

Probing stellar cores from inversions of frequency separation ratios

J. Bétrisey  and G. Buldgen

Observatoire de Genève, Université de Genève, Chemin Pegasi 51, 1290 Versoix, Switzerland
e-mail: Jerome.Betrisey@unige.ch

Received 25 March 2022 / Accepted 18 May 2022

ABSTRACT

Context. With the rapid development of asteroseismology thanks to space-based photometry missions such as CoRoT, *Kepler*, TESS, and in the future, PLATO, and the use of inversion techniques, quasi-model-independent constraints on the stellar properties can be extracted from a given stellar oscillation spectrum. In this context, inversions based on frequency separation ratios, which are less sensitive to surface effects, appear as a promising technique to constrain the properties of stellar convective cores.

Aims. We developed an inversion based on frequency separation ratios, with the goal of damping the surface effects of the oscillation frequencies. Using this new inversion, we defined a new indicator to constrain the boundary mixing properties of convective cores in solar-like oscillators.

Methods. We verified our inversion technique by conducting tests in a controlled environment, where the stellar mass and radius were known exactly, and conducted an extensive hare and hounds exercise.

Results. The inversion is not affected by surface effects. With the construction of an extensive set of models, favoured and forbidden regions can be highlighted in the parameter space. If the ratios are well fitted, unsurprisingly the inversion does not provide any additional information.

Conclusions. The indicator, coupled with the inversion based on frequency separation ratios, seems promising at probing the properties of convective cores, especially for F-type stars exhibiting solar-like oscillations.

Key words. stars: interiors – stars: fundamental parameters – asteroseismology

1. Introduction

In the last two decades, the space-based photometry missions CoRoT (Baglin et al. 2009), *Kepler* (Borucki et al. 2010), and TESS (Ricker et al. 2015) were launched, enabling a rapid development of asteroseismology. Further progress is also expected from the future PLATO mission (Rauer et al. 2014). The quality of the data from these missions made possible the use of so-called seismic inversion techniques, which had previously been applied with tremendous success in helioseismology (see Buldgen et al. 2019a; Christensen-Dalsgaard 2021, for recent reviews). The situation for asteroseismology is, however, very different from that of the Sun, as geometric cancellation forbids the detection of modes with angular degrees higher than three for solar-like oscillators. This limits the available data sets to about 50 frequencies at best, whereas thousands of modes are observed for the Sun. As a consequence, the philosophy of inversions was reoriented. Indeed, the classical inversions for the Sun are local inversions that can scan through its entire structure, but they require a large set of frequencies in order to be efficient. This can only be achieved for the best asteroseismic targets (e.g. 16CygA and 16CygB, Bellinger et al. 2017, 2019, 2021; Buldgen et al. 2022), and with the limitation that the information can only be extracted for the deep layers. Therefore, inversions based on so-called indicators were developed (Reese et al. 2012; Buldgen et al. 2015a,b, 2018), and applied to various cases (e.g. Buldgen et al. 2016a,b, 2017a, 2019b,c; Bétrisey et al. 2022). The idea was to concentrate all the information of the frequency

spectrum to conduct a global inversion that constrains a global quantity, the indicator (e.g. the mean density).

In this work, we are interested in the entropy profile in the central adiabatic convective regions to put some constraints on the properties of the convective core. With the current inversions based on individual frequencies, a large number of them are required, especially at various harmonic degrees. In addition, a very accurate characterisation of the mass and radius is also needed (Buldgen et al. 2017b). Entropy inversions were conducted for the Sun (Buldgen et al. 2017c), and for the 16Cyg binary system (Buldgen et al. 2022) using the indicators defined in Buldgen et al. (2018). However, even though the 16Cyg system is one of the best *Kepler* asteroseismic targets, Buldgen et al. (2022) found that no additional information about the central entropy could be extracted for 16CygB. In addition, inversions based on individual frequencies are sensitive to the surface regions, which is a limiting factor given that the near-surface effects are poorly modelled. In the case of 16CygA, the surface effects and the quality of the reference models prevented Buldgen et al. (2022) from extracting new constraints on the central entropy. Therefore, we chose to develop a new indicator based on frequency separation ratios instead of individual frequencies, with the aim of carrying out inversions of solar-like stars presenting convective cores. Indeed, frequency separation ratios are defined to damp surface effects (Roxburgh & Vorontsov 2003) and are more efficient at probing central stellar regions (Otí Floranes et al. 2005). Furthermore, Deheuvels et al. (2016) and Farnir et al. (2019) have shown that

frequency separation ratios could be used to constrain the extent of convective cores in solar-like oscillators.

In Sect. 2, we recall the state-of-the-art of indicator-based inversions using individual frequencies, and introduce a new inversion based on frequency separation ratios. In Sect. 3, we define a new indicator that probes central stellar regions and test it in a controlled environment. In Sect. 4, we conduct a hare and hounds exercise to simulate how the indicator would behave with a real observed target. Finally, in Sect. 5, we draw the conclusions of our study.

2. Theoretical considerations

In the inversion community, the following terminology is generally applied. The inversion aims to retrieve the properties of an ‘observed’ or ‘target’ model, which can either be a real observation or a synthetic target. To this end, the inversion takes a ‘reference’ model as an input, which is the result of a modelling procedure, and gives a small correction that is added to a quantity of interest computed from the reference model as an output. For simplicity in this manuscript, an ‘inverted’ quantity is defined as the quantity of interest that includes the correction from the inversion.

2.1. State-of-the-art of individual frequencies

The structure inversion equation is based on a perturbative analysis of the stellar oscillations at linear order. This approach is motivated by the work of Lynden-Bell & Ostriker (1967) and their predecessors (see e.g. Chandrasekhar 1964; Chandrasekhar & Lebovitz 1964; Clement 1964), who showed that the equation of motion fulfils a variational principle. In the case of individual frequencies, the frequency perturbation is directly related to the structural perturbation (Dziembowski et al. 1990):

$$\frac{\delta \nu^{n,l}}{\nu^{n,l}} = \int_0^R K_{a,b}^{n,l} \frac{\delta a}{a} dr + \int_0^R K_{b,a}^{n,l} \frac{\delta b}{b} dr + O(\delta^2), \quad (1)$$

with ν the oscillation frequency, n the radial order, l the harmonic degree, a and b two structural variables, $K_{a,b}^{n,l}$ and $K_{b,a}^{n,l}$ the structural kernels, and using the definition

$$\frac{\delta x}{x} = \frac{x_{\text{obs}} - x_{\text{ref}}}{x_{\text{ref}}}, \quad (2)$$

where the index ref stands for reference, and obs stands for observed. Historically, Eq. (1) was derived for the sound speed and density structural pair (c^2, ρ) , but the equation can be generalised to any combination of physical quantities appearing in the adiabatic oscillation equations (Elliott 1996; Kosovichev 2011; Buldgen et al. 2017b).

