 0.01\,\mu$Hz as inputs [6].

A valuable feature of the Bayesian framework employed by `BaseFM` is its ability to compare different fits (e.g., using different model grids) and quantify which model reproduces the observations most effectively. This is achieved by comparing the so-called global evidence, z, defined as the integrated posterior model probability from Eq. 7. Similar to the approach in Appendix A, different fits with varying z values can be assigned probabilities according to $p_i = z_i / \sum_j z_j$.

For KIC 4054905, the metal-poor grid yields a logarithmic global evidence of $\ln z = 3.55$. In comparison, fitting the frequencies to the original grid described in Sec. 2.2 results in a global evidence of $\ln z = 3.56$. Comparing the best-fit parameters (see Table 3) with the dynamical values shows that the metal-poor models provide a better match. However, when ignoring the dynamical measurements, the probability contrast between the two fits, $p_{G1} : p_{G2} \approx 0.49 : 0.51$, is insufficient to draw any definitive conclusions about the metallicity of KIC 4054905.

In other words, for this example, using only the observed frequencies enables accurate determination of the star's mass and radius but does not allow conclusions beyond these parameters. Nevertheless, other properties of the two best-fit models differ significantly. For instance, the metal-poor grid produces a best-fit $\Delta\Pi_1$ that deviates by more than 10% from that of the near-solar metallicity grid. To incorporate this discrepancy a Gaussian prior is introduced:

$$p(\mathcal{M}|\mathcal{I}) = \frac{1}{\sigma_{\mathcal{X}}\sqrt{2\pi}} \exp\left(-\frac{(\mathcal{X}_{\rm mod} - \mathcal{X}_{\rm obs})^2}{2\sigma_{\mathcal{X}}^2}\right), \qquad (19)$$

[6] Note that these values differ from those reported by Benbakoura et al. (2021) and Brogaard et al. (2022). As emphasized by Kallinger et al. (2014, 2018), consistent definitions of $\nu_{\rm max}$ and $\Delta\nu$ must be used throughout the process – from parameter extraction to seismic scaling relations (including solar references) – to avoid systematic uncertainties in the results.

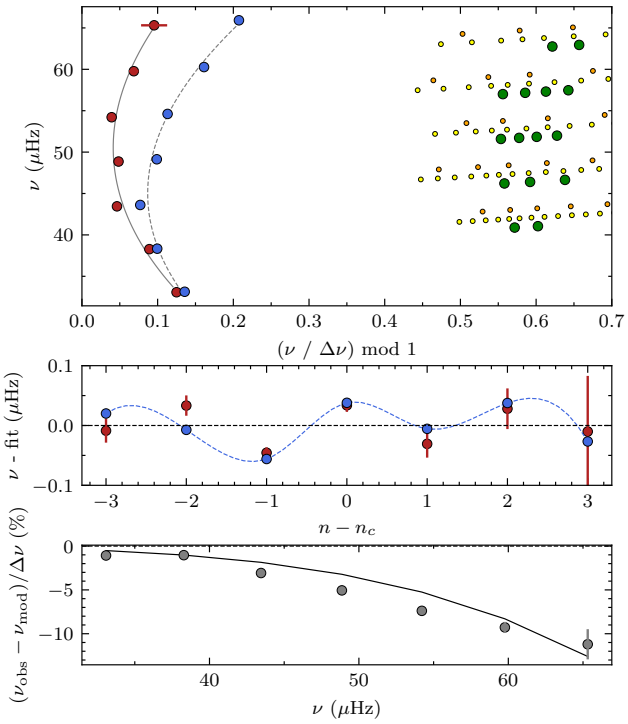

Figure 6. Same as Fig. 2 but for KIC 4054905. Green dots in the top panel indicate the observed dipole modes, which are compared to expected modes based on a $\Delta\Pi_1$ of 59 (yellow) and 160 s (orange). The latter are arbitrarily shifted in vertical direction for better visibility.

where $\mathcal{X}$ represents the target parameter (in this case, $\Delta\Pi_1$) and its uncertainty. In principle, any model parameter – or combinations of parameters – can be used. The current version of `BaseFM` includes the option to implement Gaussian priors for age, mass, radius, luminosity, effective temperature, and $\Delta\Pi_1$. With minor modifications to the code, other types of priors can be easily incorporated to accommodate specific requirements.

For the given example, applying a $\Delta\Pi_1$ prior of 59 ± 2.5s has only a marginal effect on the best-fit model's mass and radius but significantly alters the resulting global evidences. To extend the analysis, I also fit the frequencies to a metal-rich grid (see Sec. 4.1) and arrive at a final probability contrast of $p_{G1} : p_{G2} : p_{G3} \approx 0.883 : 0.117 :\sim 10^{-5}$, which clearly favors the metal-poor models.

In conclusion, a set of observed radial mode frequencies, combined with a rough estimate of $\Delta\Pi_1$, is sufficient for `BaseFM` to determine an unbiased and accurate mass and radius, as well as to identify whether a star is more metal-poor, solar-like, or metal-rich – without requiring any assumptions about the surface effect.

The seismic analysis of KIC 4054905 provides a very clear picture, allowing one to confidently assume that the identified best-fit model is a reliable representation of the actual star. It is therefore notable that the surface effect of this model remains irreproducible by the classical approaches. A power-law fit to the difference between the observed and best-fit model frequencies yields an exponent of 4.2 ± 0.4 for the K08 approach and -0.04 ± 0.03 for the B14 correction.

Table 4. Same as Table 3 but for KIC 9970396.

	$M/M_\odot$	$R/R_\odot$	p
dynamic solution[(a)]	1.18 ± 0.02	8.04 ± 0.07	
nonlinear scaling[(b)]	1.17 ± 0.05	8.06 ± 0.12	
G1	1.18 ± 0.04	7.92 ± 0.09	72
G2	1.16 ± 0.04	7.86 ± 0.09	27
G3	1.13 ± 0.03	7.80 ± 0.08	1

[(a)] Brogaard et al. (2018) – [(b)] Kallinger et al. (2018)

These results are consistent with the findings in Sec. 4.4. Forcing the surface effect to adhere to these classical approaches will inevitably introduce systematic uncertainties, as discussed in Sec. 3.6.

3.2 Red giant in the SB2 system KIC 9970396

The third example I present here is the SB2 system KIC 9970396, which hosts a red giant orbited by a main-sequence in about 235 days. This system was initially part of the sample studied by Gaulme et al. (2016), but Brogaard et al. (2018) subsequently refined its characterization. Their analysis determined the red giant component to have a mass of $1.18\pm0.02M_\odot$ and a radius of $8.04\pm0.07R_\odot$ (see Table 4). Additionally, they measured the star's effective temperature and metallicity as $T_{\rm eff} = 4860\pm80$ K and $[{\rm Fe/H}] = -0.20\pm0.02$, respectively, consistent with the APOKASC catalog values of 4795 ± 75 K and -0.16 ± 0.03 reported by Pinsonneault et al. (2018).

The seismic analysis of the *Kepler* time series, detailed in Appendix A2, reveals a sequence of six consecutive radial modes (Table A2). Using these modes, `BaseFM` identifies a best-fit model from the metal-poor grid G1 (Table 1) that perfectly matches the dynamic mass while underestimating the dynamic radius by only $1.5\pm1.4\%$. Although fits to grids G2 and G3 yield slightly poorer matches (Table 4), the probability contrast remains insufficient to definitively identify the best-fit grid – particularly when disregarding both the dynamic benchmarks and the known metal-poor nature of the star.

KIC 9970396 exhibits an even more densely populated dipole-mode spectrum than KIC 4054905, allowing for straightforward determination of $\Delta\Pi_1$ as 68 ± 1.5s. Incorporating this as a prior scarcely alters the best-fit model parameters but improves the probability contrast to $p_{G1} : p_{G2} : p_{G3} \approx 0.72 : 0.27 : 0.01$. This result effectively rules out grid G3 while moderately favoring grid G1. The best-fit model frequencies are illustrated in Figure 7.

As with previous examples, the actual surface effect of the best-fit model frequencies cannot be reproduced by classical approaches (Figure 7). Power-law fits to the frequency differences yield exponents of 3.7 ± 0.2 for the K08 approach and -0.07 ± 0.6 for the B14 correction, both significantly deviating from their canonical values of 4.9 and 3, respectively.

3.3 The planet-host red giant HR 6817

Another well-suited test object is the planet-host red giant HR 6817 (e.g., Johnson et al. 2008). Given its relative bright-

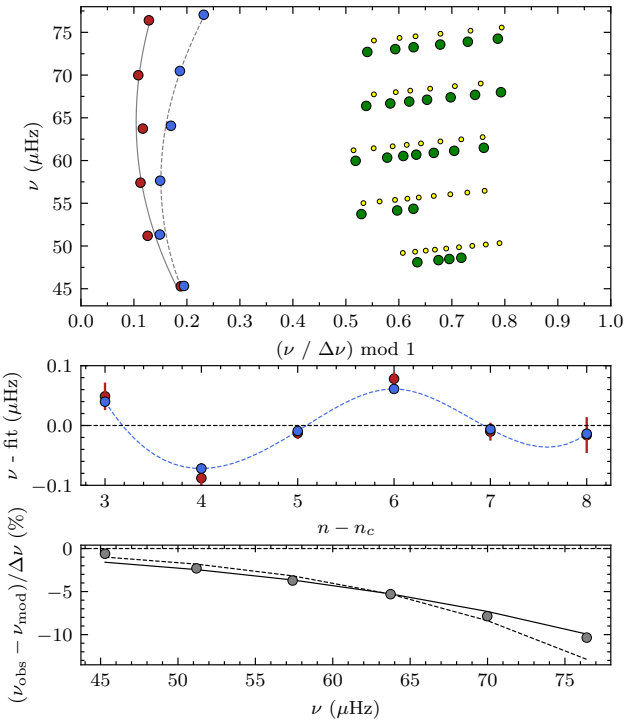

Figure 7. Same as Fig. 2 but for KIC 9970396. Yellow dots give expected dipole modes based on $\Delta\Pi_1 = 68$ s vertically shifted for better visibility.

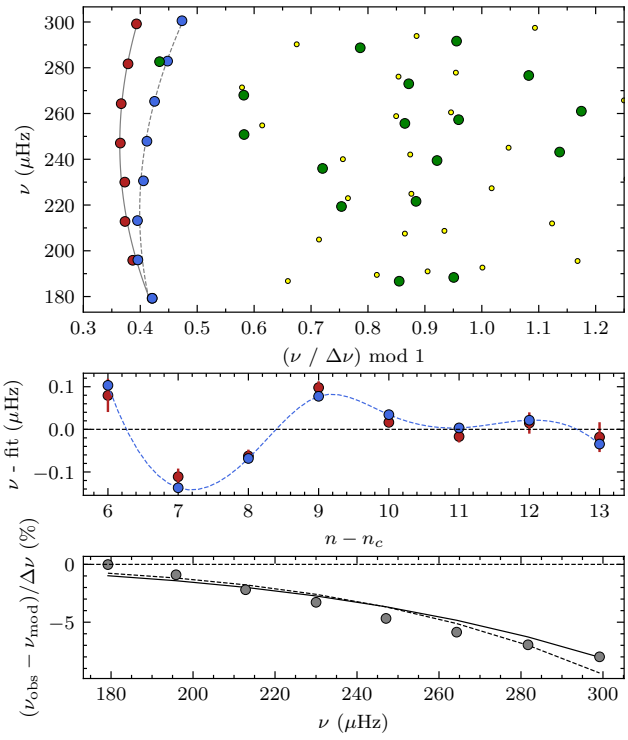

Figure 8. Same as Fig. 7 but for HR 6817 and $\Delta\Pi_1 = 82$ s.

Table 5. Mass, radius, and luminosity of HR 6817 as determined by interferometry and spectral energy distribution fitting (first line – see text for the details), nonlinear seismic scaling relation, and `BaseFM` fits to model grids from Tab. 1. The last column gives the resulting fit probability in percent.

	$M/M_\odot$	$R/R_\odot$	$L/L_\odot$	p
		4.39 ± 0.06	11.0 ± 0.4	
(a)	1.40 ± 0.04	4.40 ± 0.04	10.7 ± 0.6	
G1	1.37 ± 0.03	4.32 ± 0.03	11.8 ± 0.4	77
G2	1.30 ± 0.03	4.24 ± 0.04	9.9 ± 0.3	22
G3	1.19 ± 0.03	4.11 ± 0.03	8.2 ± 0.3	1

(a) nonlinear scalings (Kallinger et al. 2018)

ness and proximity, White et al. (2018) measured a precise interferometric angular diameter of $\theta_{\rm LD} = 0.823 \pm 0.011$mas and a bolometric flux of $F_{\rm bol} = 142.6 \pm 5.5\,{\rm pW\,m^{-2}}$. Using the Gaia DR3 (Gaia Collaboration 2022) parallax of $\varpi = 20.131 \pm 0.029$mas, the stellar radius is determined as $R = \theta_{\rm LD}/(2\varpi) = 4.39 \pm 0.06 R_\odot$, and the luminosity as $L = 4\pi F_{\rm bol}/\varpi^2 = 11.0 \pm 0.4 L_\odot$. These values are consistent with the original estimates from White et al. (2018), which were based on the Hipparcos parallax.

The metallicity of HR 6817 is somewhat uncertain. High-resolution spectroscopic studies by Mortier et al. (2013) and Ghezzi et al. (2018) suggest [Fe/H] values between 0.03 and 0.1, while the APOGEE-2 DR17 (Abdurro'uf et al. 2022) and Gaia DR3 (Gaia Collaboration 2022) catalogs list values as low as -0.16 and -0.41, respectively. Additionally, the LAMOST catalog (Xiang et al. 2019) reports [Fe/H] = -0.23, identifying the star as relatively metal-poor.

The effective temperature of HR 6817 is more consistently determined. As a bright star, there are numerous measurements of $T_{\rm eff}$ available in the literature. After searching the VizieR database and removing obvious outliers, a median $T_{\rm eff} = 4975 \pm 62$K is obtained based on 20 independent sources.

HR 6817 is also located in the Continuous Viewing Zone of NASA's Transiting Exoplanet Survey Satellite (TESS; Ricker et al. 2015) and has been observed in 33 sectors between June 18, 2019, and September 30, 2024. Initial results were presented by Hon et al. (2022), but no individual frequencies have been extracted from these data. For the present analysis, data processing and peak bagging are described in Appendix A3.

Using the eight radial modes listed in Table A3, `BaseFM` identifies best-fit models with radii ranging from 4.11 to $4.32 R_\odot$ (see Table 5), depending on the model grid used. As in the case of KIC 4054905, the probability contrast of the individual fits is insufficient to draw conclusions about the star's metallicity. However, the dipole mode spectrum of HR 6817 allows $\Delta\Pi_1$ to be constrained to 82 ± 2s. Including this as a prior improves the probability contrast of the fits. From a seismic perspective, HR 6817 is more likely a metal-poor star than a near-solar metallicity star.

While not definitive, this conclusion is supported by the fact that the best-fit metal-poor model's mass and luminosity are more consistent with independent measurements than those from other fits (see Table 5). The observed and best-fit model frequencies are shown in Fig. 8.

In summary, `BaseFM` identifies a representative stellar

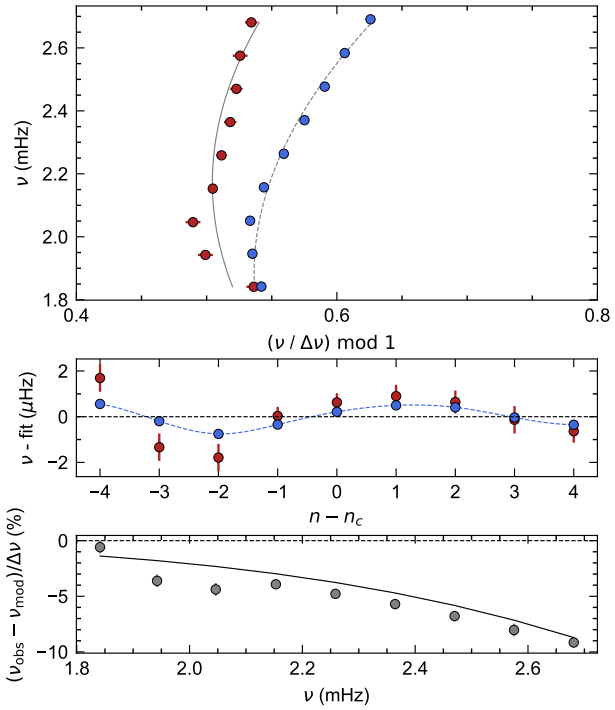

Figure 9. Same as Figure 6 but for α Cen A.

Table 6. Various parameters of α Cen A. References: (1) –de Meulenaer et al. (2010); (2) – Bazot et al. (2007); (3) – Kervella et al. (2016); (4) – Kervella et al. (2017); (5) – Chmielewski et al. (1992); (6) – Eggenberger et al. (2004)

	`BaseFM` fit	literature	source
$\Delta\nu_{\rm cor}$ [μHz]	105.27 ± 0.08	105.9 ± 0.3	(1) (2)
$M/M_\odot$	1.088 ± 0.030	1.106 ± 0.003	(3)
$R/R_\odot$	1.213 ± 0.014	1.223 ± 0.005	(4)
$L/L_\odot$	1.57 ± 0.10	1.52 ± 0.02	(4)
$T_{\rm eff}$ [K]	5857 ± 66	5795 ± 19	(4)
$[{\rm Fe/H}]_s$	-	0.22 ± 0.02	(5)
age [Gyr]	6.2 ± 1.4	6.5 ± 0.3	(6)

model with a radius consistent with the interferometric radius to within one sigma, or 1.6%. The surface effect of this model, however, remains incompatible with classical surface corrections. The difference between the observed and best-fit model frequencies requires a power-law correction term with an exponent of 3.9 ± 0.4, as opposed to the canonical value of 4.9 in the K08 approach.

3.4 The main sequences star α Cen A

Even though `BaseFM` was developed to model evolved red giants, there is no reason it cannot be applied to main-sequence stars as well.

Apart from the Sun, the stars in the α Cen system are likely the solar-type oscillating main-sequence stars with the most accurate, seismology-independent fundamental parameters available (e.g., Guenther & Demarque 2000). This nearby visual binary system consists of two stars exhibiting solar-type oscillations that bracket the Sun in mass and radius. Kervella et al. (2016) revised the orbital parameters of the system and, using the new reduction of the original Hipparcos data (van Leeuwen 2007), determined a mass of $1.106 \pm 0.003 M_\odot$ for the primary component, α Cen A. Compared to older literature values, the updated parallax also affects the radius, which was derived from the interferometric angular diameter (Kervella et al. 2017). A summary of α Cen A's fundamental parameters is provided in Table 6.

For the seismic analysis of α Cen A, a new stellar model grid was computed (see Sec. 2.2), consisting of approximately 12,000 models with masses ranging from 0.9 to $1.3 M_\odot$ (in steps of $0.025 M_\odot$) and radii from 1.0 to $1.5 R_\odot$. Given [Fe/H] = 0.22 (e.g., Chmielewski et al. 1992), the initial chemical composition was set to (Y, Z) = (0.30, 0.03). As with the original grid, $\alpha_{\rm MLT} = 2.1$ was used, and overshooting was turned off.

Following the procedure described in Sec. 2, nine consecutive radial mode frequencies determined by de Meulenaer et al. (2010) were used to search for a best fit in the model grid. Despite the relatively large frequency uncertainties (typically $0.5\,\mu$Hz), `BaseFM` accurately constrained the mass and radius of α Cen A to $1.088 \pm 0.030 M_\odot$ and $1.213 \pm 0.014 R_\odot$, respectively, using only the observed frequencies. These results are in excellent agreement with the binary mass and interferometric radius (see Table 6). While the luminosity and age are less well constrained (uncertainties of approximately ±6% and ±22%, respectively), they remain consistent with independent measurements.

The observed and best-fit model frequencies are illustrated in Figure 9, which again shows that the surface effect (i.e., the frequency difference) is not fully compatible with the K08 power-law correction. A fit to the frequency shifts reveals a power-law exponent of 4.0 ± 0.5, differing from the canonical value of 4.9.

A larger and more precise set of frequencies will soon become available from TESS, allowing for seismic testing of model physics for a star similar to, but different from, the Sun. `BaseFM` is an ideal tool for such analyses, as it is not biased by our incomplete understanding of the surface effect.

Initial tests have shown that with minor modifications[7], `BaseFM` can also be used to fit solar frequencies. However, properly calibrating a solar model grid is a complex task beyond the scope of this paper and will be addressed in a future publication.

So far, only radial modes have been used in forward modeling with `BaseFM`. However, the method can also handle nonradial modes, at least for less-evolved stars where significant mixing between pressure and gravity modes is absent. α Cen A, however, is not an ideal example for this due to the ambiguity of its observed modes caused by daily aliases, which would require a more detailed analysis. A better example is discussed in the following section.

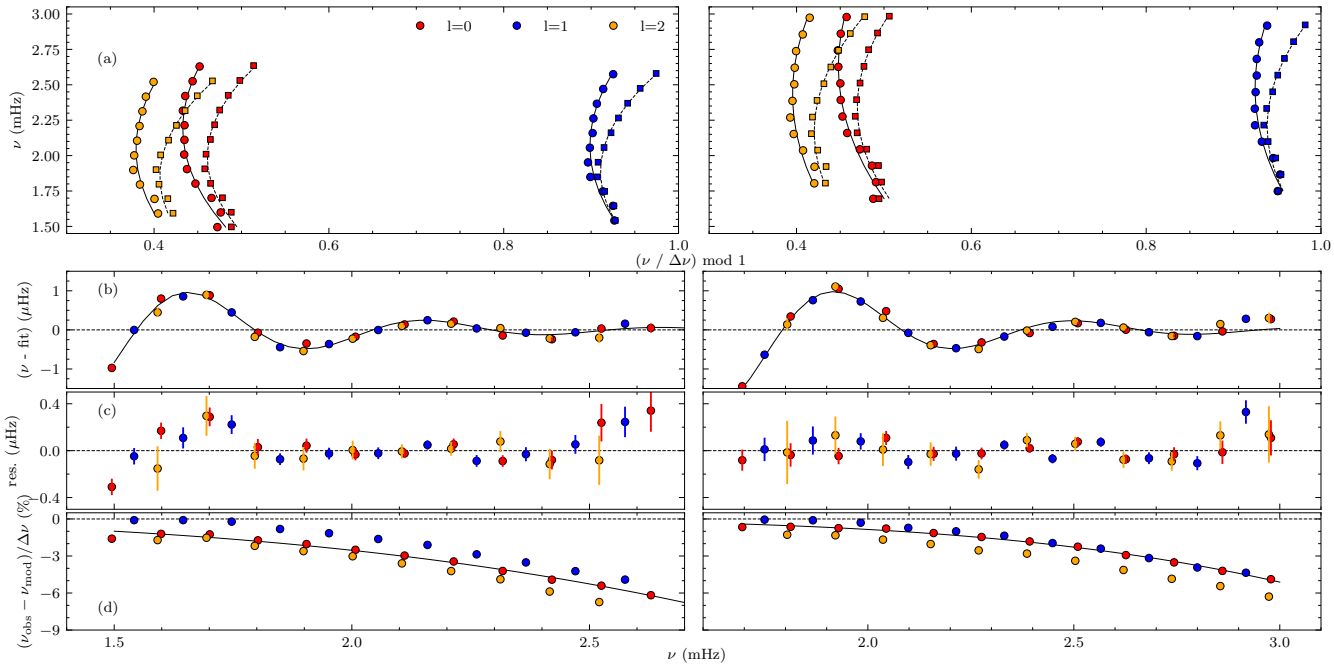

Figure 10. Seismic analysis of the solar analogs 16 Cyg A (left) and B (right). Panels (a) show Echelle diagrams with colored circles and squares indicating the observed and best-fit model frequencies. Lines indicate curvature fits. Panels (b) give the curvature corrected frequenices with black lines corresponding to fits to all modes (see text for an explaination). Panels (c) illustrate the residuals between the curvature corrected observed and model frequencies from panels (b). Panels (d) show the differences between the (uncorrected) observed and model frequencies (i.e., the surface effect) with black lines indicating powerlaw fits to the radial modes.

Table 7. Various parameters of 16 Cyg A&B, where the non-seismic values (first line) for the mass, radius and luminosity are determined here (see text) and only the age (in Gyr) is from Bazot (2020). The `BaseFM` fit is based on a small grid with (Y, Z) = (0.25, 0.024).

	16 Cyg A				16 Cyg B			
	$M/M_\odot$	$R/R_\odot$	$L/L_\odot$	age	$M/M_\odot$	$R/R_\odot$	$L/L_\odot$	age
	1.027 ± 0.024	1.192 ± 0.009	1.51 ± 0.04	6.9 ± 0.1	0.976 ± 0.024	1.087 ± 0.009	1.22 ± 0.04	6.9 ± 0.1
(a)	1.035 ± 0.007	1.211 ± 0.003	1.46 ± 0.02		0.988 ± 0.007	1.097 ± 0.003	1.16 ± 0.02	
`BaseFM` fit	1.033 ± 0.010	1.211 ± 0.010	1.46 ± 0.03	9.6 ± 0.3	0.989 ± 0.010	1.099 ± 0.010	1.14 ± 0.03	10.2 ± 0.4

$^{(a)}$nonlinear scalings (Kallinger et al. 2018)

Table 8. Glitch signal parameters of the two 16 Cyg stars according to Eq. 20.

	A [μHz]	η [μHz]	τ_g [s]	ϕ [rad]
16 Cyg A	89 ± 35	367 ± 31	1006 ± 16	-0.7 ± 0.3
16 Cyg B	114 ± 50	405 ± 37	864 ± 18	5.6 ± 0.4

3.5 The solar analogs 16 Cyg A&B

The two solar analogs 16 Cyg A and B are among the brightest stars in the *Kepler* field and were continuously observed for over 2.5 years. The resulting high-precision time series provided some of the best seismic data available for main-sequence solar-type oscillators (aside from the Sun).

[7] Eq. 15 must be extended to higher-order terms to accurately follow the global trend of solar frequencies.

This data inspired numerous seismic studies (e.g., Metcalfe et al. 2012, 2015; Bellinger et al. 2016; Silva Aguirre et al. 2017; Farnir et al. 2020), leading to a relatively consistent set of stellar parameters. These include seismic masses of 1.080 ± 0.016 and $1.030 \pm 0.015 M_\odot$ for 16 Cyg A and B, respectively (Bellinger et al. 2017), and an estimated system age of 6.88 ± 0.12 Gyr (Bazot 2020).

In addition to seismic data, a set of complementary non-seismic parameters has been established over time. Frequently cited stellar radii are derived from the interferometric measurements of White et al. (2013) combined with Hipparcos parallaxes, though these values are found to be about 2% too large. Updated estimates use more precise interferometric angular diameters from Karovicova et al. (2022), who reported $\theta_{\rm LD} = 0.524 \pm 0.004$ mas for 16 Cyg A and 0.479 ± 0.004 mas for 16 Cyg B. When combined with Gaia DR3 parallaxes of $\varpi = 47.324 \pm 0.020$ mas and 47.320 ± 0.017 mas, these yield refined non-seismic radii of $1.192 \pm 0.009 R_\odot$ and $1.087 \pm 0.009 R_\odot$, respectively (see Table 7). Karovicova et al. (2022) also provided bolometric

flux measurements, which lead to luminosity estimates of $1.51 \pm 0.04 L_\odot$ for 16 Cyg A and $1.22 \pm 0.05 L_\odot$ for 16 Cyg B. While these values are consistent with earlier estimates (e.g., Metcalfe et al. 2012), they are approximately 3% smaller overall.

The two primary components of the 16 Cyg hierarchical triple system have an orbital period exceeding 18,000 years (Hauser & Marcy 1999), making it impossible to estimate their individual dynamic masses. Instead, seismic estimates have been utilised, though these are only marginally consistent with independent constraints. A straightforward approach to estimating a star's mass involves using its surface gravity (g) and radius, based on $M = gR^2$. However, this method often provides only a rough estimate because surface gravity measurements are typically difficult to determine with high precision. This is not the case for the two stars in 16 Cyg, where several independent, consistent, and highly precise $\log g$ values are available in the literature. I adopt measurements from ten sources listed in Table 2 of Ryabchikova et al. (2022), averaging to $\log g = 4.297 \pm 0.008$ for 16 Cyg A and 4.355 ± 0.008 for 16 Cyg B. Using these values, along with the previously determined radii, yields masses of $1.027 \pm 0.024 M_\odot$ for 16 Cyg A and $0.976 \pm 0.024 M_\odot$ for 16 Cyg B[8], which are about 5% smaller than previous estimates. Interestingly, these results align well with the estimates by Hauser & Marcy (1999), who derived masses of 1.05 ± 0.01 and 0.99 ± 0.01 for the two stars using spectral types and B–V colors as mass indicators.

In addition, I derive the peak frequencies for 16 Cyg A and B from the Kepler time series, obtaining $\nu_{\rm max} = 2215.5 \pm 2.5\,\mu$Hz for 16 Cyg A and $2585.0 \pm 2.8\,\mu$Hz for 16 Cyg B, following the methodology of Kallinger et al. (2014). Using effective temperatures of $T_{\rm eff} = 5816 \pm 20$K for 16 Cyg A and 5756 ± 20K 16 Cyg B (average values from the sources listed by Ryabchikova et al. 2022), the nonlinear scaling relations from Kallinger et al. (2018) yield the masses and radii listed in Table 7. These values align well with the non-seismic estimates.

In the realm of seismic studies, Bellinger et al. (2017) demonstrated that even the most refined models fail to fully reproduce the observed frequencies. Their analysis revealed a significant systematic deviation between the surface-corrected model frequencies and the observations. Specifically, the model frequencies tend to overestimate the observed frequencies at low and high radial orders while underestimating them at intermediate radial orders. This systematic discrepancy suggests a structural mismatch between the models and the actual stars. To quantify this difference, Bellinger et al. (2017) employed a model-independent structure inversion technique.

For the `BaseFM` analysis, I use the $l = 0$ to 2 frequencies from Davies et al. (2015) to identify the best-fit models within grids G1, G2, and G3 (from Table 1). The resulting best-fit parameters yield masses and radii ranging from 1.05 to $0.92 M_\odot$ and 1.22 to $1.16 R_\odot$ for 16 Cyg A and 1.01 to $0.87 M_\odot$ and 1.11 to $1.05 R_\odot$ for 16 Cyg B. More critically, the metal-poor grid G1 is preferred for both stars, with probability contrasts of $p_{G1} : p_{G2} : p_{G3} = 0.96 : 0.04 :\sim 0.001$ for 16 Cyg A and $0.71 : 0.23 : 0.06$ for 16 Cyg B. However, this contradicts the observed chemical compositions of the stars. According to Ryabchikova et al. (2022), the metallicities are $[M/H] = 0.11 \pm 0.04$ for star A and 0.07 ± 0.03 for star B, making grid G2 a more plausible match. Unfortunately, G2 is based on a helium abundance of Y = 0.28, which is inconsistent with the Y ≈ 0.25 determined by Verma et al. (2015) for both stars.