Given the frequency differences, the idea is then to solve Eq. (1) to have access to the structure differences and compute a global integrated quantity, the so-called indicator t (e.g. the mean density), which concentrates all the information of the frequency spectrum. The general form of the indicator is

$$t = \int_0^R f(r)g(a)h(b)dr, \quad (3)$$

where f is a weight function that depends on the radius, and g and h are two functions of the structural variables. Except in rarer cases (see e.g. Buldgen et al. 2015a), $h(b)$ is set to one,

and the indicator is only a function of one of the structural variables. In practice, the options for g are limited, because the target function \mathcal{T}_t in Eq. (4) must be reproducible with the number of modes available in asteroseismology. Therefore, simple functions such as $g(a) = a$, $g(a) = 1/a$, etc. are good candidates. A more detailed explanation of the determination of f and g is provided in Sect. 3, where we define the indicator used in this work.

This approach has the advantage of being quasi-model independent because it does not rely on the physics assumed to generate the reference model that serves as starting point for the inversion. In addition, the inversion is also quasi-independent from the starting point in the sense that, if we start from another point in the parameter space, the inversion will still correct towards the exact value.

The inversion can be performed with the subtractive optimally localised averages (SOLA) method (Pijpers & Thompson 1994). This minimisation technique is an adaptation of the OLA approach of Backus & Gilbert (1968, 1970), and allows the inversion to be conducted in a numerically more efficient way. In practice, for a given indicator t , the following cost function is minimised:

$$\mathcal{J}_t(c_i) = \int_0^1 (\mathcal{K}_{\text{avg}} - \mathcal{T}_t)^2 dx + \beta \int_0^1 \mathcal{K}_{\text{cross}}^2 dx + \lambda \left[k - \sum_i c_i \right] + \tan \theta \frac{\sum_i (c_i \sigma_i)^2}{\langle \sigma^2 \rangle} + \mathcal{F}_{\text{Surf}}(\nu), \quad (4)$$

where the averaging and cross-term kernels are related to the structural kernels,

$$\mathcal{K}_{\text{avg}} = \sum_i c_i K_{a,b}^i, \quad (5)$$

$$\mathcal{K}_{\text{cross}} = \sum_i c_i K_{b,a}^i. \quad (6)$$

The normalisation k depends on the properties of the indicator. The balance between the amplitudes of the different terms during the fitting is adjusted with trade-off parameters, β and θ . Here, the idea is to provide a good fit of the target function \mathcal{T}_t while reducing the contribution of the cross-term and of the observational errors to the individual frequencies. We introduce a short notation for the identification pair of a frequency, $i \equiv (n, l)$. The 1σ uncertainty of the relative frequency differences is denoted by σ_i , and $\langle \sigma^2 \rangle = \sum_i^N \sigma_i^2$ is defined, with N the number of observed frequencies. The variable λ is a Lagrange multiplier, and c_i are the inversion coefficients.

The function $\mathcal{F}_{\text{Surf}}(\nu)$ is an empirical term to describe the surface effects. Besides introducing additional free parameters in the minimisation at the expense of the fit of the target function, the treatment of the surface effects constitutes a limitation of the inversion. Indeed, structural kernels based on individual frequencies have a large amplitude in the surface regions, as shown in the upper panel of Fig. 1, and are therefore sensitive to the treatment of these layers. In practice, this limitation can be translated into a systematic uncertainty (see e.g. Bétrisey et al. 2022), but it can become significant enough to prevent the inversion of from correcting accurately.

2.2. Adaptation for frequency separation ratios

Roxburgh & Vorontsov (2003) introduced frequency separation ratios to damp surface effects. They are defined as the ratio of the small frequency separations divided by the large frequency

separations. Because the modes with a higher harmonic degree than $l = 2$ are difficult to observe in asteroseismology, we only consider the following ratios:

$$r_{01}(n) = \frac{d_{01}(n)}{\Delta_1(n)}, \quad (7)$$

$$r_{10}(n) = \frac{d_{10}(n)}{\Delta_0(n+1)}, \quad (8)$$

$$r_{02}(n) = \frac{d_{02}(n)}{\Delta_1(n)}, \quad (9)$$

where $\Delta_l(n)$ are the large separations, and $d_{xy}(n)$ are the small separations:

$$\Delta_l(n) = \nu_{n,l} - \nu_{n-1,l}, \quad (10)$$

$$d_{01}(n) = \frac{1}{8} (\nu_{n-1,0} - 4\nu_{n-1,1} + 6\nu_{n,0} - 4\nu_{n,1} + \nu_{n+1,0}), \quad (11)$$

$$d_{10}(n) = -\frac{1}{8} (\nu_{n-1,1} - 4\nu_{n,0} + 6\nu_{n,1} - 4\nu_{n+1,0} + \nu_{n+1,1}), \quad (12)$$

$$d_{02}(n) = \nu_{n,0} - \nu_{n-1,2}. \quad (13)$$

The first version of a kernel of a ratio was presented in [Otá Floranes et al. \(2005\)](#), with the introduction of the r_{02} kernels, and modulo a different scaling than in this work. We present here the derivation of the r_{01} kernels. The derivations for the r_{10} and r_{02} kernels can be found in [Appendix A](#). As for the individual frequencies, the relative differences of the ratios can be related to the structure differences

$$\frac{\delta r_{01}}{r_{01}}(n) = \frac{\delta d_{01}(n)}{d_{01}(n)} - \frac{\delta \Delta_1(n)}{\Delta_1(n)}, \quad (14)$$

$$= \int_0^R \left(K_{a,b}^{r_{01}}(n) \frac{\delta a}{a} + K_{b,a}^{r_{01}}(n) \frac{\delta b}{b} \right) dr + \mathcal{O}(\delta^2). \quad (15)$$

Because a frequency ratio is a linear combination of individual frequencies, its kernel is a linear combination of the structural kernels of the frequencies used in the definition of the ratio. For example, the r_{01} kernels are the result of the following combination:

$$K_{a,b}^{r_{01}}(n) = \frac{\nu_{n-1,0} K_{a,b}^{n-1,0} - 4\nu_{n-1,1} K_{a,b}^{n-1,1} + 6\nu_{n,0} K_{a,b}^{n,0} - 4\nu_{n,1} K_{a,b}^{n,1} + \nu_{n+1,0} K_{a,b}^{n+1,0}}{\nu_{n-1,0} - 4\nu_{n-1,1} + 6\nu_{n,0} - 4\nu_{n,1} + \nu_{n+1,0}} - \frac{\nu_{n,1} K_{a,b}^{n,1} - \nu_{n-1,1} K_{a,b}^{n-1,1}}{\nu_{n,1} - \nu_{n-1,1}}. \quad (16)$$

The kernels $K_{b,a}^{r_{01}}(n)$ are computed by switching a and b in the previous expression.

[Figure 1](#) shows a comparison of the kernels for the individual frequencies and for the frequency separation ratios, for a $1.3 M_\odot$ star (model *1.3MoAGSS* from [Sect. 3.2](#)). The frequency kernels have a large amplitude at the surface of the star, and are therefore well suited to probe this region, but sensitive to surface effects. On the contrary, ratio kernels have an amplitude that is damped at the surface, suppressing the surface effects. In addition, they have a larger amplitude in the central stellar regions, allowing these regions to be probed more efficiently. We chose to display the kernels of the structural pair (c^2, ρ) because it nicely illustrates this behaviour, but in practice, another structural pair should be used. Indeed, the form of these kernels is not adapted with our indicator, and the contribution of the cross-term is important for this structural pair. For inversions based on frequency separation ratios, structural pairs such as $(S_{5/3}, Y)$ or $(S_{5/3}, \Gamma_1)$ are better suited. The variable $S_{5/3} = P/\rho^{5/3}$ is a proxy

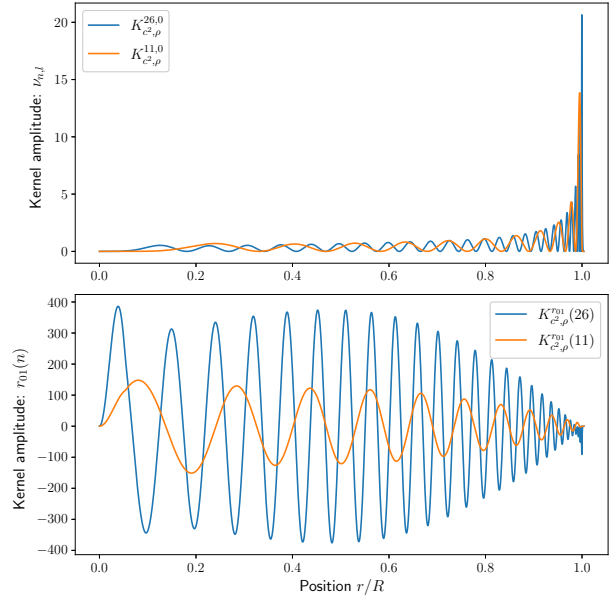


Fig. 1. Frequency kernels versus ratio kernels of the model *1.3MoAGSS* from [Sect. 3.1](#). For both panels, we show the kernel amplitude as a function of the position, and for the structural variable pair (c^2, ρ) . *Upper panel:* low-order ($l = 0, n = 11$) and high-order ($l = 0, n = 26$) kernels for the individual frequencies, respectively in orange and blue. *Lower panel:* low-order ($n = 11$) and high-order ($n = 26$) kernels for the frequency separation ratios, respectively in orange and blue.

of the entropy, P is the pressure, ρ is the density, Y is the helium mass fraction, and $\Gamma_1 = \left(\frac{\partial \ln P}{\partial \ln \rho} \right)_{\text{ad}}$ is the first adiabatic exponent.

There are two main differences for the SOLA cost function. We cannot assume that the ratio uncertainties are uncorrelated because a ratio is, by definition, a combination of frequencies, and the surface effects can be neglected by construction (as we show later in [Sect. 3.1](#)).

$$\mathcal{J}_t(c_i) = \int_0^1 (\mathcal{K}_{\text{avg}} - \mathcal{T}_t)^2 dx + \beta \int_0^1 \mathcal{K}_{\text{cross}}^2 dx + \tan \theta \sum_{p=1}^{N_r} \sum_{q=1}^{N_r} \frac{c_p c_q \Sigma_{pq}}{\langle \Sigma^2 \rangle}. \quad (17)$$

As for the individual frequencies, β and θ are trade-off parameters to balance the amplitudes of the different terms during the fitting. Again, the idea is to provide a good fit of the target function \mathcal{T}_t while reducing the contribution of the cross-term and the observational errors. The variable λ is a Lagrange multiplier, and c_p are the inversion coefficients. The normalisation $\langle \Sigma^2 \rangle = \frac{1}{N_r} \|\Sigma\|_{\text{F}}^2$, where $\|\Sigma\|_{\text{F}}^2$ is the square of the Frobenius norm of the covariance matrix of the ratios, and N_r is the number of observed ratios. The entries of the covariance matrix of the ratios are given by [Eq. \(B.14\)](#) from [Appendix B](#): $\Sigma_{pq} = \text{Cov} \left(\left(\frac{\delta r}{r} \right)_p, \left(\frac{\delta r}{r} \right)_q \right)$.

3. A new indicator probing stellar cores

In this section, we aim to define an indicator that is sensitive to the properties of convective cores. For the first variable, the entropy proxy $S_{5/3}$ appears to be a natural choice. Indeed, this proxy is motivated by the Sackur-Tetrode equation for the

entropy in the case of an ideal gas,

$$S = \frac{3k_B}{2} \left[\mu m_u \ln \left(\frac{P}{\rho^{5/3}} \right) + g(\mu) \right], \quad (18)$$

where k_B is the Boltzmann constant, μ is the mean molecular weight, m_u is the atomic mass unit, and $g(\mu)$ is a function that depends only on the mean molecular weight. They both form a plateau in adiabatic convective regions, and the height and extension of this plateau are related to the boundary mixing properties, such as the temperature and mean molecular weight gradients. Because these boundary regions are generally poorly modelled, an indicator sensitive to their properties can provide new constraints on the physical processes acting in these regions.

For the second structural variable, we chose the helium mass fraction. In comparison to Γ_1 , the amplitude of the cross-term is significantly smaller, which naturally reduces the cross-term contribution in the SOLA cost function. Even if the structural pair including Γ_1 is less well suited, an inversion based on these variables is feasible, although with a lessened stability. Indeed, we can define error measures to quantify the sources of error in an inversion (Buldgen et al. 2015b). The first sources of error are the errors on the averaging and cross-term kernels, which quantify how well the associated target functions are reproduced. The last source of error is the residual error that quantifies the unknown remaining error after taking into account the two previous error measures. In an ideal case, the averaging kernel would perfectly reproduce the target function, the cross-term would be zero, and the inversion could therefore retrieve the exact value. In practice, the cross-term error is negligible, and compensations occur when the averaging and residuals errors are of the same magnitude but with opposite sign. In such cases, the inversion is unstable because a significant unknown source of error affects the inversion result.

In the following section, we define a new indicator that probes central stellar regions, and test it in a controlled environment where the stellar mass and radius between the reference and target models are identical. In these conditions, the mean density is well controlled, which strengthens the stability of the inversion.

3.1. Definition and calibration

We define our indicator as follows:

$$S_{\text{core}}^{r_{01}} = \int_0^R \frac{-f(r)}{S_{5/3}(r)} dr, \quad (19)$$

where $f(r)$ is a weight function. The minus sign is purely cosmetic to ensure that the indicator is positive. Because only a small number of modes are available in asteroseismology, the target function must be constructed so that it can be fitted by the structural kernels. Therefore, the weight function is chosen such that the resulting target function is close to the natural form of the kernels.

We compute the eulerian linear perturbation of the indicator:

$$\frac{\delta S_{\text{core}}^{r_{01}}}{S_{\text{core}}^{r_{01}}} = \int_0^R \frac{f(r)}{S_{\text{core}}^{r_{01}} \cdot S_{5/3}(r)} \frac{\delta S_{5/3}(r)}{S_{5/3}(r)} dr, \quad (20)$$

$$= \int_0^R \mathcal{T}_{S_{\text{core}}^{r_{01}}}(r) \frac{\delta S_{5/3}(r)}{S_{5/3}(r)} dr. \quad (21)$$

Hence, the target function is:

$$\mathcal{T}_{S_{\text{core}}^{r_{01}}}(r) = \frac{f(r)}{S_{\text{core}}^{r_{01}} \cdot S_{5/3}(r)}. \quad (22)$$

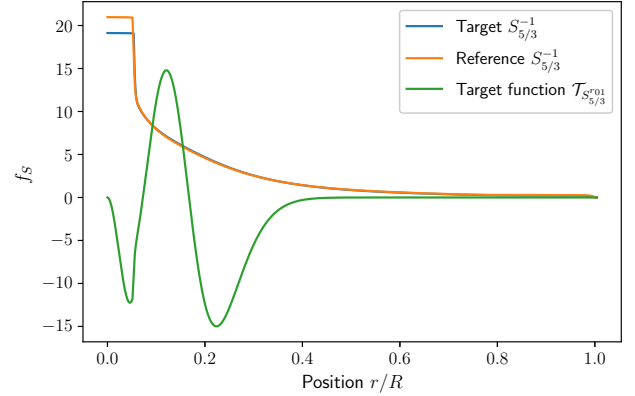


Fig. 2. Target function and entropy profiles for the calibration model. The observed and modelled entropy profiles are in blue and orange, respectively. The target function is in green.