Notably, `BaseFM` does select the "best" helium abundance among the available options (0.26, 0.28, and 0.31), but the overall chemical composition of the best-fit models results in masses and radii that are significantly smaller than the independent measurements. To explore whether a more realistic chemical composition would yield better results, I constructed a new grid with a primordial composition of (Y,Z)=(0.25,0.024) and masses ranging from 0.9 to $1.2 M_\odot$ from the ZAMS to $1.5 R_\odot$. All other modelling details are as described in Section 2.2.

Searching within this new grid (G4) produces improved fits, reproducing the non-seismic masses, radii, and luminosities to better than $1.3 \pm 2.7\%$, $1.6 \pm 1.2\%$, and $6 \pm 4\%$, respectively (see Table 7). For 16 Cyg A, the probability contrast is also moderately in favour of G4, with $p_{G1} : p_{G4} = 0.3 : 0.7$. For 16 Cyg B, the probability contrast alone does not allow for a clear conclusion.

Since the new seismic masses are significantly lower than previously accepted, it is unsurprising that 16 Cyg appears significantly older than previous estimates. However, it is reassuring that `BaseFM` predicts ages that differ only by $6 \pm 5\%$ (or 1.2σ). The identified best-fit models are not necessarily the definitive models for 16 Cyg A and B, as the quality of the fit is inherently limited by the representativeness of the model grid. However, the goal of this analysis is not to conduct an in-depth seismic study of 16 Cyg but rather to showcase the capabilities of `BaseFM`.

When multiple mode sequences are used, `BaseFM` performs a simultaneous curvature fit (Eq.,15) to all selected modes, employing a single curvature parameter α but allowing for individual $\Delta\nu_{\rm cor}$. The observed modes, their corresponding curvature fits, and the curvature-corrected modes are displayed in panels (a) and (b) of Figure 10. To confirm that the residual signal, after correcting for curvature, arises from the He II ionization glitch, I fit a damped harmonic function:

$$\delta\nu(\nu) = A e^{-(\nu_{l,n})/\eta} \sin\left(4\pi\tau_g \nu_{l,n} + \phi\right), \tag{20}$$

where the fit yields acoustic depths for the glitch (measured from the surface) of $\tau_g = 1005 \pm 16$s for 16 Cyg A and 864 ± 18s for 16 Cyg B (see Table 8). These results are consistent with the acoustic depths of He II ionization zones observed in other main-sequence stars (e.g., Mazumdar et al. 2014).

In summary, four independent measurements of the masses of 16 Cyg A and B – two seismic and two non-seismic – show agreement at the 1–2% level. This consistency strongly indicates that previous seismic analyses have overestimated the masses of both stars by approximately 3–5% (or 2–3σ), with significant implications for subsequent in-depth follow-up studies.

3.6 Systematic uncertainties

To investigate potential systematic uncertainties in different forward modelling methods, I repeat the seismic analyses

[8] applying a $\log g_\odot = 4.4381$ (Mamajek et al. 2015)

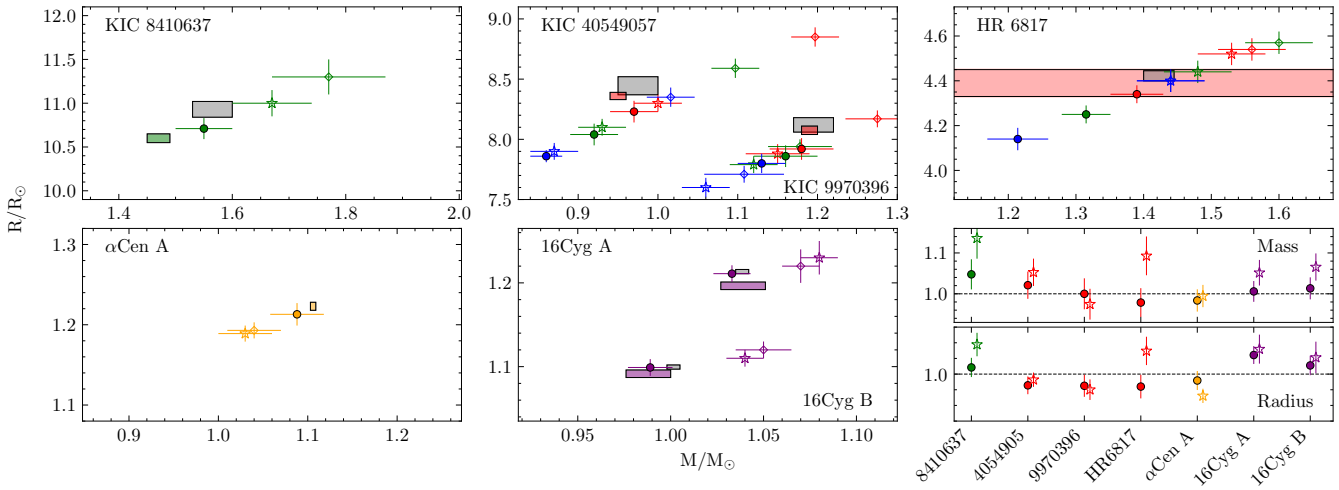

Figure 11. Seismic and non-seismic masses and radii for the seven tested stars. Color-filled squares represent the 1σ uncertainty range of the non-seismic parameters, while grey-filled squares indicate the error range for M and R derived from nonlinear scalings. Color-filled dots denote the best-fit `BaseFM` parameters, with red, green, and blue corresponding to fits to the grid G1, G2, and G3, respectively. Orange and purple indicate grids used to fit α Cen A and the 16 Cyg stars. Star and diamond symbols indicate the best-fit parameters based on the K08 and B14 surface corrections, respectively. The two panels in the bottom right corner show seismic masses and radii relative to the non-seismic values.

of the seven test stars using the classical surface correction approaches. The results are shown in Figure 11, which compares non-seismic masses and radii with seismic values derived from `BaseFM` and the K08 and B14 surface corrections.

Notably, `BaseFM` provides the closest match to the non-seismic parameters in all cases. Specifically, `BaseFM` reproduces the non-seismic masses and radii with ratios of 1.007 ± 0.012 and 0.999 ± 0.005, respectively (see bottom left panel of Figure 11). In other words, it overestimates the mass by $0.7 \pm 1.2\%$ and underestimates the radius by $0.1 \pm 0.5\%$, both well within the average uncertainties of approximately 3% for mass and 1% for radius.

By contrast, the K08 surface correction shows a more significant deviation, overestimating the mass and radius on average by $5.3 \pm 1.4\%$ (i.e., $> 3\sigma$) and $0.9 \pm 0.6\%$, respectively, which implies underestimating the stellar age by more than 20%.

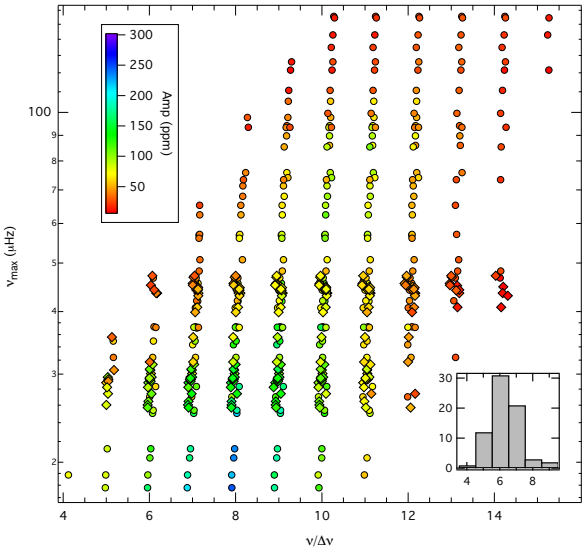

Figure 12. Radial mode frequencies for 45 RGB (dots) and 25 RC (diamonds) stars in NGC 6791 and NGC 6819 with at least four modes with $p > 0.95$ and $\sigma_\nu < 70$nHz, with colour-coded mode amplitdes. The insert shows a histogram for the number of significant modes per star.

4 FORWARD MODELLING OF CLUSTER GIANTS

An ideal test bench to demonstrate the capabilities of `BaseFM` are the red giants observed by *Kepler* in the two open clusters NGC 6791 and NGC 6819 (e.g., Stello et al. 2011a,b). In a cluster, main-sequence stars, which share the same age, distance, and primordial chemical composition, exhibit a tight relation between their initial mass and evolutionary stage. However, stars on the red giant branch (RGB) evolve much faster than those on the main sequence. As a result, very luminous RGB stars are only slightly more massive (typically less than 1%) than low-luminosity RGB stars. Consequently, one can safely assume that all members of a cluster's RGB approximately share a specific characteristic mass ($M_{\rm RGB}$), which is a key property of the cluster. More importantly, this provides an opportunity for a fundamental test: independent seismic modeling of individual RGB stars within a given cluster should yield consistent values for mass and age across all stars.

Several studies have attempted to determine $M_{\rm RGB}$ for the two clusters (e.g., Basu et al. 2011; Miglio et al. 2012; Wu et al. 2014b,a). These studies primarily rely on seismic scaling relations for the global oscillation parameters $\nu_{\rm max}$ and $\Delta\nu$ (e.g., Kjeldsen & Bedding 1995; Kallinger et al. 2010a), yielding masses of approximately $1.23 M_\odot$ for NGC 6791 and $1.65 M_\odot$ for NGC 6819. However, it has since been estab-

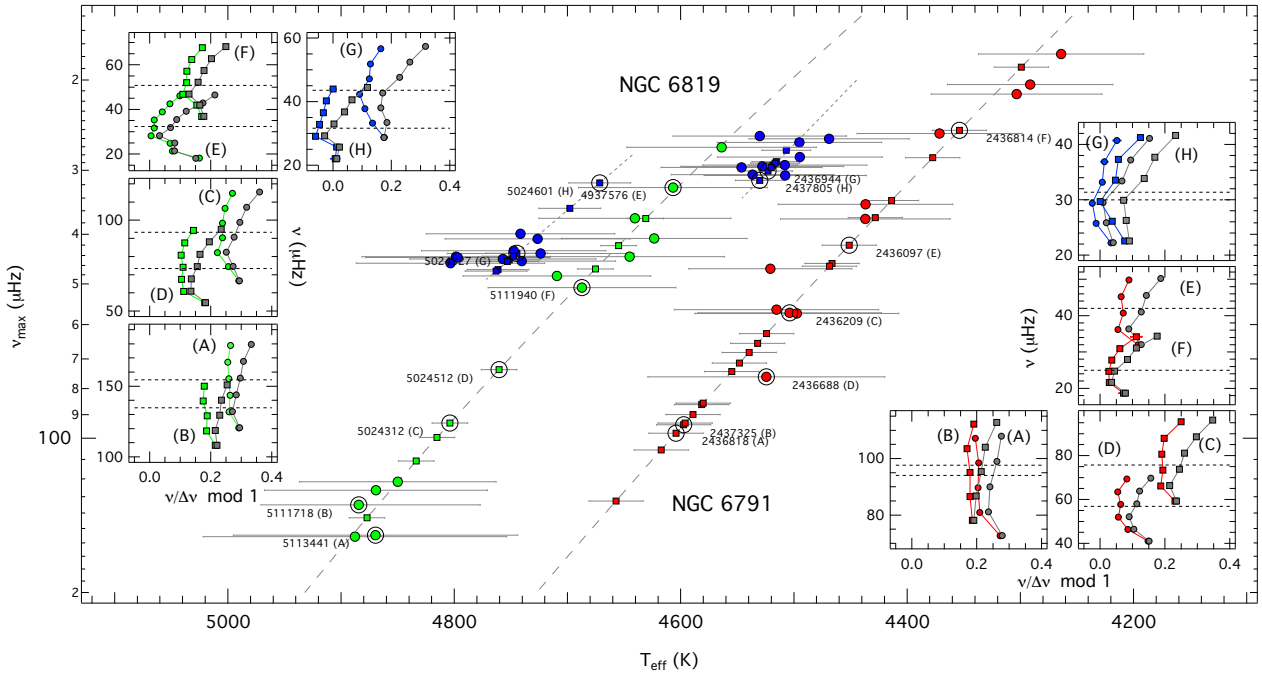

Figure 13. Peak frequency vs. effective temperature for the analysed sample of RGB (red and green symbols) and RC (blue symbols) stars in NGC6791 and NGC6819. Circles and squares indicate stars with the effective temperature taken from the APOKASC catalogue and derived from polynomial fits (grey dashed and dotted lines), respectively. Horizontal grey lines represent the uncertainties in $T_{\rm eff}$. The inserts examplarily show echelle diagrams for 16 stars from both clusters (marked by their KIC number). The coloured and grey symbols indicate the observed and corresponding best-fit model frequencies, respectively. Horizontal dashed lines give the peak frequency.

lished that classical seismic scaling relations tend to significantly overestimate the mass and radius of stars on the RGB (e.g., Gaulme et al. 2016; Themeßl et al. 2018). To address this, Kallinger et al. (2018) reformulated the classical seismic scaling relations in a model-independent way, deriving $M_{\rm RGB} = 1.10 \pm 0.03 M_\odot$ for NGC 6791 and $1.45 \pm 0.06 M_\odot$ for NGC 6819.

Since `BaseFM` operates independently of the seismic scaling relations, confirming these findings would not only validate the revised (non-linear) scaling relations but also provide compelling evidence for the reliability of Bayesian forward modeling.

Details about the *Kepler* data for the cluster stars are provided by K18. Radial mode frequencies are extracted following the procedure described in Sec. A and are shown in Figure 12. The sample includes 45 RGB stars and 25 RC stars, each with at least four significant radial modes, covering $\nu_{\rm max}$ values from approximately 17 to 155 μHz.

The positions of the stars in the "seismic" HR diagram are illustrated in Figure 13, where $\nu_{\rm max}$ serves as a proxy for stellar luminosity. Since not all stars in the sample are listed in the APOKASC catalog (Pinsonneault et al. 2014), missing effective temperatures are interpolated using polynomial fits in the $\nu_{\rm max}$– $T_{\rm eff}$ plane (see K18 for further details).

4.1 Independent forward modelling

The stars in NGC 6819 have a metallicity of [M/H] $= 0.05 \pm 0.04$ (Pinsonneault et al. 2014), which is practically identical to that of KIC 8410637. Consequently, the model grid constructed in Sec. 2 can also be used for these stars. On the other hand, stars in NGC 6791 exhibit a much higher metallicity of [M/H] $= 0.38 \pm 0.054$, requiring the use of metal-rich models for a proper seismic analysis. To address this, I compute a metal-rich grid consisting of approximately 108,800 models with an initial chemical composition of (Y, Z) = (0.31, 0.035) and $\alpha_{\rm MLT} = 2.1$. Note that $\alpha_{\rm MLT}$ is slightly smaller than that used for the original grid. While this adjustment has no significant impact on the analysis, it better reproduces the stars' positions in the HR diagram. All other parameters are set as described in Sec. 2, and the basic grid parameters are listed in Table 1.

In the first step, `BaseFM` uses 4–9 significant radial mode frequencies (see Figure 12) to identify a representative model for each star, treating each star independently. The best-fit masses and radii are illustrated in Figure 14, while the observed and corresponding best-fit model frequencies are shown for eight representative stars from each cluster in Figure 13. The average masses and ages (with corresponding uncertainties) for the 25 RGB and 13 RC stars in NGC 6791 and the 20 RGB and 12 RC stars in NGC 6819 are summarized in Table 9. These results are in good agreement with those derived from the non-linear seismic scaling relations.