The determination of the weight function is the most difficult part when developing a new indicator. Indeed, there is no straightforward procedure to determine it, except trial and error, and one has to find a good balance between the fine-tuning of the weight function and the trade-off parameters. In practice, to find the weight function, we start with a guess for $f(r)$ and iterate until we find a target function that is reproducible with the available kernels. There is no recipe for the initial guess, but we started with an r^2 term (because the kernels follow approximately this behaviour in the very central regions), an exponential decay (to suppress the signal at the surface), and two Gaussian functions. After a few iterations, we found that including a third Gaussian function to match the number of ‘lobes’ of the target function (illustrated in Fig. 2), as well as the hyperbolic tangent (to more efficiently suppress the signal beyond the central region), produces better results. Because we use Y as structural variable, the cross-term is naturally damped and the value of β has very little impact on the inversion. To find an adequate balance, we calibrated the weight function and the trade-off parameters for one model, the model *1.3MoAGSS* of Sect. 3.2. Then, we verified the indicator in a controlled environment (Sect. 3.2) and with an hare and hounds exercise (Sect. 4). We found that adopting $\beta = 10^{-4}$, $\theta = 10^{-4}$, and

$$f(r) = r^2 e^{-13r} (f_1(r) + f_2(r) + f_3(r)) \tanh(3(1-r)^7), \quad (23)$$

where

$$f_1(r) = -\exp(-150(r-0.02)^2), \quad (24)$$

$$f_2(r) = 3 \exp(-150(r-0.15)^2), \quad (25)$$

$$f_3(r) = -6 \exp(-50(r-0.29)^2), \quad (26)$$

is a good compromise in most of the cases.

In Fig. 2, we show the target function and the entropy profiles for the calibration model to illustrate where the information is localised and to which stellar regions the target function is the most sensitive. The structural differences are mainly located in the central region, up to one-tenth of the stellar radius. Therefore, most of the contribution in the inversion will originate from this region. The first inner peak and the inner part of the second peak of the target function are the most important because they can ‘see’ where the structural differences are localised. Also, it is not necessary to efficiently reproduce the last peak of the target function because it is sensitive to a region where the structural differences are almost non-existent. In addition, we recall

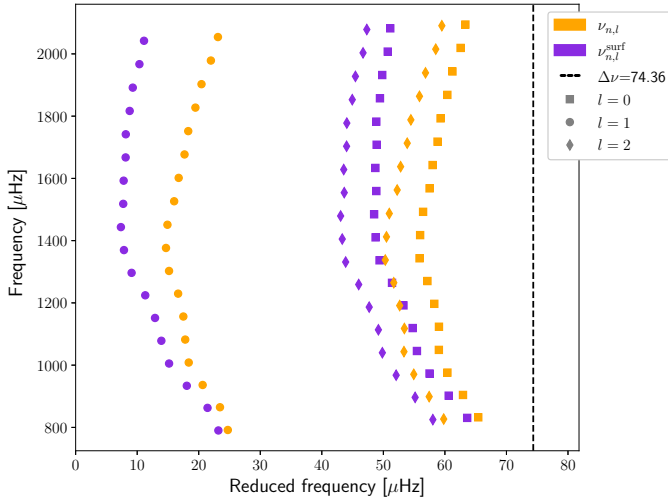


Fig. 3. Impact of the surface effects on the frequencies. The frequencies before (orange) and after (purple) the addition of the surface effects are shown for different harmonic degrees l .

that the kernels are basis vectors that are combined to reproduce the target function. From a linear algebra point of view, some of them will generate the global shape of the target function, while some others will reproduce the details. In our inversion, the global shape is generated by the low-order kernels and they have the highest coefficients. The high-order kernels retrieve some of the details and have smaller coefficients. In this sense, low-order kernels are thus an asset for the inversion. We note that extreme orders ($n < 5$ or $n > 30$) showed a numerically unstable behaviour. This might be a numerical effect that could be fixed or the sign of an intrinsic non-linear behaviour. From a theoretical point of view, further investigations beyond the scope of this article are required. However, it is currently unrealistic to measure modes with $n < 5$, and high-order modes ($n > 30$), as well as being very difficult to observe, will be very sensitive to surface effects.

For the calibration model, we tested that the inversion is indeed not affected by surface effects. To this end, we added surface effects on the frequencies of the synthetic observed model using the [Sonoï et al. \(2015\)](#) prescription, which generates strong surface corrections, especially at high order, as shown in [Fig. 3](#). Even though the frequencies are significantly affected, the impact on the ratios is very small and the relative ratio differences are modified on average by about 0.08%, and much less for low-order modes that carry the inversion. Because this effect is much smaller than any optimistic observational uncertainty, especially at low order, the impact of the surface effects on the inversion is completely negligible.

3.2. Behaviour in a controlled environment

For the preliminary tests in a controlled environment, we generated synthetic ‘observed’ targets in the main sequence with masses of $1.1 M_{\odot}$, $1.2 M_{\odot}$, $1.3 M_{\odot}$, $1.4 M_{\odot}$, and $1.5 M_{\odot}$, and modified one physical ingredient at a time. For the observed targets, we used the OPAL opacities ([Iglesias & Rogers 1996](#)), supplemented by the [Ferguson et al. \(2005\)](#) opacities at low temperature and the electron conductivity by [Potekhin et al. \(1999\)](#). We used the GN93 abundances ([Grevesse & Noels 1993](#)), the FreeEOS equation of state ([Irwin 2012](#)), and microscopic diffusion was described using the formalism of [Thoul et al. \(1994\)](#),

Table 1. Mode sets for the inversions in a controlled environment.

	Radial order n
set 1	11–26
set 2	13–26
set 3	15–26
set 4	18–26
set 5	11–23

Notes. For the models with a mass of $1.4 M_{\odot}$ and $1.5 M_{\odot}$, the maximal radial order is $n_{\max} = 27$ (instead of 26).

but with the screening coefficients of [Paquette et al. \(1986\)](#), and taking partial ionisation into account. The nuclear reaction rates were from [Adelberger et al. \(2011\)](#). We fixed the mixing-length parameter α_{MLT} at a solar calibrated value of 2.05, and following the implementation of [Cox & Giuli \(1968\)](#). For the atmosphere modelling, we used the $T(\tau)$ relation from [Eddington \(1959\)](#). For all the models in this work, we used the Liège Evolution Code (CLES, [Scuflaire et al. 2008a](#)), and the frequencies were computed with the adiabatic Liège Oscillation Code (LOSC, [Scuflaire et al. 2008b](#)).

For the reference models, we changed the abundances for the AGSS09 abundances ([Asplund et al. 2009](#)), the opacities for the OPLIB opacities ([Colgan et al. 2016](#)), and the atmosphere by using the $T(\tau)$ relation described by Model-C in [Vernazza et al. \(1981\)](#) (hereafter VAL-C). We also introduced overshooting and undershooting, respectively with $\alpha_{\text{ov}} = 0.1$ and $\alpha_{\text{under}} = 0.3$, and assuming an adiabatic stratification in both cases. We selected the reference models in the evolutionary sequence by matching the radius with the observation. Concerning the uncertainty of the relative ratios differences, we adopted an error of 20% (hereafter equal-weighted uncertainties) that corresponds to the expectation for high-quality data.

For the inversions, we considered several sets of modes, summarised in [Table 1](#). As shown in [Fig. 4](#), with the addition of low-order modes, the averaging kernel better reproduces the target function, especially in the central regions. Moreover, the contribution of high-order modes is small. Indeed, the target function of sets 1 and 5 is almost identical. For sets 3 and 4, the target function is poorly reproduced by the averaging kernel and it is meaningless to carry out an inversion in such unstable conditions. In practice, we would recommend using mode sets similar to sets 1, 2, or 5, and stress that such mode sets can be observed (and were, for example, for some *Kepler* targets).

The inversion results are shown in [Fig. 5](#) for lower mass stars ($1.1 M_{\odot}$ to $1.2 M_{\odot}$) and higher mass stars ($1.3 M_{\odot}$ to $1.5 M_{\odot}$). The atmosphere model and undershooting affect surface regions, and have no impact on the central entropy profile. The indicator is therefore insensitive to these quantities. It shows promising results for the abundances, and especially for the opacities and the overshoot. The latter indeed induce more significant structural changes, especially for the higher masses, which can be captured by the inversion. For the models with the AGSS09 abundances and masses of $1.2 M_{\odot}$, $1.4 M_{\odot}$, and $1.5 M_{\odot}$, the inversion was less successful or unsuccessful because the relative ratio differences were too small. An inversion has a resolution limit linked with the averaging error and how well the target function is fitted. If the reference model is too close to the observed model, the error on the averaging kernel dominates the inversion and prevents a meaningful correction. Hence, if the ratios are well fitted, it is not appropriate to carry out an inversion.

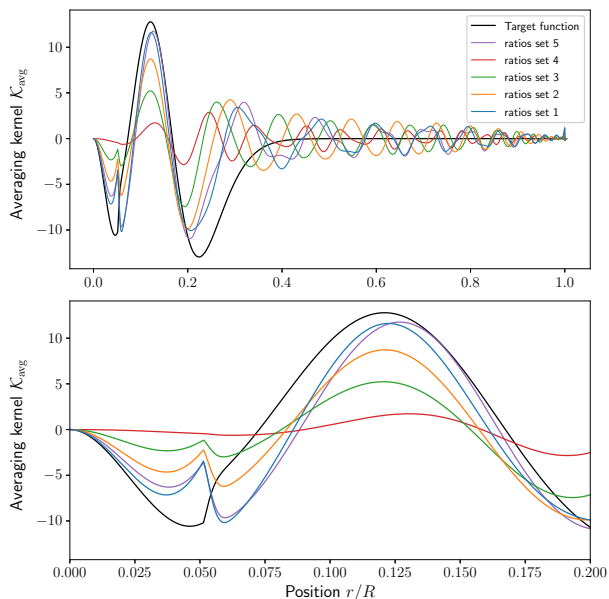


Fig. 4. Averaging kernel considering several sets of modes. *Upper panel:* averaging kernels over the full radius domain. *Lower panel:* zoom on the region that contributes to the inversion.

4. Hare and hounds

4.1. Presentation of the hares

In this hare and hounds exercise, we aimed at computing models with different physical ingredients, unknown to the hounds, to reproduce as realistically as possible the presence of additional physical processes in observed targets unaccounted for in the seismic modelling procedure. To avoid biases, the hares were generated independently and only the frequencies, as well as the spectroscopic constraints (effective temperature, surface metallicity, and luminosity), were provided to the modeller of the hounds. After all the inversions were carried out, the modeller was given access to the stellar properties and internal profiles of the hares to draw the conclusions of this hare and hounds exercise.

The properties of the hares are presented in Table 2. We varied the opacity tables using either the OPAL or the OPLIB opacities, the reference solar abundances using either the AGSS09 or the GN93 abundances, and the formalism for convection using either the classical mixing-length theory or the full spectrum of turbulence theory (Canuto et al. 1996). The effects of microscopic and turbulent diffusion were considered, and the effects of radiative accelerations on metals were simulated for one hare by solely taking the diffusion of helium and hydrogen into account. Other ingredients were kept unchanged, such as the equation of state (for which we used the FreeEOS equation of state), and the atmosphere model (for which we used a VAL-C atmosphere). To generate realistic frequency uncertainties, we searched in the *LEGACY* sample for stars with similar frequency ranges and adopted their observed uncertainties, namely KIC8394589 (Target 1 and 6), KIC12258514 (Target 2), KIC8228742 (Target 3), KIC6508366 (Target 4), and KIC8938364 (Target 5).

4.2. Results of the hounds

To generate the hounds as realistically as possible, we used an advanced modelling strategy that involved first conducting a

Markov chain Monte-Carlo (MCMC) in a grid, fitting the classical constraints (effective temperature, metallicity, and luminosity) and the individual frequencies. We used the AIMS software (Rendle et al. 2019) to perform the MCMC and two grids that differed in their overshooting value. The global properties of the grids are summarised in Table 3 and the physical ingredients of the grid models are the same as for the observed targets in Sect. 3.2, but with one difference that has no influence; partial ionisation was not taken into account. With our indicator, we study central stellar regions of main-sequence stars, where accounting for partial ionisation (or not) only has a very small impact (Turcotte et al. 1998a,b). If radiative accelerations are included or if the target is close to the end of the main sequence, partial ionisation is then not negligible (Schlatterl 2002; Deal et al. 2018). However, inversions are quasi-model independent, especially they do not rely on the choice of the transport processes. It is, therefore, not problematic if the reference model does not include the most detailed physics possible for the hare and hounds exercise. Next, we carried out a mean density inversion on the optimal MCMC model to add the inverted mean density to the set of constraints, with a conservative precision of 0.6%, and we conducted new MCMC fitting the classical constraints, the inverted mean density, and frequency separation ratios (r_{01} , r_{02} , or both). Finally, we performed local minimisations with a Levenberg-Marquardt algorithm (see e.g. Roweis 1996) by leaving the overshooting parameter free. We summarised the constraints in Table 4 and included the r_{01} in the constraints of some of the hounds to verify that the correction predicted by the inversion was significantly smaller or negligible, as we expected. This also allows us to estimate the impact of the resolution limit. This advanced modelling strategy provided the optimal model of Kepler-93 (which was confirmed by a more elaborated modelling procedure, Bétrisey et al. 2022). For stars with physical ingredients that are too different from the physics of the grids, a series of local minimisations are necessary. This was the case for target 5, for which the MCMC was not successful, and for which we conducted local minimisations by starting with guesses of different masses (models *LM*, *M126*, *M133*, *M135*, and *M140*), and with different physical ingredients (models *Ov020/GN93* and *OPLIB/GN93*). In general, if we fit frequency separation ratios, only the r_{02} ratios can be fitted most of the time. Indeed, the fit of both r_{01} and r_{02} ratios tends to be unstable. These ratios are sensitive to different stellar properties and a minimisation including both of them has to deal with too many constraints, preventing a successful convergence into an optimal solution. An inversion based on the r_{01} ratios is therefore relevant at this stage, because it can provide quasi-model-independent constraints that can highlight favoured and forbidden regions in the parameter space, and provide new insight for the modelling and understanding of the star.