The scatter in the best-fit masses, at around 4%, is relatively low and compares well with the average uncertainty of individual measurements, approximately 3%. This suggests no significant intrinsic dispersion beyond this level of accuracy. In contrast, the average ages are less well constrained, with 7.0 ± 1.2 Gyr for NGC 6791 and 2.6 ± 0.3 Gyr for NGC 6819. For example, best-fit ages for stars in NGC 6791

Table 9. Average mass (in solar units) and ages (in Gyr) of RGB and RC stars as given in the literature and determined in the present analysis. The finally adopted values are in bold face.

	NGC 6791				NGC 6819			
	$M_{\rm RGB}$	$M_{\rm RC}$	$t_{\rm RGB}$	$t_{\rm RC}$	$M_{\rm RGB}$	$M_{\rm RC}$	$t_{\rm RGB}$	$t_{\rm RC}$
eclipsing binaries	$1.15 \pm 0.02^{\dagger}$	-	$8.3 \pm 0.3^{\dagger}$		$1.54 \pm 0.02^{\dagger\dagger}$	-	$2.62 \pm 0.25^{\dagger\dagger}$	
nonlinear scaling†††	1.10 ± 0.03	1.02 ± 0.05	10.1 ± 0.9		1.45 ± 0.06	1.54 ± 0.07	2.9 ± 0.3	
independent fit	1.08 ± 0.04	1.05 ± 0.03	7.0 ± 1.2	6.1 ± 0.5	1.51 ± 0.06	1.56 ± 0.03	2.6 ± 0.3	2.2 ± 0.2
fit with age prior	$\mathbf{1.14 \pm 0.02}$	1.07 ± 0.02	$\mathbf{7.4 \pm 0.4}$	7.4 ± 0.3	$\mathbf{1.51 \pm 0.02}$	1.51 ± 0.02	$\mathbf{2.6 \pm 0.1}$	2.5 ± 0.1

†Brogaard et al. (2021) – †† Sandquist et al. (2013) – †††Kallinger et al. (2018)

range from 5.0 to 8.6 Gyr, with correspondingly large uncertainties. Such a wide spread is unlikely to be real and instead reflects the challenges of accurately placing the best-fit models in parameter space when only a limited number of oscillation modes are available. To address this issue, another advantage of the Bayesian framework outlined in Sec. 2.4 can be exploited, as described in Sec. 4.2.

An additional interesting outcome of this analysis is its implications for mass loss on the giant branch. For NGC 6819, no significant mass loss is observed between RGB and RC stars. In contrast, stars in NGC 6791 that have ignited core-helium burning should have lost about 7% of their initial mass (Miglio et al. 2012; Kallinger et al. 2018). While `BaseFM` appears to partially reproduce this for NGC 6791 (i.e., $M_{\rm RC}$ is smaller than $M_{\rm RGB}$), the ages of RC stars are systematically underestimated. As a result, it is challenging to directly compare the masses of best-fit models for stars on the RGB and RC, as they must have the same age for a meaningful comparison.

4.2 Forward modelling with an age prior

Originally, all models are treated equally by using a uniform prior in Eq. 7. However, given that all cluster members are expected to share a common age, an age prior is applied instead, following Eq. 19, to encode this information. The prior includes a target age t and its uncertainty σ_t. For the reasons discussed earlier, $t_{\rm RC}$ is excluded from the initial modeling. Instead, slightly relaxed age priors of 7 ± 2 and 2.6 ± 0.5 Gyr are adopted for NGC 6791 and NGC 6819, respectively, and the analysis is repeated. The resulting best-fit masses and radii are illustrated in Figure 14, while the average masses and ages are listed in Table 9.

Although the adopted priors allow for considerable dispersion, they significantly reduce the scatter in the best-fit masses and ages – by a factor of two to three. The age prior aligns the RC stars with the same age range as the RGB stars, enabling an exploration of the mass difference between these two populations.

For NGC 6791, the difference between $M_{\rm RGB}$ and $M_{\rm RC}$ is $0.07 \pm 0.03 M_{\odot}$, equivalent to $6 \pm 2\%$. This result is fully consistent with the mass loss reported in previous studies. Conversely, no mass loss is detected in NGC 6819, which aligns with theoretical expectations.

The masses determined here are consistent within uncertainties with those reported by K18, despite the use of entirely independent methods and different evolutionary models. They also compare well with measurements of several eclipsing binaries near or on the clusters' main sequences

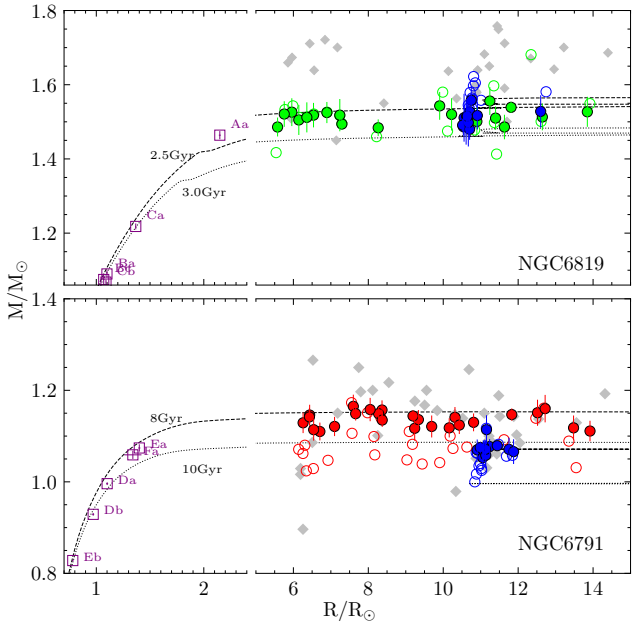

Figure 14. Seismic mass versus radius from Bayesian model fitting. Open and filled circles indicate results from independent model fitting and those determined with a common age prior, respectively. Uncertainies are only shown for the latter for better visibility. Grey diamond symbols show M and R from fits with the B14 surface correction. Red (top) and green (bottom) symbols indicated RGB stars and blue symbols give RC stars. Also plotted are BaSTI isochrones for solar metallicity in the top panel and [Fe/H] = 0.26 in the bottom panel with the numbers indicating the age of the isochrones as well as a number of eclipsing binaries (purple squares), with the labels indicating: Aa → WOCS23009 A; Ba, Bb → WOCS24009 A & B; Ca, Cb → WOCS40007 A & B; Da, Db → V18 A & B; Ea, Eb → V20 A & B; Fa → V80 A.

(e.g., Grundahl et al. 2008; Brogaard et al. 2012; Jeffries et al. 2013; Brewer et al. 2016). For a detailed discussion of this comparison, I refer to K18.

In summary, the lower $M_{\rm RGB}$ values derived from seismic forward modeling (and the non-linear scaling relations) are in better agreement with binary measurements than the previously adopted higher $M_{\rm RGB}$ values of approximately $1.23 M_{\odot}$ for NGC 6791 and $1.65 M_{\odot}$ for NGC 6819 (e.g., Basu et al. 2011; Miglio et al. 2012; Wu et al. 2014b; Handberg et al. 2017; Rodrigues et al. 2017). This is qualitatively illustrated in Figure 14, which shows that the binaries follow

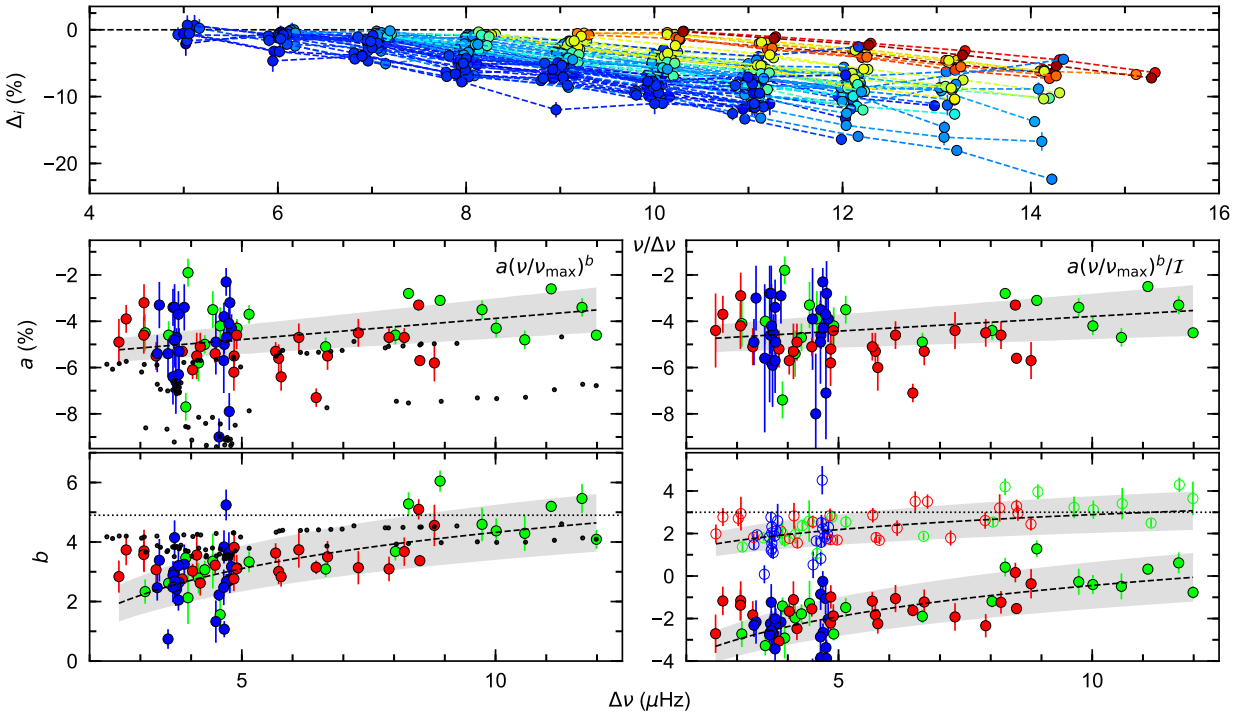

Figure 15. The top panel gives the measured surface effect (in the sense of observed minus model frequencies as a fraction of $\Delta\nu$) for the sample of cluster stars. Symbol colours indicate $\Delta\nu$ with red and blue corresponding to the largest and smallest values in the sample, respectively. The middle and bottom panels show the scaling factor (a in per cent of $\Delta\nu$) and exponent (b) of fits with Eq. 5 (left) and 6 (right), respectively, to the measured surface effects. Symbol colours are as in Fig. 13. The open symbols in the bottom right panel give exponents determined from model fitting with the classical model grid (Sec. 2.7). Black dots indicate predicted values from Sonoi et al. (2015). Dashed lines correspond to linear (middle) and log-linear (bottom) fits to the scaling factors and exponents, respectively, with the grey-shaded areas indicating 1σ uncertainties. Dotted lines in the bottom panels give the originaly adopted exponents for K08 and B14 corrections.

the same range of BaSTI[9] isochrones as the red giants. The only exception is WOCS23009, which is an outlier. However, its mass was derived from isochrone fitting rather than from binary orbit measurements and is therefore questionable.

Quantitatively, Brogaard et al. (2021) reported that the isochrone fitting the binaries in NGC 6791 implies $M_{\rm RGB}$ should be around $1.15 M_\odot$, which matches the average mass of the RGB stars in this cluster determined here. Similarly, the detailed seismic analysis of McKeever et al. (2019) also yields $M_{\rm RGB} \approx 1.15 M_\odot$. A similar agreement is found for NGC 6819, where Sandquist et al. (2013) reported that the isochrone fitting an eclipsing binary at the cluster turnoff corresponds to a giant branch mass of $1.52 - 1.56 M_\odot$, which is consistent with the present findings within one sigma.

Determining the cluster ages is more challenging. While the average seismic age of NGC 6819, 2.6 ± 0.1 Gyr, is in excellent agreement with independent measurements (e.g., Sandquist et al. 2013), `BaseFM` appears to underestimate the age of NGC 6791 compared to other studies (e.g., Brogaard et al. 2021). However, age estimates for NGC 6791 in the literature vary widely, ranging from approximately 6.8 Gyr (Basu et al. 2011) to over 10 Gyr (K18). Such a large spread is unlikely to be real and instead reflects the difficulty of determining stellar ages, a persistent challenge in stellar astrophysics that will not be resolved here. Absolute age estimates depend heavily on the adopted models, the implemented physics, and input parameters such as chemical composition, mixing length, overshooting, and mass loss. For example, neglecting mass loss on the giant branch alone can increase the inferred cluster age by about 2 Gyr. As a result, any age determination for NGC 6791 with better than $\sim 10\%$ precision should be treated cautiously, as it is likely to reflect the internal accuracy of the method rather than the true cluster age.

Finally, I determine M and R using the B14 surface corrections within the Bayesian framework and present the results in Figure 14. While the scatter is relatively large, the average RGB mass derived for NGC 6791, $1.14 \pm 0.08 M_\odot$, is fairly consistent with the value obtained from the surface-effect-independent approach. The situation is different for NGC 6819, where the individual best-fit masses and radii are systematically larger than those obtained with `BaseFM`. The B14 correction yields $M_{\rm RGB} = 1.61 \pm 0.09 M_\odot$, consistent with the findings of Ball et al. (2018), who reported that commonly adopted surface corrections lead to systematically overestimated masses and radii for red giants.

This discrepancy likely arises because the classical surface corrections are unable to accurately reproduce the true surface effect in red-giant models, thereby introducing sys-

[9] BaSTI isochrones

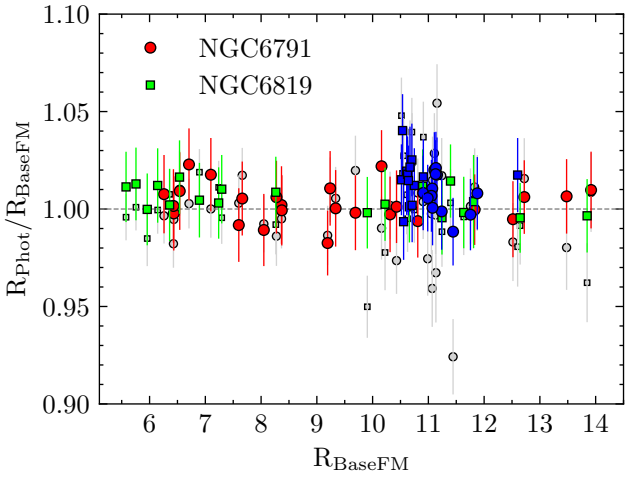

Figure 16. Photometric versus seismic radii from `BaseFM` for RGB (red/green) and RC (blue) stars in NGC 6791 (dots) and NGC 6819 (squares). Grey symbols correspond to seismic radii from the nonlinear scalings of K18.

tematic errors in the model matching process. This issue is further explored in the next section.

4.3 Systematic uncertainties

The systematic uncertainties in the mass of cluster stars are with less than $2 \pm 2\%$ practically negligible, as illustrated in Figure 14 and Table 9. A reliable method to also validate the radii obtained from `BaseFM` is to compare them with those derived from the star's distance (d) and apparent bolometric magnitude ($m_{\rm bol}$) using:

$$R_{\rm phot} = 8.935 \times 10^{(2-0.2m_{\rm bol})} d\, T_{\rm eff}^{-2}, \qquad (21)$$

where $T_{\rm eff}$ is expressed in solar units and d in kiloparsecs. The apparent bolometric magnitudes are sourced from K18 and are based on 2MASS JHK photometry (Skrutskie et al. 2006), with bolometric corrections applied via interpolation using the tables from Bonatto et al. (2004). K08 also determined the cluster distances to be 4.19 ± 0.06 kpc for NGC 6791 and 2.41 ± 0.03 kpc for NGC 6819.

Figure 16 compares the resulting photometric radii with those from `BaseFM`, revealing no significant discrepancies. Specifically, the average ratio between photometric and seismic radii for RGB stars in both clusters is 1.00 ± 0.01, indicating equivalence within 1%. For RC stars in NGC 6791, the ratio is 1.01 ± 0.01, which is also nearly perfect. However, the seismic radii of RC stars in NGC 6819 are on average $2 \pm 1\%$ smaller than the corresponding photometric radii. Although this difference is not statistically significant, it likely results from RC stars in NGC 6819 having undergone minimal mass loss, contrary to the assumptions of the stellar models used which assume some mass loss.

In summary, the cluster stars' masses and radii determined from the `BaseFM` fit, incorporating an age prior, are consistent with independent values, showing systematic differences smaller than 1% for radii and 2% for mass.

4.4 The surface effect on the giant branch

Assuming that the best-fit models identified above are representative of the analyzed stars, their actual surface effect corresponds to the difference between the observed and best-fit model frequencies. These differences are shown in Figure 15 and exhibit a wide range of characteristics. For stars near the base of the giant branch (i.e., with large $\Delta\nu$), the surface effect appears relatively shallow, contributing less than 10% of the large frequency separation. In contrast, for more evolved stars, the frequency shifts exceed 20% of $\Delta\nu$.

To quantify this, I fit Eq. 5 (K08 correction) and Eq. 6 (B14 correction with a free exponent) to the frequency differences and plot the corresponding scaling factors (a) and exponents (b) in Figure 15. One might expect that the varying surface effect along the giant branch is driven by the scaling factor a. However, no clear trend with $\Delta\nu$ is observed, although the scatter is substantial. Linear fits yield small gradients of 0.18 ± 0.05 and $0.15 \pm 0.05\%/\,\mu$Hz for the K08 power law and the inertia-scaled B14 power law, respectively. Furthermore, no significant dependence on the mass or metallicity of the stars is evident.

The exponent, however, shows a clear decrease with decreasing $\Delta\nu$. While the K08 power-law exponents are consistent with the canonical value of 4.9 for stars near the lower giant branch, they decrease to approximately 3 for more evolved stars. Again, no significant dependence on the evolutionary stage, mass, or metallicity is apparent. To quantify this trend, I fit a log-linear function to the measured exponents, obtaining $b = (4.0 \pm 0.5) \log \Delta\nu + (0.3 \pm 0.5)$.

A similar trend is found for the inertia-weighted frequencies (i.e., the B14 correction). However, the absolute values of the fitted exponents differ substantially. As discussed in Sec. 2.7, the peculiar behavior of the model frequencies' inertia forces the exponents to deviate significantly from the expected value of three, with even negative values for stars with $\Delta\nu$ smaller than approximately $8\,\mu$Hz. A fit reveals $b = (4.9 \pm 0.6) \log \Delta\nu - (5.3 \pm 0.4)$.

Using a standard model grid brings the exponents closer to three but retains the same trend, with $b = (2.3 \pm 0.5) \log \Delta\nu + (0.5 \pm 0.4)$. This indicates that the B14 correction, like the K08 correction, is unable to accurately predict or correct the surface effect in red-giant models. Using either correction in forward modeling will inevitably introduce unknown systematic errors.

Sonoi et al. (2015) argue that empirical surface corrections should not be calibrated solely on the Sun but should instead be constrained using realistic physical modeling. To this end, they employed 3D hydrodynamical simulations patched onto 1D stellar models to compute realistic model frequencies (supposedly free of the surface effect) and compared these to the eigenfrequencies of the original 1D models. Through this approach, they characterized the surface effect for various positions in the HR diagram and provided scaling relations for the parameters of several correction functions, such as the K08 correction.

However, their model grid focuses primarily on the main sequence and includes only one model on the giant branch. Consequently, their scalings are not expected to provide reliable predictions for red giants. This is demonstrated in Figure 15, where the predicted parameters of the K08 correction are compared with the actual measurements, showing only

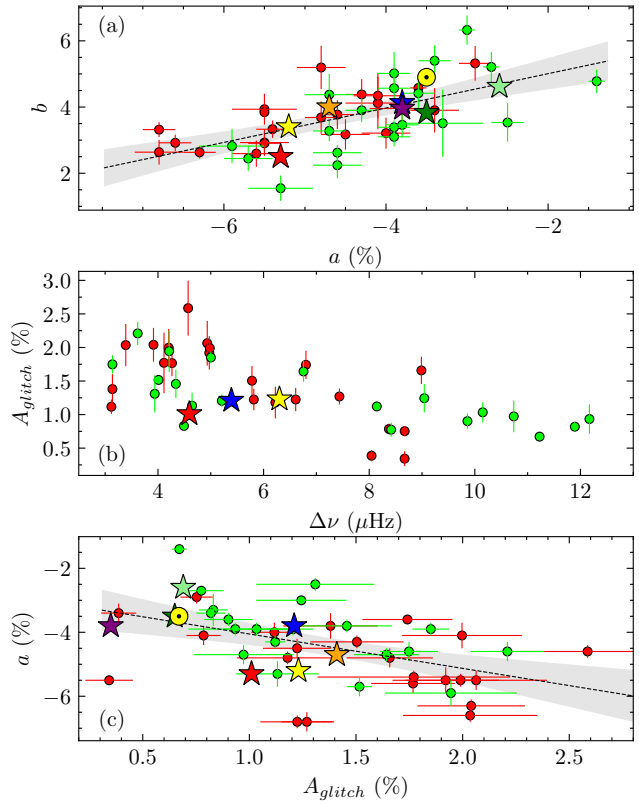

Figure 17. Connection between the acoustic glitch and the surface effect. Symbol colors are the same as in Fig. 13. Panel (a) shows the correlation between the exponent and scaling factor (in per cent of $\Delta\nu$) of the the power law fitted to difference between observed and best-fit model frequencies for RGB stars in the two clusters. Panel (b) gives the amplitude to the acoustic glitch signal (in per cent of $\Delta\nu$) as a function of the large separation. Panel (c) describes the correlation between the scale of the surface effect and amplitude of the glitch signal. Dashed lines indicate linear fits with the grey-shaded area corresponding to the 90% confidence interval. Large symbols mark the Sun (a and b taken from K08), α CenA (orange star), 16 Cyg A&B (green and lightgreen), KIC 8410637 (red star), KIC9970396 (yellow star), KIC 4054905 (blue star), HR 6817 (purple star), which are not used for the fits.

marginal agreement. Notably, the scalings produce significantly different scaling factors for stars in different clusters, suggesting a strong dependence on metallicity or mass. However, their scalings are a function of $T_{\rm eff}$, which differs by more than 100 K for stars with similar surface gravity (the other independent variable in their scaling) but belonging to different clusters. Such an effect is not evident in the actual measurements.

A more conclusive parameterization of the surface effect properties is presented in Figure 17. It shows that the power-law scaling factors and exponents are correlated, with the exponent decreasing (i.e., the shape becoming shallower) as the scaling factor increases (i.e., for a "deeper" surface effect). A fit reveals $b = (6.0 \pm 0.2) + a \cdot (0.5 \pm 0.05)$, where a is expressed as a percentage of $\Delta\nu$. This correlation is also consistent with values derived for the Sun (using a and b from K08) and the stars from Sec. 3.

I then test various parameters ($\nu_{\rm max}$, $\Delta\nu$, R, M, $T_{\rm eff}$, etc.) for correlations with either a or b and find the strongest correlation between a and the amplitude of the acoustic glitch signal, $A_{\rm glitch}$. As mentioned earlier, the amplitude of the periodic deviations from a strictly regular mode pattern is a measure of the width of the region in the stellar interior responsible for the acoustic glitch (in the case of red giants, the second helium ionisation zone).

Normally, fitting a damped harmonic function to the frequency deviations is used to determine the properties of the glitch signal. However, given the limited number of observed modes per star, this approach becomes problematic. An alternative method to estimate the glitch amplitude leverages the statistical property $\lim_{n\to\infty} {\rm Var}(\sin x) = 0.5$ (where n is the number of elements in x). Using this, $A_{\rm glitch}$ can be approximated as $\sqrt{2}$ times the standard deviation of ν' from Eq. 15, representing the average glitch amplitude over the observed frequency range.

The resulting amplitudes are shown in Figure 17 as a function of $\Delta\nu$. These results suggest that the width of the second helium ionisation zone increases during stellar evolution, consistent with the findings of Vrard et al. (2015). More importantly, $A_{\rm glitch}$ correlates with the scale of the surface effect. A linear fit yields:

$$a = (-3.0 \pm 0.2) - A_{\rm glitch} \cdot (1.1 \pm 0.2). \quad (22)$$

Notably, members of the metal-rich (and presumably helium-rich) cluster NGC 6791 tend to group at larger glitch amplitudes and deeper surface effects, while members of the roughly solar-metallicity cluster NGC 6819 are located toward smaller glitch amplitudes and shallower surface effects in the models. This suggests that both the observed glitch signal and the surface effect in representative models are influenced by the chemical composition.

5 THE SURFACE EFFECT OF DIPOLE MODES

Unlike less-evolved stars, seismic model fitting of RGB stars is typically restricted to radial modes, as the spectrum of non-radial mixed modes is too complex to handle effectively. However, various applications, such as the analysis of internal rotation profiles (e.g., Beck et al. 2012, 2014), require a comparison between observed and theoretical mixed dipole modes, necessitating precise mode identification – that is, determining which theoretical mode corresponds to which observed mode.

This task becomes increasingly challenging as stars evolve up the giant branch. The mixed-mode spectra grow denser, and conventional surface corrections can potentially exceed the g-mode spacing. To address this, Ball et al. (2018) proposed modifying the first adiabatic index of the stellar model to suppress gravity modes in the stellar core. This adjustment forces the model to oscillate in pure pressure modes, which can then be corrected using standard surface corrections. A less invasive approach is presented in the following discussion.

None of the red giants analyzed so far are suitable for studying the surface effect on dipole modes, as their mixed-mode spectra are either too dense (e.g., KIC 8410637) or lack significant dipole modes. A much better target for this purpose is the RGB star KIC 10513837. The four-year *Kepler* observations reveal an almost complete spectrum of $l = 0$

and 1 modes spanning over five radial orders, with clearly resolved dipole modes that show no significant rotational splittings (which could complicate mode identification).

KIC 10513837 is part of the APOKASC sample and has a metallicity of approximately [M/H] = 0.03. Therefore, the model grid G2 from Table 1 can be used for the seismic analysis. The radial and dipole modes are extracted as described in Appendix A and are shown in Figure 18. A `BaseFM` fit to the six central radial modes yields a mass and radius of $1.31 \pm 0.01 M_\odot$ and $4.74 \pm 0.02 R_\odot$, respectively.

The best-fit model has a $\Delta\Pi_1$ value of 82.5 ± 0.5 s (calculated using Eq. 3), which agrees very well with the value of 82.3 ± 1.1 s determined from the observed modes. The latter is derived by fitting the observed dipole mode period spacings (ΔP) using:

$$\Delta P = \Delta\Pi_1 \zeta(\nu), \tag{23}$$

where $\zeta(\nu)$ represents the ratio between the kinetic energy of the mode in the g-mode cavity and the total kinetic energy. This parameter describes whether a mode is predominantly trapped in the g-mode cavity (g-dominated modes with ζ close to 1) or in the p-mode cavity (p-dominated modes with $\zeta \ll 1$) and can be expressed as (e.g., Mosser et al. 2015; Hekker & Christensen-Dalsgaard 2017):

$$\zeta(\nu) = \left[1 + \frac{\nu^2}{\Delta\nu} \frac{q\,\Delta\Pi_1}{q^2 + (1-q^2)\sin^2\left(\pi \frac{\nu - \nu_{n_p}}{\Delta\nu} + \epsilon_g\right)} \right]^{-1}. \tag{24}$$

Here, the local minima are defined by ν_{n_p}, the frequency of the pure p-mode of radial order n_p (approximately the midpoint between two consecutive radial modes). The parameter q describes the coupling between the p- and g-mode components, and ϵ_g is a phase term.

For KIC 10513837, the fit yields a coupling factor of $q = 0.17 \pm 0.02$ and a phase term of $\epsilon_g = 0.07 \pm 0.01$, which are typical for stars at this evolutionary stage (e.g., Mosser et al. 2017). The observed and best-fit model period spacings, along with the corresponding fits to Eq. 23, are shown in Figure 18.

It is noteworthy that while the fits yield almost identical parameters, the surface effect causes ν_{n_p} to increase more rapidly in the model frequencies than in the observations.

Since nonradial modes are more conveniently analyzed in period space, I define the difference between the observed dipole modes and their corresponding theoretical counterparts as:

$$\Delta_{P_1} = \frac{1}{\nu_{\rm obs}} - \frac{1}{\nu_{\rm mod}}, \tag{25}$$

and plot the period differences of the 25 observed dipole modes in Figure 18. Intuitively, one might expect the most p-mode-like mixed modes (i.e., the p-dominated modes) to exhibit a trend similar to the surface effect of the radial modes, expressed in terms of period differences. Indeed, the period difference increases with frequency, with an amplitude comparable to that of the radial modes but in the opposite direction.

Interestingly, the g-dominated modes follow the radial mode trend more closely than the p-dominated modes. The period differences of the p-dominated modes appear to

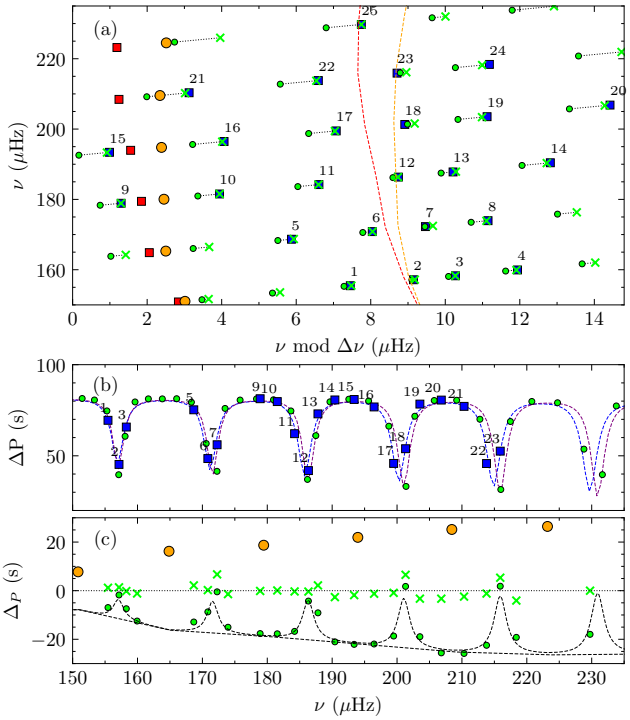

Figure 18. Panel (a) shows an echelle diagram of the observed radial (red squares) and dipole (numbered blue squares) modes and the corresponding best-fit model frequencies (orange and green dots) of KIC 10513837. The red and orange lines indicate the midpoints between adjecent radial modes connected by line segments. Green crosses indicate the surface-effect-corrected dipole modes of the model (connected to original modes by dotted-line segments). Panel (b) gives the corresponding period spacings along with fits to Eq. 23. Panel (c) shows the period difference between the observed and best-fit model dipole modes before (green dots) and after (green crosses) the surface correction with Eq. 26. Orange dots indicate the surface effect of the radial modes in period and dashed lines are the correction function (Eq. 26), with and without the ζ-function multiplied.

be modulated similarly to the mode's period spacings, described by the ζ-function. To account for this, I correct the period of a given dipole mode ν using:

$$\Delta_{P_1}(\nu) = -\Delta_{P_0}(\nu)\,\zeta(\nu), \tag{26}$$

where Δ_{P_0} is the period difference of the corresponding radial mode, which is determined in practice through linear interpolation between adjacent radial modes. For improved accuracy, the ζ-function is rescaled to a range [0,1] for the considered frequency range.