Two main conditions are required for the inversion to provide a meaningful correction (here, we already assume that the linear regime assumption is fulfilled, as it is typically the case with the usual modelling strategies of solar-like stars). First, the indicator must be sensitive to some structural differences; in our case, the differences in the entropy profile in the central stellar regions. Due to our modelling strategy, some of the hounds presented a quasi-identical entropy profile and were thus identical for our indicator. We therefore discarded the redundant models to avoid bias. In the left panels of Fig. 6, we illustrate the non-redundant structural differences of the hounds of targets 2 and 5. The hare profiles are in orange. The entropy profiles of the hounds differ significantly from each other, the height of the entropy plateau is different, the extent of the convective core is

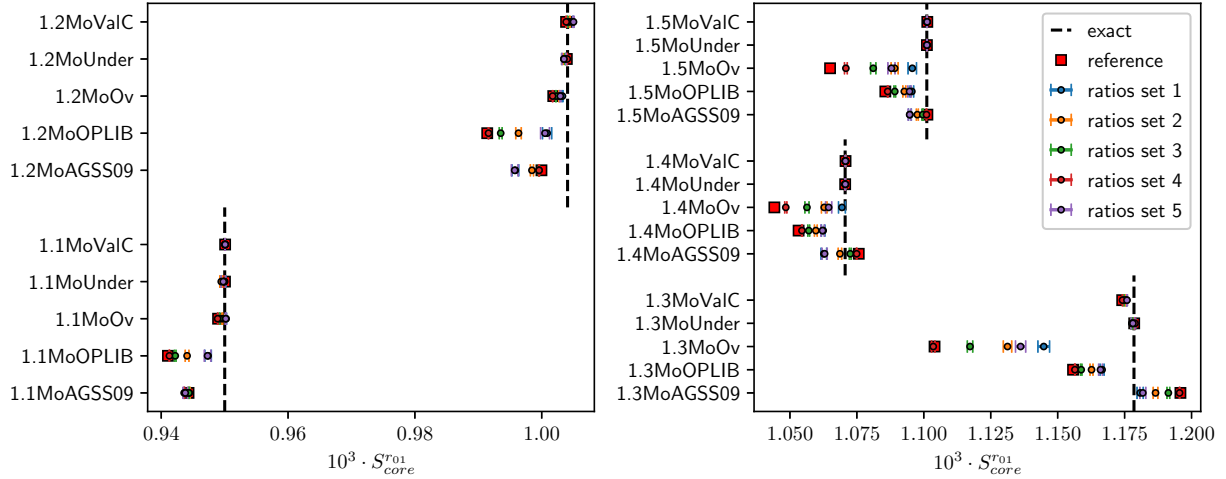


Fig. 5. Behaviour of the indicator in a controlled environment. *Left panel:* results for the lower masses ($1.1 M_{\odot}$ to $1.2 M_{\odot}$). *Right panel:* results for the higher masses ($1.3 M_{\odot}$ to $1.5 M_{\odot}$).

Table 2. Characteristics of the hares.

	Target 1	Target 2	Target 3	Target 4	Target 5	Target 6
Mass (M_{\odot})	1.10	1.26	1.42	1.48	1.35	1.23
Age (Gyr)	3.70	3.55	1.67	2.09	2.09	3.16
Radius (R_{\odot})	1.165	1.515	1.734	2.274	1.816	1.444
T_{eff} (K)	5978	5951	6604	6374	6629	6212
Z_0	0.020	0.028	0.023	0.015	0.0135	0.020
X_0	0.71	0.68	0.69	0.70	0.685	0.71
Abundances	GN93	AGSS09	GN93	AAG21	AGSS09	GN93
Opacities	OPAL	OPAL	OPLIB	OPAL	OPAL	OPAL
EOS	FreeEOS	FreeEOS	FreeEOS	FreeEOS	FreeEOS	FreeEOS
α_{Conv}	1.90	1.75	2.15	2.05	0.85	1.90
α_{Ov}	0.05	0.15	0.10	0.15	0.13	0.10
α_{Under}	0.10	0.15	0.10	0.10	0.10	0.10
Convection	MLT	MLT	MLT	MLT	FST	MLT
Microscopic diffusion	YES	YES	YES	YES – only X, Y	YES	YES
Turbulent diffusion	YES	NO	YES	NO	YES	YES
D_T	7500	0	2500	0	50	7500
β	3	0	2	0	2	3
Atmosphere	VAL-C	VAL-C	VAL-C	VAL-C	VAL-C	VAL-C
Radial order n	13–25	13–25	11–23	11–25	12–23	10–22

Table 3. Global properties of the grids for the MCMC.

	Minimum	Maximum	Step
Mass (M_{\odot})	1.20	1.50	0.02
X_0	0.68	0.72	0.01
Z_0	0.010	0.025	0.001
Overshooting 1		$\alpha_{\text{ov},1} = 0.0$	
Overshooting 2		$\alpha_{\text{ov},2} = 0.1$	

different, and the mixing boundary properties are also different. The corresponding indicator values are displayed in the right panels of Fig. 6 (red squares). We observe that the indicator is indeed sensitive to the structural differences. For example, if the overshooting parameter is changed (e.g. models *NuOv000* and *NuOv010*), the indicator is significantly affected by the resulting structural differences. Second, if the ratios are well fitted,

Table 4. Constraints for the targets 1, 2, 3, 4, and 6.

Label	Constraints	Overshooting
LM	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{02}	Free
LM1	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{01}	Free
NuOv000	[Fe/H], T_{eff} , L , $\nu_{n,l}$	$\alpha_{\text{ov}} = 0.0$
NuOv010	[Fe/H], T_{eff} , L , $\nu_{n,l}$	$\alpha_{\text{ov}} = 0.1$
R01Ov000	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{01}	$\alpha_{\text{ov}} = 0.0$
R01Ov010	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{01}	$\alpha_{\text{ov}} = 0.1$
R02Ov000	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{02}	$\alpha_{\text{ov}} = 0.0$
R02Ov010	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{02}	$\alpha_{\text{ov}} = 0.1$
RatiosOv000	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{01} , r_{02}	$\alpha_{\text{ov}} = 0.0$
RatiosOv010	[Fe/H], T_{eff} , L , $\bar{\rho}_{\text{Inv}}$, r_{01} , r_{02}	$\alpha_{\text{ov}} = 0.1$

the relative ratios differences are too small. The inversion combines these relative differences to compute the correction and

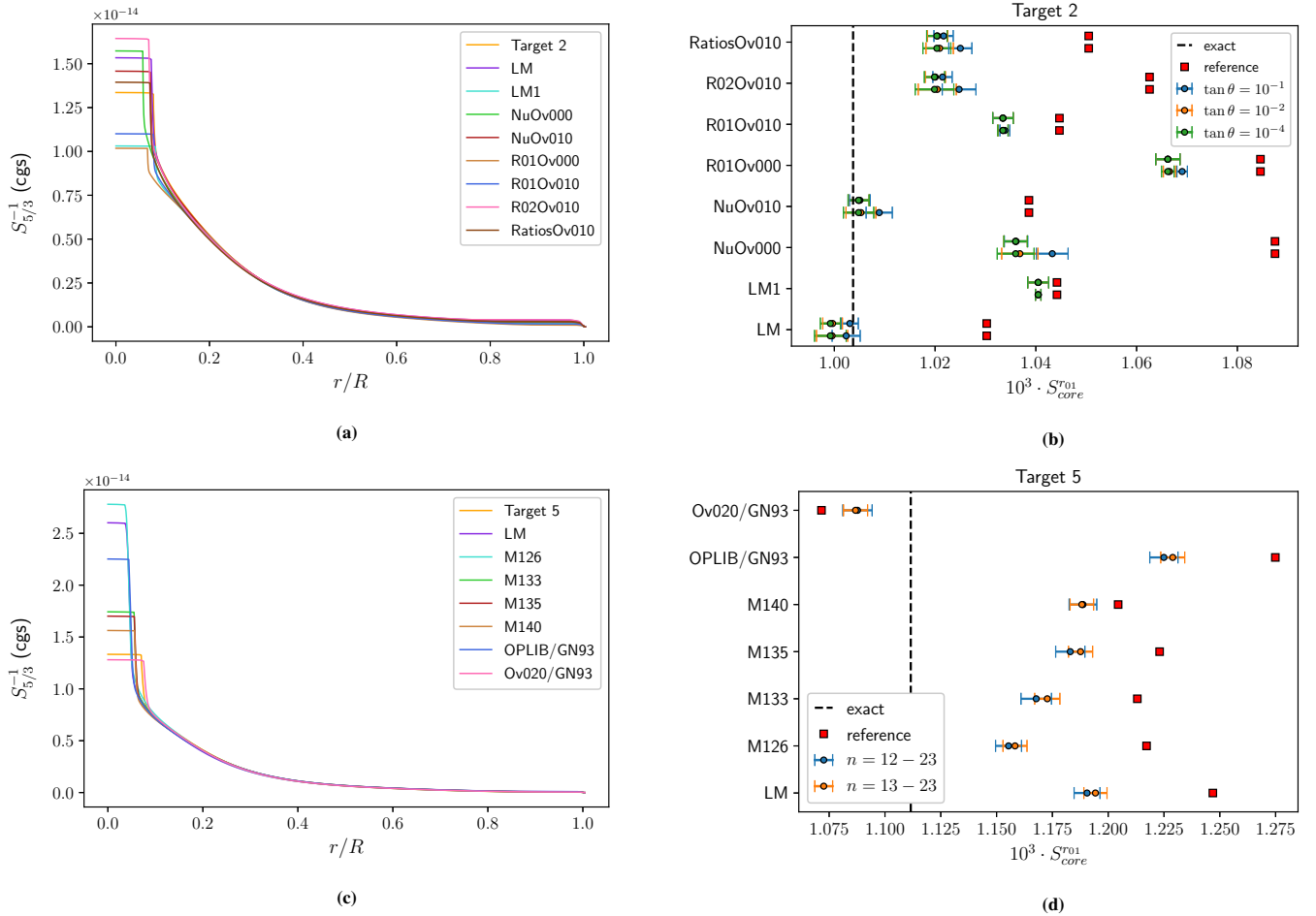


Fig. 6. Entropy profiles and hare and hounds results of targets 2 and 5. The entropy proxy $S_{5/3}$ is in cgs units ($\frac{\text{cm}^4}{\text{g}^{-2/3}\text{s}^{-2}}$). (a) Entropy profiles of target 2. (b) Hares and hounds results of target 2. (c) Entropy profiles of target 5. (d) Hares and hounds results of target 5.

if they are too small, the correction is therefore damped. In an ideal case, this behaviour means that the seismic information is well modelled and that additional information cannot be extracted with an inversion. In other words, if the inversion predicts a tiny correction, it confirms that the seismic information is well described by the reference model. However, inversions are numerical methods, and we must ensure that the correction is not driven by numerical noise. In particular, if the ratios are well fitted, the target function may not be well reproduced because of the limited number of modes available. In that case, a non-negligible correction is driven by the bad fit of the target function and should not be mistaken as a sign that the physics of the reference model is lacking. In Fig. 7, we show examples of relative ratio differences that can and cannot be used for an inversion. We recall that we are conducting linear inversions, and we therefore discarded relative ratios differences that were larger than 0.5 in absolute values, as these were not compatible with the linear formalism. Such large relative differences occur for high-order modes and are purely a numerical artefact. For these modes, the frequency separation ratios are close to zero and automatically produce a large relative difference. This numerical issue is not limiting because the contribution of high-order modes to the correction predicted by inversion is very small. These modes can therefore be discarded without losing any information.

The hares fall into two categories of results. For targets 1, 3, 4, and 6, all the ratios (r_{01} and r_{02}) were well fitted, resulting

in relative ratio differences that were too small for the inversion. In this situation, the reference value stays within the 1σ interval of the inverted value, and no new knowledge can be extracted. The most interesting results were obtained for targets 2 and 5, for which it was not possible to simultaneously fit the r_{01} and r_{02} ratios. In this case, the inversion achieved significant improvements.

The entropy profiles and results for the first category of targets are shown in Fig. 8. Target 3 shows the reason for which we stressed that an inversion should not be carried out on well fitted ratios. Theoretically, we would expect a negligible correction in such conditions, as seen for targets 1 and 4. However, some of the hounds of target 3 exhibit non-negligible corrections, which are a combination of physical structural differences and the resolution limit of the inversion. We observe that the inversion corrects in the right direction, but its magnitude is difficult to trust because it is probably largely dominated by the quality of the fit of the target function. The behaviour of the inversions for target 6 is similar to the results for targets 1 and 4, with one exception; the ratios were not well fitted by the LM model and the entropy profiles were significantly different (Fig. 8g). The conditions for a meaningful inversion were therefore met and the inversion behaves accordingly with our expectations, as illustrated in Fig. 8h.