The corrected dipole modes are illustrated in Figure 18 and generally agree well with the observations. In this example, the average period difference is reduced from approximately -13.7 s to 0 ± 3 s.

Since the mode inertia is expected to follow the ζ-function, this surface correction for dipole modes resembles the B14 correction. The key difference is that the modes are corrected in period space rather than frequency space. Moreover, this correction can be seamlessly integrated into the Bayesian model fitting framework outlined in Sec. 2. How-

ever, further testing is required, and this will be addressed in a follow-up study.

6 CONCLUSIONS

This paper presents a novel Bayesian approach to asteroseismic forward modeling that is based on the work of Gruberbauer et al. (2012, 2013) but significantly improves the accuracy and reliability of stellar parameter estimations, for stars with solar-type oscillations. The method's primary strength lies in its ability to overcome the inherent uncertainties in stellar models, especially the near-surface effect, without resorting to often problematic empirical corrections. Traditional methods rely on latter corrections, which are calibrated to solar data and may not generalize well to other stellar types, especially evolved stars. The study demonstrates a superior methodology by directly incorporating the uncertainties into the Bayesian framework. The key advancements are as follow:

(i) The core innovation of `BaseFM` is the probabilistic treatment of the near-surface effect. Instead of applying a correction, the method introduces an additional parameter to represent the systematic frequency shift. This parameter is then marginalized over, effectively integrating out its influence while retaining its impact on the overall model probability. This avoids the assumptions inherent in empirical corrections and produces more robust and unbiased results. The impact is substantial, as it allows for reliable comparisons between observed and model frequencies without relying on approximations that may not be valid for all stars.

(ii) The inclusion of the glitch signal (periodic deviations from regular mode patterns due to sharp structural variations in the stellar interior) further refines the parameter estimations. This aspect enhances the discriminative power of the method by providing additional constraints on the model selection process. The improved results demonstrate the usefulness of incorporating additional seismic features to improve model fits.

(iii) The second major advantage of `BaseFM` is its adoption of nested sampling for model selection, parameter estimation, and uncertainty quantification. Unlike traditional grid-search methods – such as those used in AIMS (Rendle et al. 2019) or BASTA (Aguirre Børsen-Koch et al. 2022) – nested sampling focuses computational resources on high-likelihood regions of parameter space, while also providing properly weighted posterior samples and robust evidence estimates.

(iv) The efficacy of the method is rigorously tested and validated across diverse stellar types from Sun-like stars to evolved red giants (including 70 RGB and RC stars in NGC 6791 and NGC 6819). Consistent results across these stars strongly support the method's robustness and general applicability. The successful reproduction of independently measured parameters with negligible systematic uncertainties significantly enhances confidence in the methodology.

(v) The study provides a critical assessment of existing empirical surface corrections (K08 and B14). It finds that these corrections are often inadequate, highlighting a systematic overestimation of stellar masses and radii when these corrections are utilized. This analysis emphasizes the limitations of relying on empirically-derived corrections and motivates the development of model-independent techniques.

(vi) The research presents a comprehensive analysis of the surface effect, demonstrating its dependency on various stellar parameters (effective temperature, large separation, and possibly metallicity). This detailed exploration reveals a complex relationship between the surface effect and other stellar characteristics, adding to the broader understanding of stellar structure and oscillations. The correlation between surface effect magnitude and acoustic glitch amplitude is a significant finding that supports the physical understanding of the phenomenon.

(vii) Also presented is a method to correct for the surface effect on dipole modes by modifying the period of the dipole mode based on the period difference of the corresponding radial mode and a factor that accounts for the mode's mixed p- and g-mode character. This approach provides a more effective way to compare observed and model dipole mode frequencies in evolved stars.

The presented `BaseFM` approach represents a superior method for asteroseismic forward modeling that has the potential to significantly contribute to our understanding of stellar structure and evolution, especially in evolved stars. The refined parameter estimations will significantly improve the accuracy of stellar modeling, allowing to identify more representative models for stars with solar-type oscillations than this was previously possible. Consequently, `BaseFM` opens new avenues to more precise tests of various physical phenomena, from the puzzling angular momentum transport in evolved stars to stellar evolution theory in general.

This study's results are not intended as exhaustive analyses of the individual stars presented. Instead, these examples demonstrate and validate a method for asteroseismic fitting that circumvents the limitations of near-surface effect corrections. The realism of any best-fit model is inherently limited by the underlying stellar models and their physical assumptions. During the last ten to twenty years numerous published seismic analyses rely on classical surface correction. Frankly, one has to question them all, as these corrections are now demonstrably flawed. `BaseFM` offers a crucial first step to overcome this.

In a broader context, the findings presented here also have implications for the frequently used seismic scaling relations (e.g., Kallinger et al. 2010a). The scaling relations $\Delta\nu \propto \sqrt{\rho}$ and $\nu_{\rm max} \propto g/\sqrt{T_{\rm eff}}$ provide a convenient way to estimate a star's mass and radius from its global seismic parameters. However, these relations are not exact (e.g., Gaulme et al. 2016), leading to the introduction of correction terms. The $\Delta\nu$ correction is typically derived from the eigenfrequencies of stellar models, which are in the worst case uncorrected for the surface effect (e.g., White et al. 2011; Guggenberger et al. 2016; Sharma et al. 2016). Even when applying traditional surface corrections, this introduces unknown systematic uncertainties into the $\Delta\nu$ scaling[10]. While

[10] Note that even in the case of the Sun, the surface effect contributes to roughly half of the discrepancy between the observed $\Delta\nu$ and the value derived from theoretical eigenfrequencies. Introducing a model-dependent correction to the scaling relation thus artificially incorporates the surface effect into a relation that is, by nature, independent of it.

one might compensate for these effects statistically by empirically calibrating the $\nu_{\rm max}$ scaling (e.g., Pinsonneault et al. 2025), the validity of this approach remains questionable (see also Valle et al. 2025). The results presented here lay the groundwork for improving the model-independent nonlinear scalings proposed by Kallinger et al. (2018), but a detailed exploration of this will be deferred to a follow-up study.

ACKNOWLEDGEMENTS

The author thanks Joel Ong for providing a script to compute the large separation estimator already before publishing it and Werner W. Weiss for fruitful discussions. This paper includes data collected by the Kepler and TESS missions and obtained from the MAST data archive at the Space Telescope Science Institute (STScI). Funding for the Kepler mission is provided by the NASA Science Mission Directorate. Funding for the TESS mission is provided by the NASA Explorer Program. STScI is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–26555.

This research made use of `UltraNest` (Buchner 2021) and the Python libraries `numpy` (Harris et al. 2020), `pandas` (Wes McKinney 2010), `mpmath` (mpmath development team 2023), `SciencePlots` (Garrett 2021), and `scipy` (Virtanen et al. 2020).

REFERENCES

Abdurro'uf et al., 2022, ApJS, 259, 35
Aguirre Børsen-Koch V., et al., 2022, MNRAS, 509, 4344
Baglin A., Auvergne M., Barge P., Deleuil M., Catala C., Michel E., Weiss W., COROT Team 2006, in Fridlund M., Baglin A., Lochard J., Conroy L., eds, ESA Special Publication Vol. 1306, The CoRoT Mission Pre-Launch Status - Stellar Seismology and Planet Finding. p. 33
Ball W. H., Gizon L., 2014, A&A, 568, A123
Ball W. H., Gizon L., 2017, A&A, 600, A128
Ball W. H., Themeßl N., Hekker S., 2018, MNRAS, 478, 4697
Basu S., et al., 2011, ApJ, 729, L10
Bazot M., 2020, A&A, 635, A26
Bazot M., Bouchy F., Kjeldsen H., Charpinet S., Laymand M., Vauclair S., 2007, A&A, 470, 295
Beck P. G., et al., 2012, Nature, 481, 55
Beck P. G., et al., 2014, A&A, 564, A36
Bellinger E. P., Angelou G. C., Hekker S., Basu S., Ball W. H., Guggenberger E., 2016, ApJ, 830, 31
Bellinger E. P., Basu S., Hekker S., Ball W. H., 2017, ApJ, 851, 80
Benbakoura M., et al., 2021, A&A, 648, A113
Böhm-Vitense E., 1958, Z. Astrophys., 46, 108
Bonatto C., Bica E., Girardi L., 2004, A&A, 415, 571
Borucki W. J., et al., 2010, Science, 327, 977
Brewer L. N., et al., 2016, AJ, 151, 66
Brogaard K., et al., 2012, A&A, 543, A106
Brogaard K., et al., 2018, MNRAS, 476, 3729
Brogaard K., et al., 2021, A&A, 649, A178
Brogaard K., et al., 2022, A&A, 668, A82
Brown T. M., 1984, Science, 226, 687
Buchner J., 2021, The Journal of Open Source Software, 6, 3001
Buldgen G., et al., 2019, MNRAS, 482, 2305
Chaplin W. J., Miglio A., 2013, ARA&A, 51, 353
Chmielewski Y., Friel E., Cayrel de Strobel G., Bentolila C., 1992, A&A, 263, 219
Christensen-Dalsgaard J., 2008, Ap&SS, 316, 113
Christensen-Dalsgaard J., Dappen W., Lebreton Y., 1988, Nature, 336, 634
Christensen-Dalsgaard J., et al., 1996, Science, 272, 1286
Christensen-Dalsgaard J., et al., 2010, ApJ, 713, L164
Clara M., Cunha M. S., Avelino P. P., Campante T. L., Deheuvels S., Reese D. R., 2025, A&A, 694, A314
Corsaro E., De Ridder J., García R. A., 2015, A&A, 579, A83
Corsaro E., De Ridder J., García R. A., 2018, A&A, 612, C2
Davies G. R., et al., 2015, MNRAS, 446, 2959
Deheuvels S., et al., 2012, ApJ, 756, 19
Dziembowski W., 1977, Acta Astron., 27, 95
Eggenberger P., Charbonnel C., Talon S., Meynet G., Maeder A., Carrier F., Bourban G., 2004, A&A, 417, 235
Elsworth Y., Hekker S., Basu S., Davies G. R., 2017, MNRAS, 466, 3344
Farnir M., Dupret M. A., Buldgen G., Salmon S. J. A. J., Noels A., Pinçon C., Pezzotti C., Eggenberger P., 2020, A&A, 644, A37
Frandsen S., et al., 2013, A&A, 556, A138
Gaia Collaboration 2022, VizieR Online Data Catalog: Gaia DR3 Part 1. Main source (Gaia Collaboration, 2022), VizieR Online Data Catalog: I/355. Originally published in: Astron. Astrophys., in prep. (2022), doi:10.26093/cds/vizier.1355
Garrett J. D., 2021,] 10.5281/zenodo.4106649
Gaulme P., et al., 2016, ApJ, 832, 121
Ghezzi L., Montet B. T., Johnson J. A., 2018, ApJ, 860, 109
Goldreich P., Murray N., Willette G., Kumar P., 1991, ApJ, 370, 752
Gough D. O., 1990, in Osaki Y., Shibahashi H., eds, Lecture Notes in Physics, Berlin Springer Verlag Vol. 367, Progress of Seismology of the Sun and Stars. p. 283, doi:10.1007/3-540-53091-6
Grevesse N., Sauval A. J., 1998, Space Sci. Rev., 85, 161
Gruberbauer M., Guenther D. B., Kallinger T., 2012, ApJ, 749, 109
Gruberbauer M., Guenther D. B., MacLeod K., Kallinger T., 2013, MNRAS, 435, 242
Grundahl F., Clausen J. V., Hardis S., Frandsen S., 2008, A&A, 492, 171
Guenther D. B., Brown K. I. T., 2004, ApJ, 600, 419
Guenther D. B., Demarque P., 2000, ApJ, 531, 503
Guggenberger E., Hekker S., Basu S., Bellinger E., 2016, MNRAS, 460, 4277
Handberg R., Miglio A., Brogaard K., Bossini D., Elsworth Y. P., 2016, Astronomische Nachrichten, 337, 799
Handberg R., Brogaard K., Miglio A., Bossini D., Elsworth Y., Slumstrup D., Davies G. R., Chaplin W. J., 2017, MNRAS, 472, 979
Harris C. R., et al., 2020, Nature, 585, 357
Hauser H. M., Marcy G. W., 1999, PASP, 111, 321
Hekker S., Christensen-Dalsgaard J., 2017, A&ARv, 25, 1
Hekker S., et al., 2010, ApJ, 713, L187
Hon M., Kuszlewicz J. S., Huber D., Stello D., Reyes C., 2022, AJ, 164, 135
Houdek G., Trampedach R., Aarslev M. J., Christensen-Dalsgaard J., 2017, MNRAS, 464, L124
Huber D., et al., 2019, AJ, 157, 245
Jaynes E. T., 2003, Probability theory: The logic of science. Cambridge University Press, Cambridge
Jeffreys H., 1998, Theory of probability. Oxford Classic Texts in the Physical Sciences, The Clarendon Press Oxford University Press, New York
Jeffries Jr. M. W., et al., 2013, AJ, 146, 58
Jenkins J. M., et al., 2016, in Chiozzi G., Guzman J. C., eds, Society of Photo-Optical Instrumentation Engineers (SPIE)

Conference Series Vol. 9913, Software and Cyberinfrastructure for Astronomy IV. p. 99133E, doi:10.1117/12.2233418
Johnson J. A., Marcy G. W., Fischer D. A., Wright J. T., Reffert S., Kregenow J. M., Williams P. K. G., Peek K. M. G., 2008, ApJ, 675, 784
Kallinger T., 2019, arXiv, p. arXiv:1906.09428
Kallinger T., et al., 2010a, A&A, 509, A77
Kallinger T., Gruberbauer M., Guenther D. B., Fossati L., Weiss W. W., 2010b, A&A, 510, A106
Kallinger T., et al., 2010c, A&A, 522, A1
Kallinger T., et al., 2012, A&A, 541, A51
Kallinger T., et al., 2014, A&A, 570, A41
Kallinger T., Beck P. G., Stello D., Garcia R. A., 2018, A&A, 616, A104
Karovicova I., White T. R., Nordlander T., Casagrande L., Ireland M., Huber D., 2022, A&A, 658, A47
Kervella P., Mignard F., Mérand A., Thévenin F., 2016, A&A, 594, A107
Kervella P., Bigot L., Gallenne A., Thévenin F., 2017, A&A, 597, A137
Kjeldsen H., Bedding T. R., 1995, A&A, 293, 87
Kjeldsen H., Bedding T. R., Christensen-Dalsgaard J., 2008, ApJ, 683, L175
Li T., Bedding T. R., Huber D., Ball W. H., Stello D., Murphy S. J., Bland -Hawthorn J., 2018, MNRAS, 475, 981
Li Y., et al., 2023, MNRAS, 523, 916
Mamajek E. E., et al., 2015, arXiv, p. arXiv:1510.07674
Mathur S., et al., 2012, ApJ, 749, 152
Mazumdar A., et al., 2014, ApJ, 782, 18
McKeever J. M., Basu S., Corsaro E., 2019, ApJ, 874, 180
Metcalfe T. S., et al., 2012, ApJ, 748, L10
Metcalfe T. S., Creevey O. L., Davies G. R., 2015, ApJ, 811, L37
Miglio A., et al., 2010, A&A, 520, L6
Miglio A., et al., 2012, MNRAS, 419, 2077
Mortier A., Santos N. C., Sousa S. G., Adibekyan V. Z., Delgado Mena E., Tsantaki M., Israelian G., Mayor M., 2013, A&A, 557, A70
Mosser B., et al., 2011, A&A, 525, L9
Mosser B., et al., 2012, A&A, 540, A143
Mosser B., et al., 2014, A&A, 572, L5
Mosser B., Vrard M., Belkacem K., Deheuvels S., Goupil M. J., 2015, A&A, 584, A50
Mosser B., Pinçon C., Belkacem K., Takata M., Vrard M., 2017, A&A, 600, A1
Ong J. M. J., Basu S., 2019a, ApJ, 870, 41
Ong J. M. J., Basu S., 2019b, ApJ, 885, 26
Paxton B., et al., 2019, ApJS, 243, 10
Pinsonneault M. H., et al., 2014, ApJS, 215, 19
Pinsonneault M. H., et al., 2018, ApJS, 239, 32
Pinsonneault M. H., et al., 2025, ApJS, 276, 69
Reimers D., 1975, Memoires of the Societe Royale des Sciences de Liege, 8, 369
Rendle B. M., et al., 2019, MNRAS, 484, 771
Ricker G. R., et al., 2015, Journal of Astronomical Telescopes, Instruments, and Systems, 1, 014003
Rodrigues T. S., et al., 2017, MNRAS, 467, 1433
Rosenthal C. S., Christensen-Dalsgaard J., Nordlund Å., Stein R. F., Trampedach R., 1999, A&A, 351, 689
Roxburgh I. W., 2005, A&A, 434, 665
Ryabchikova T., Pakhomov Y., Mashonkina L., Sitnova T., 2022, MNRAS, 514, 4958
Sandquist E. L., et al., 2013, ApJ, 762, 58
Schlattl H., Weiss A., Ludwig H. G., 1997, A&A, 322, 646
Sharma S., Stello D., Bland-Hawthorn J., Huber D., Bedding T. R., 2016, ApJ, 822, 15
Silva Aguirre V., et al., 2017, ApJ, 835, 173
Skilling J., 2006, Bayesian Analysis, 1, 833
Skrutskie M. F., et al., 2006, AJ, 131, 1163
Sonoi T., Samadi R., Belkacem K., Ludwig H.-G., Caffau E., Mosser B., 2015, A&A, 583, A112
Stein R. F., Nordlund Å., 1991, in Gough D., Toomre J., eds, , Vol. 388, Challenges to Theories of the Structure of Moderate-Mass Stars. p. 195, doi:10.1007/3-540-54420-8_67
Stello D., et al., 2011a, ApJ, 737, L10
Stello D., et al., 2011b, ApJ, 739, 13
Tang Y. K., Gai N., 2011, A&A, 526, A35
Tassoul M., 1980, ApJS, 43, 469
Themeßl N., et al., 2018, MNRAS, 478, 4669
Townsend R. H. D., Teitler S. A., 2013, MNRAS, 435, 3406
Valle G., Dell'Omodarme M., Prada Moroni P. G., Degl'Innocenti S., 2025, A&A, 698, A111
Verbunt F., Phinney E. S., 1995, A&A, 296, 709
Verma K., et al., 2015, in European Physical Journal Web of Conferences. p. 06066, doi:10.1051/epjconf/201510106066
Virtanen P., et al., 2020, Nature Methods, 17, 261
Vrard M., et al., 2015, A&A, 579, A84
Vrard M., Kallinger T., Mosser B., Barban C., Baudin F., Belkacem K., Cunha M. S., 2018, A&A, 616, A94
Walker G., et al., 2003, PASP, 115, 1023
Weiss W. W., et al., 2014, PASP, 126, 573
Wes McKinney 2010, in Stéfan van der Walt Jarrod Millman eds, Proceedings of the 9th Python in Science Conference. pp 56 – 61, doi:10.25080/Majora-92bf1922-00a
White T. R., Bedding T. R., Stello D., Christensen-Dalsgaard J., Huber D., Kjeldsen H., 2011, ApJ, 743, 161
White T. R., et al., 2013, MNRAS, 433, 1262
White T. R., et al., 2018, MNRAS, 477, 4403
Wu T., Li Y., Hekker S., 2014a, ApJ, 781, 44
Wu T., Li Y., Hekker S., 2014b, ApJ, 786, 10
Xiang M., et al., 2019, ApJS, 245, 34
Yang J. Y., Li Y., 2007, MNRAS, 375, 403
Zhugzhda Y. D., Stix M., 1994, A&A, 291, 310
de Meulenaer P., Carrier F., Miglio A., Bedding T. R., Campante T. L., Eggenberger P., Kjeldsen H., Montalbán J., 2010, A&A, 523, A54
di Mauro M. P., et al., 2011, MNRAS, 415, 3783
mpmath development team T., 2023, mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 1.3.0)
van Leeuwen F., 2007, A&A, 474, 653

APPENDIX A: AUTOMATED PEAK BAGGING

Seismic forward modelling requires accurate oscillation frequencies extracted from the time series observations as input. So-called "peak bagging" (i.e., the extraction of individual mode parameters from the observed power density spectrum – PDS) has proven itself extensively for main-sequence and sub-giant stars, but has so far only been applied to a handful of red giants (e.g. Corsaro et al. 2018; di Mauro et al. 2011; Handberg et al. 2016; Themeßl et al. 2018). This is mainly due to complications introduced by the complex structure of the mixed dipole modes (e.g. Mosser et al. 2012). However, Kallinger (2019) presented a new approach to overcome this difficulties and applied it to more than 6 000 red giants from the APOKASC sample (Pinsonneault et al. 2018).

The peak bagging algorithm used here, is part of the `STELLAR`[11] pipeline, which also characterises – prior the

[11] https://github.com/tkallinger/KeplerRGpeakbagging/tree/master/STELLAR

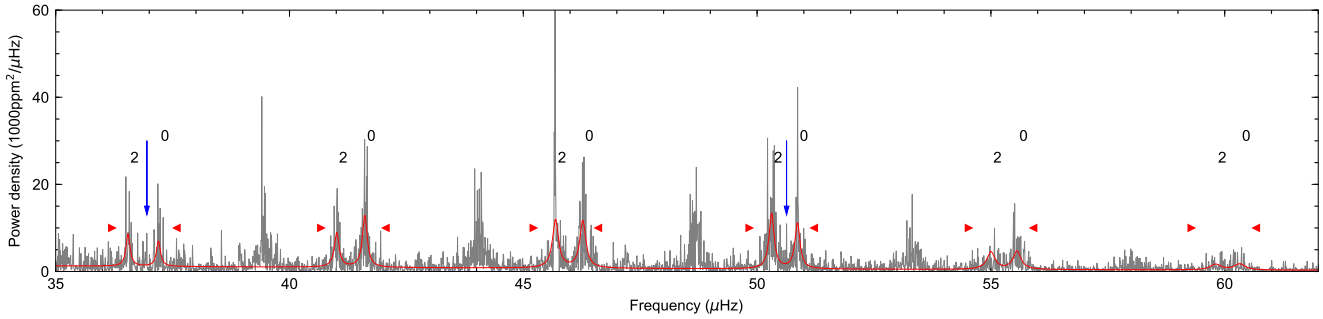

Figure A1. Power density spectrum (grey line) of KIC 8410637 in the frequency range of the oscillations with the red line representing a fit to $l = 0$ and 2 modes. Red triangle symbols enclose the range in which the modes are searched and blue arrows mark possible dipole modes in the vicinity of even modes, which potentially distort the fit.

mode extraction – the granulation background and global properties of the oscillation power excess (i.e., the frequency of maximum oscillation power $\nu_{\rm max}$, the frequency of the central radial mode $\nu_{\rm c}$, the local large and small frequency separations $\Delta\nu_{\rm c}$ and $\delta\nu_{02}$, respectively, and the evolutionary stage, in case of a red giant). See Kallinger et al. (2010c,a, 2012, 2014) for more details. Unlike other methods that fit all modes at once, `STELLAR` fits them sequentially. This approach is well justified since, for instance, the $l = 0$ and 2 modes of consecutive radial orders can be treated as independent, leading to a significant increase in computational speed.

Based on the prerequisite, the peak bagging starts in the centre of the power excess and fits two superposed Lorentzians (to cover the $l = 0$ and 2 modes),

$$P_{\rm fit}(\nu) = B(\nu) + \sum_i \frac{a_i^2 \tau_i}{1 + 4[\pi\tau_i(\nu - \nu_i)]^2}, \tag{A1}$$

on top of the granulation background (B) to the PDS in the frequency range $[-1.5, 0.5]\delta\nu_{02}$ around the estimated position of a given radial mode. For the fit, the Bayesian nested sampling algorithm `UltraNest` (Buchner 2021) is used. It delivers the posterior probability distributions for the frequency (ν_i), rms amplitude (a_i), and mode lifetime (τ_i) of the i-th profile, from which the best-fit parameters and their 1σ uncertainties are determined. It thereby uses a logarithmic likelihood function,

$$\ln\mathcal{L} = -\sum_\nu \left(\frac{P_{\rm obs}(\nu)}{P_{\rm fit}(\nu)} + \ln P_{\rm fit}(\nu) \right), \tag{A2}$$

where $P_{\rm obs}$ and $P_{\rm fit}$ are the observed and fit power density, respectively.

However, mixed dipole modes that sometimes appear in close vicinity to the radial and quadrupole modes potentially distort the fit (see Fig. A1). To identify such modes the fit-to-background ratio (FBR) is used. Corsaro et al. (2015) found that a FBR of eight compares well to a 99% probability of a peak to be not due to noise. All sharp peaks with FBR > 8 are therefore very likely dipole modes and the corresponding frequency bins are excluded from the PDS before re-fitting with Eq. A1.

To rate the significance of the individual modes, the algorithm considers four scenarios:

$\mathcal{S}1$: Both modes are statistically significant.
$\mathcal{S}2$: Only the $l = 0$ mode is significant ($l = 2$ is absent).
$\mathcal{S}3$: Only the $l = 2$ mode is significant ($l = 0$ is absent).
$\mathcal{S}4$: No significant signal.

Given that $\mathcal{S}1$ to $\mathcal{S}4$ cover all possible cases the total probability has to be equal to one. The individual probabilities are then given as, $p_{\mathcal{S}i} = z_{\mathcal{S}i}/\sum_j z_{\mathcal{S}j}$, where $z_{\mathcal{S}i}$ is the global evidence of scenario i as delivered by `UltraNest`. To determine $z_{\mathcal{S}4}$ is quite simple as it requires the "fit" to follow the background (i.e., no additional signal is present) and the corresponding model has no free parameter. The global evidence simplifies to the likelihood defined by Eq. A2, with $P_{\rm fit} = B$. In oder to determine the global evidences for $\mathcal{S}2$ and $\mathcal{S}3$, one could fit only one profile to the PDS, which is, however, computationally expensive. A much faster but equivalent solution is not to test the presence of one mode but the absence of the other. This again simplifies to models with no free parameter for which the global evidences directly result from Eq. A2. In case of $\mathcal{S}2$, $P_{\rm fit}$ is equal to B plus the radial mode profile from $\mathcal{S}1$. For $\mathcal{S}3$, $P_{\rm fit}$ turns into B plus the quadrupole mode profile from $\mathcal{S}1$. The resulting $p_{\mathcal{S}2}$ and $p_{\mathcal{S}3}$ give the probabilities that only the $l = 2$ and $l = 0$ mode is insignificant, respectively.

The mode significance can therefore be defined with the odds ratios $\mathcal{O}_{l=0} = (p_{\mathcal{S}1} + p_{\mathcal{S}2})/(p_{\mathcal{S}3} + p_{\mathcal{S}4})$ and $\mathcal{O}_{l=2} = (p_{\mathcal{S}1} + p_{\mathcal{S}3})/(p_{\mathcal{S}2} + p_{\mathcal{S}4})$. In practice a mode is considered to be significant[12] if $\mathcal{O} > 5$. The significant modes are then prewhitened from the spectrum according to $P_{\rm obs,i} = P_{\rm obs,i-1} \cdot B/P_{\rm fit}$ for the i-th iteration, with $P_{\rm fit}$ (from Eq. A1) including only the significant mode profiles.

The search for radial and quadrupole modes is continued from the centre of the power excess outwards until no significant modes in two consecutive radial orders if found. The individual radial mode positions are thereby predicted from $\nu_{0,n}$ that results from the previous step as $\nu_{0,n\pm1} = \nu_{0,n} \pm \Delta\nu_{\rm c}$.

Once all significant $l = 0$ and 2 modes are prewithened, the residual PDS then only contains $l = 1$ and 3 modes, which are then sequentially extracted following the method of Kallinger (2019) until no significant mode is left. `STELLAR` can also fit rotationally split $l = 1$ and 2 modes, but

[12] Note that in probability theory $p \gtrsim 0.8$ (i.e. an odds ratio of 5:1) is considered as substantial evidence (Jeffreys 1998).

Table A1. Global seismic parameters (taken from Kallinger et al. 2018) and significant radial and quadrupole mode frequencies of KIC 8410637. All values are in μHz.

	$\nu_{\rm max}$	$\Delta\nu_{\rm c}$	$\Delta\nu_{\rm cor}$
	46.87±0.23	4.648±0.011	4.594±0.005
l	this work	$\nu_{l,n}$ Themeßl et al. (2018)	Li et al. (2018)
2	-	32.13±0.05	-
0	-	32.79±0.02	-
2	36.55±0.02	36.56±0.03	-
0	37.20±0.02	37.18±0.02	-
2	41.01±0.01	41.02±0.01	41.01±0.03
0	41.61±0.01	41.62±0.01	41.63±0.03
2	45.69±0.01	45.69±0.01	45.68±0.01
0	46.27±0.02	46.28±0.02	46.29±0.03
2	50.31±0.02	50.31±0.02	50.32±0.04
0	50.86±0.02	50.85±0.01	50.86±0.02
2	55.00±0.07	55.06±0.04	55.05±0.05
0	55.56±0.02	55.54±0.03	55.53±0.05
2	59.81±0.05	59.74±0.05	59.70±0.09
0	60.32±0.02	60.28±0.05	60.37±0.03

this option is not used here. A detailed description of the automated peak bagging approach including extensive tests and comparisons with results from others will be subject to a separate paper. First results for radial mode amplitudes and lifetimes were presented by Vrard et al. (2018).

A1 KIC 8410637

To demonstrate the automatic peak bagging I utilise the four year-long *Kepler* time series of KIC 8410637 (see Sec. 2.1). Fig. A1 shows the PDS of KIC 8410637 along with the co-added best-fit radial and quadrupole mode profiles. The extracted frequencies (see Tab. A1) compare well with those determined by Themeßl et al. (2018) and Li et al. (2018) except for the mode pair around 32 μHz from Themeßl et al. (2018). However, this was also picked up by my algorithm but found to be not significant.

A2 KIC 9970396

As for KIC 8410637, the four year-long *Kepler* time series of KIC 9970396 are analysed with `STELLAR`, where the eclipses of the companion star are left out for simplicity. The results are shown in Fig. reffig:9970396spec and the significant mode frequencies are listed in Tab. A2, which compare reasonably well with those of Li et al. (2018).

A3 HR 6817

For the present study, the 2-min cadence Pre-search Data Conditioning Simple Aperture Photometry (PDC_SAP) data are used provided by the TESS Science Processing Operations Center pipeline (Jenkins et al. 2016, SPOC) and accessible via the Mikulski Archive for Space Telescopes (MAST) at the Space Telescope Science Institute. Individual observations are only used if their quality flag is equal to zero and no further outliers are removed. The resulting 1901 day-long time series consists of about 492 000 data points with a duty cycle of about 38%. Even tough there are long gaps in the time series, the spectral window is clean enough to only marginally influence the subsequent analysis.

Data preparation and the subsequent peak bagging follow the approach described above and the result is shown in Fig. A3. Significant mode frequencies are listed in Tab. A3.

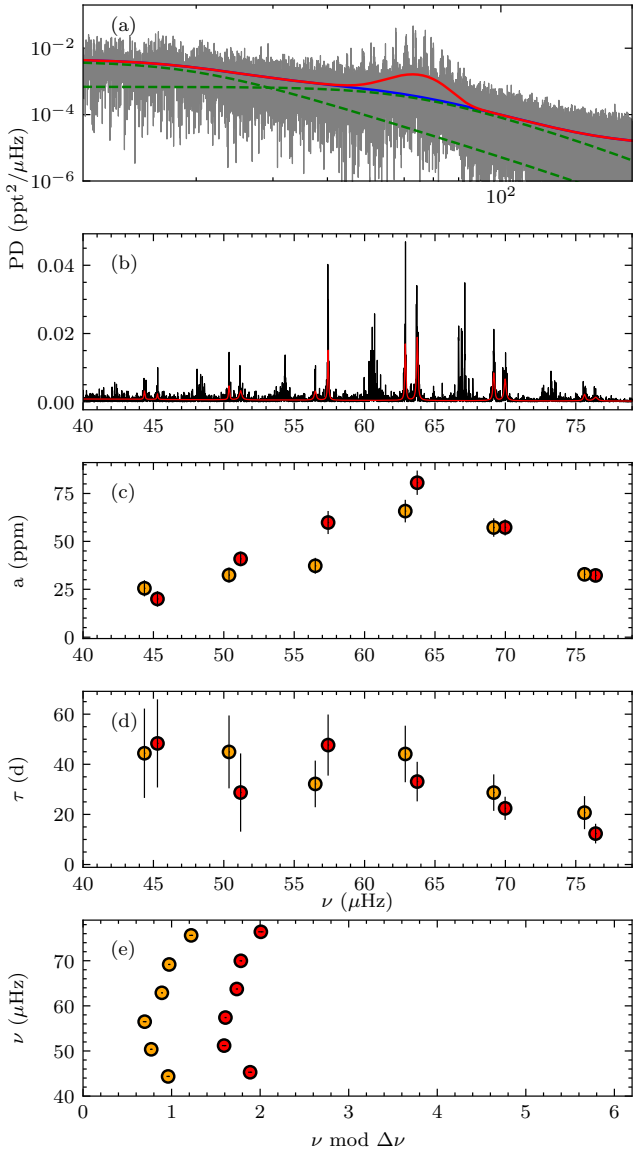

Figure A2. Seismic analysis of KIC 9970396 from `STELLAR`. Panel (a) shows the power density spectrum (grey) of the *Kepler* time series along with a global fit (red line). The blue line is the fit without the Gaussian component and serves as background for the subsequent peak bagging. Dashed green lines give the two granulation components. Panel (b) shows the power density (black) in the frequency range of the oscillations. The red line indicates the fit with Eq. A1. Panel (c) and (d) are the amplitude and lifetime of the significant radial (red) and quadrupole (orange) modes, respectively. Panel (e) shows and Echelle diagram of these modes.

Table A2. Global seismic parameters and significant $l = 0$, 1, and 2 mode frequencies determined from the *Kepler* observations of KIC 9970396. All values are in μHz.

$\nu_{\rm max}$	$\Delta\nu_{\rm c}$	$\Delta\nu_{\rm cor}$
63.42 ± 0.25	6.301 ± 0.007	6.286 ± 0.006
$l = 0$	$l = 2$	$l = 1$
45.29 ± 0.02	44.36 ± 0.02	48.10 ± 0.01
		48.35 ± 0.02
		48.48 ± 0.01
		48.62 ± 0.01
51.19 ± 0.03	50.37 ± 0.01	53.74 ± 0.01
		54.16 ± 0.01
		54.35 ± 0.01
57.41 ± 0.01	56.50 ± 0.02	59.97 ± 0.01
		60.34 ± 0.01
		60.53 ± 0.02
		60.69 ± 0.01
		60.86 ± 0.01
		61.14 ± 0.01
		61.49 ± 0.01
63.74 ± 0.01	62.89 ± 0.01	66.39 ± 0.01
		66.68 ± 0.01
		66.90 ± 0.01
		67.12 ± 0.01
		67.40 ± 0.01
		67.69 ± 0.01
		67.99 ± 0.01
69.98 ± 0.02	69.17 ± 0.01	72.71 ± 0.02
		73.04 ± 0.01
		73.26 ± 0.01
		73.57 ± 0.01
		73.90 ± 0.01
		74.26 ± 0.01
76.41 ± 0.03	75.62 ± 0.02	

Table A3. Same as Table A2 but for the TESS observations of HR 6817. All values are in μHz.

$\nu_{\rm max}$	$\Delta\nu_{\rm c}$	$\Delta\nu_{\rm cor}$
248.7 ± 0.5	17.158 ± 0.005	17.206 ± 0.009
$l = 0$	$l = 2$	$l = 1$
179.24 ± 0.04	177.04 ± 0.05	186.70 ± 0.08
		188.35 ± 0.01
195.85 ± 0.02	193.75 ± 0.10	
212.82 ± 0.02	210.73 ± 0.05	219.36 ± 0.05
		221.61 ± 0.07
230.01 ± 0.02	228.07 ± 0.04	235.99 ± 0.02
		239.45 ± 0.02
		243.15 ± 0.02
247.08 ± 0.01	245.13 ± 0.04	250.82 ± 0.10
		255.68 ± 0.01
		257.30 ± 0.04
		261.01 ± 0.03
264.30 ± 0.01	262.49 ± 0.04	268.01 ± 0.01
		272.99 ± 0.01
		276.62 ± 0.02
281.71 ± 0.03	279.76 ± 0.05	282.66 ± 0.02
		288.73 ± 0.02
		291.64 ± 0.03
299.17 ± 0.04	297.18 ± 0.07	

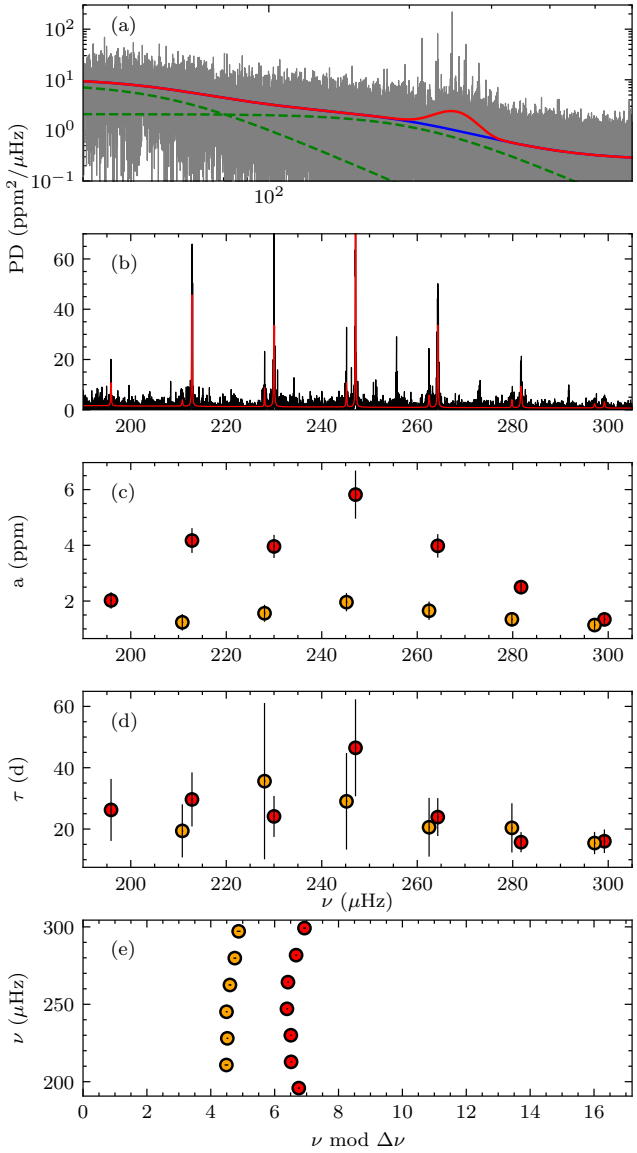

Figure A3. Same as Fig. A2 but for the TESS observations of HR 6817.

APPENDIX B: GIRD INTERPOLATION AND NESTED SAMPLING

The stellar model grids used in the present analysis are not sufficiently dense to produce well-sampled posterior probability distributions (see Fig. 3) from which reliable best-fit parameters and their uncertainties can be determined. This limitation arises because the eigenfrequencies of adjacent models differ by more than the typical observational uncertainties of the measured frequencies (approximately 10 to 50 nHz for the stars analyzed here). While increasing the grid density is a potential solution, it is computationally impractical.

A more practical solution is to interpolate the missing models. The existing grids are dense enough that the parameter variations (including eigenfrequencies) between adjacent models can be accurately approximated through

linear interpolation. This is demonstrated in Fig. B1, which compares real and interpolated model frequencies along a $1.525\,M_\odot$ evolutionary track. For testing, this additional "test track" was computed, located between the 1.5 and $1.55\,M_\odot$ tracks of grid G2 used in Sec. 2. This setup represents a worst-case scenario for interpolation accuracy. Along the test track, pairs of mass and radius were used to bilinearly interpolate all other parameters from the four surrounding models on the 1.5 and $1.55\,M_\odot$ tracks.

During normal stellar evolution, the absolute differences between real and interpolated model frequencies is 2 ± 0.8 nHz, which is much smaller than typical observational uncertainties and comparable to the findings of Clara et al. (2025). Only during the first dredge-up phase – occurring in the range between about 11.3 and $12.2\,R_\odot$ – do interpolation support points become inconsistently allocated, leading to larger uncertainties. Even then, the differences remains on average below 10 nHz, comparable to or smaller than observational uncertainties. Importantly, this represents a worst-case scenario, meaning the interpolation uncertainties for other regions of the mass-radius plane within the model grid are smaller than those shown in Fig. B1. For other interpolated parameters, such as age and luminosity, the uncertainties are similarly small, averaging around 0.1%, while the effective temperature shows deviations of about 1 K.

In practice, `BaseFM` generates pairs of mass and radius values and interpolates the remaining model parameters at these points. Rather than increasing the grid resolution through interpolation and performing a grid search, it employs a nested sampling approach to efficiently identify the best-fit model.

Nested Sampling (e.g., Skilling 2006) is a Monte Carlo method primarily used to compute the Bayesian evidence $\mathcal{Z}$ (i.e., the marginal likelihood) and weighted posterior samples for parameter estimation. The key idea involves transforming the high-dimensional integral of the evidence into a one-dimensional integral over the prior volume $\mathcal{X}$.

Initially, a set of N random live points is sampled from the prior distribution $\pi(\theta)$. At each iteration, the point with the lowest likelihood $\mathcal{L}^*$ is removed and replaced by a new point drawn from the prior under the condition $\mathcal{L}(\theta) > \mathcal{L}^*$. This shrinks X exponentially, while $\mathcal{Z}$ is approximated as a weighted sum over the discarded points. The algorithm terminates when the remaining prior volume contributes negligibly to the evidence.

The final evidence and posterior weights (w) are computed from the discarded and live points as,

$$\mathcal{Z} \approx \sum_k \mathcal{L}_k \Delta X_k \qquad\qquad w_k = \frac{\mathcal{L}_k \Delta X_k}{\mathcal{Z}} \qquad\qquad \text{(B1)}$$

where ΔX_k represents the change in prior volume at each step.

Normally, the main challenge lies in efficiently sampling from the constrained prior, requiring sophisticated methods. In the present case, the prior volume reduces to a 2D space (i,.e. with the model mass and radius as parameters), where the likelihood can only be high along ridges of equal mean density. During live point replacement, it is sufficient enough to search for new points in the vicinity of these ridges.

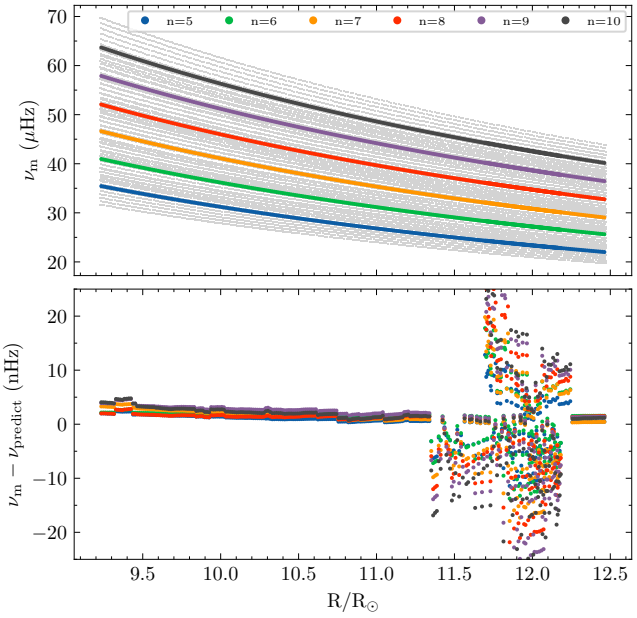

Figure B1. Comparison between actual and interpolated model frequencies in the vicinity of the best-fit model for KIC 8410637. The top panel shows the evolotion of the central six radial modes (colored symbols) along a $1.525\,M_\odot$ evolotionary track. The grey symbols indicate model frequencies of the original Z=0.023 grid. The bottom panel gives the difference between actual and the for models along $1.525\,M_\odot$ track interpolated frequencies. The region with the large scatter is due to the first dredge-up phase where evolution follows a small loop in the HR-diagram (i.e. models with near equal mass and radius but different inner structure are located next to each other).

APPENDIX C: THE GLITCH AMPLITUDE IN STELLAR MODELS

To verify that the amplitudes of the frequency variations that result from acoustic glitches in the stellar models in fact depend on the model mass I apply Eq. 15 to the radial model frequencies covering ±3 orders around ν_c, which is assumed to be the radial mode closest to the theoretical ν_{max} of the model. To extract the properties of the glitch signal one would now need to fit a damped harmonic function to the resulting curvature corrected frequencies $\nu'(n)$. In practice, such a fit is ill-defined and tends to not converge properly. However, the standard deviation of $\nu'(n)$ is already a good proxy for the amplitude of the glitch signal A_{glitch}.

The resulting glitch amplitudes are exemplarily shown in Fig. C1 for the RGB part of the model grid used in Sec. 2. The figure shows that for the used model grid, the glitch amplitude varies between about 0.2 and 1.7% of $\Delta\nu$, which corresponds to the range that is observed in real stars. More importantly, it also shows that for models with a given $\Delta\nu$ (i.e. constant mean density), A_{glitch} correlates with the model mass. The glitch signal therefore in fact improves the otherwise weak discriminative power between mass and radius of `BaseFM`.

This paper has been typeset from a TEX/LATEX file prepared by the author.

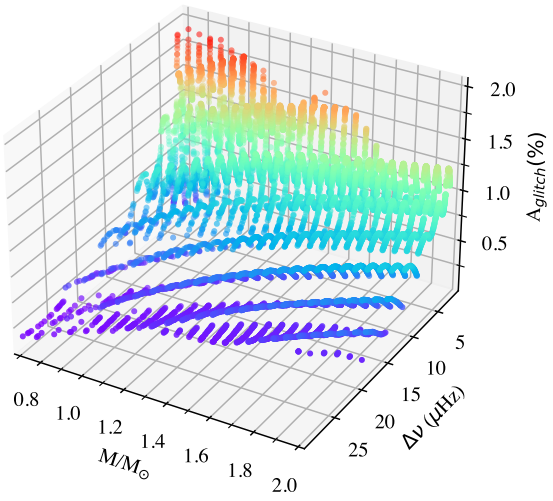

Figure C1. Glitch amplitudes (given in per cent of $\Delta\nu$) of stellar models as a function of the model mass and large frequency separation.