The results for targets 2 and 5 are displayed in Figs. 6b and 6d, respectively. Because the relative ratio differences were significant enough to carry out an inversion, we conducted more

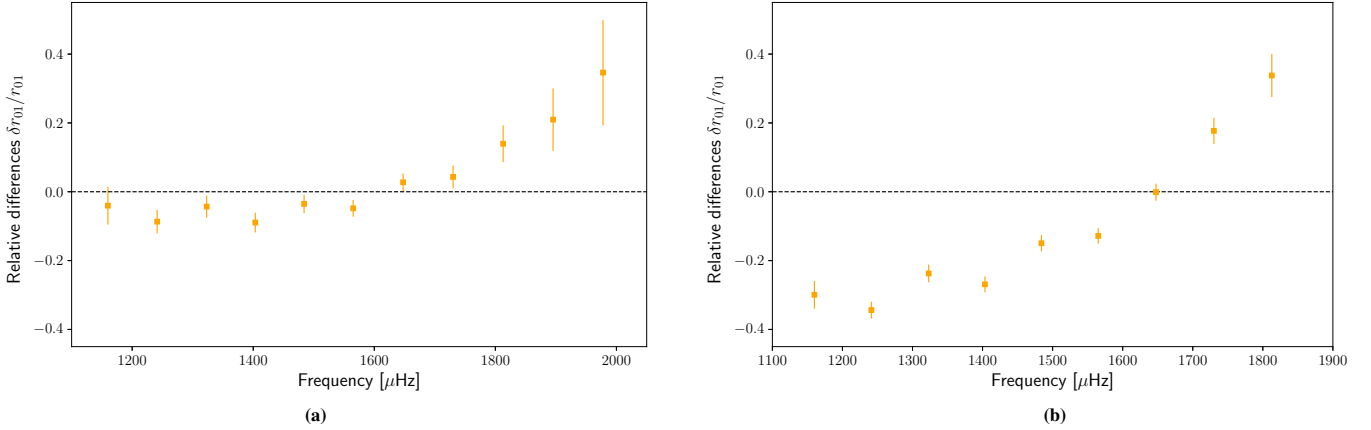


Fig. 7. Comparison of the relative r_{01} differences for models that can (*right*) and cannot (*left*) be used for a meaningful inversion. The relative differences above 0.5 are discarded because they are not appropriate for the linear formalism of the inversion. (a) Model *LM1* of target 2, well fitted r_{01} ratios that are *not* suited for an inversion. (b) Model *LM* of target 2, relative r_{01} differences suited for an inversion.

Table 5. Results for targets 2 and 5.

	$S_{\text{core}}^{r_{01}}$	Accuracy	Precision
<i>Target 2</i>			
Reference	1.057×10^{-3}	5.3%	2.1%
Inverted	1.026×10^{-3}	2.2%	2.2%
Exact	1.004×10^{-3}		
<i>Target 5</i>			
Reference	1.207×10^{-3}	8.6%	5.3%
Inverted	1.171×10^{-3}	5.4%	3.6%
Exact	1.112×10^{-3}		

intensive tests. For target 2, we tested the impact of the trade-off parameter θ , and our choice of $\theta = 10^{-4}$ appears to be a good compromise. Indeed, the inversion result is slightly affected only if a much higher value is considered. In such cases, the target function is less well reproduced because it is equivalent to assuming high observational uncertainties. We also tested two uncertainty sets, where we considered the standard observational uncertainties and checked that equal-weighted uncertainties produce similar results. In this fashion, we could discard the possibility of the presence of highly non-linear modes. For target 5, we tested two mode sets by removing the lowest order mode. For both sets, the inversion significantly improved the reference values, which shows that it is not necessary to use excessively low order modes that could be difficult to observe. Moreover, model *Ov020/GN93* strengthens the consistency of the inversion, because it lies at the opposite side of the exact value with respect to the other models, and therefore corrects in the opposite direction. We point out that, even if we do not have the same convection theory in the hounds of target 5, which strongly impacts the stellar structure, the inversion is still able to correct significantly and in the right direction.

In Table 5, we tried to estimate the improvement achieved by the inversion. We computed the average value¹ of the indicator and defined an accuracy measure as the relative difference of

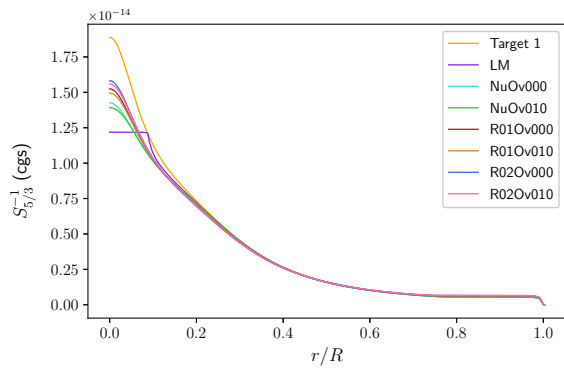
¹ For target 2, we excluded the model *LM1* because it was a fit of the r_{01} ratios, which made the inversion meaningless. We still kept the other models including the r_{01} in the constraints because they produced relative ratio differences large enough to carry out an inversion.

the reference or inverted value with respect to the exact value. For target 2, the accuracy is improved from 5.3% to 2.2%, and for target 5, from 8.6% to 5.4%. This result should be put in perspective with the precision achieved by the inversion. From a pure statistical point of view, the precision is on the order of $\sim 0.1\%$ to 0.2% , which is negligible in comparison with the systematic uncertainty due to the choice of the physical ingredients, and the resolution limit of the inversion. In practice, a modelling strategy similar to that of Bétrisey et al. (2022) should be considered. An extensive set of models with various physical ingredients should be constructed and the precision measured with the standard deviation of the inversion results. With this approach, the precision is significantly lower than the statistical precision, but it is generally smaller or equal to the precision of the reference models. Hence, the inversion can shift and in some cases shrink the reference range towards the exact value.

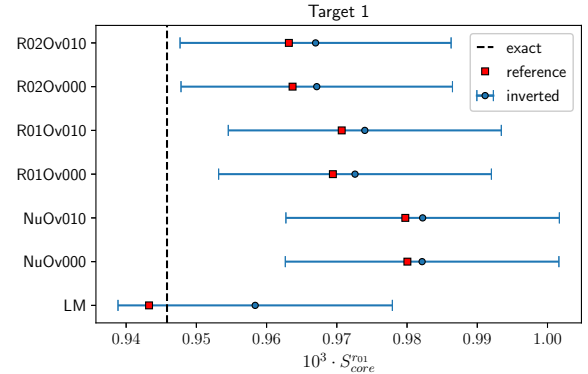
5. Conclusions

In this work, we adapted the SOLA method for frequency separation ratios in Sect. 2 and introduced a new indicator probing the properties of stellar convective cores in Sect. 3.1. Thanks to this approach, the inversion is not affected by surface effects. In Sect. 3.2, we verified our technique in a controlled environment, where the stellar mass and radius were identical in both the reference and observed models. Then, we conducted an intensive hare and hounds exercise in Sect. 4.

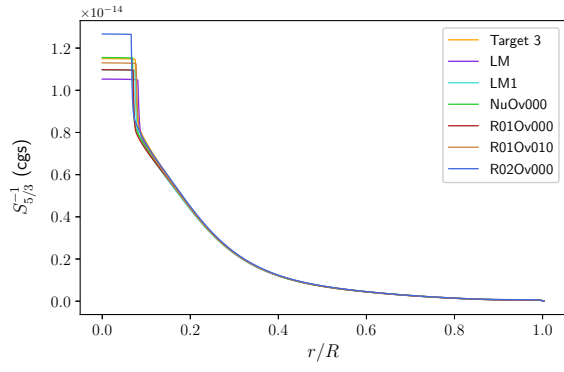
Our inversion technique showed promising results for probing the convective core of main-sequence stars, and is especially suited for F-type stars. Indeed, they have an adequate mass span for the indicator, and the inversion is not sensitive to the surface effects that are difficult to treat for this type of star. Because the entropy differences are localised in the very central stellar regions that are difficult to probe, even with the ratios, we found that a modelling strategy similar to that of Bétrisey et al. (2022) should be considered in a practical application of the method. An extensive set of models should be constructed, and the properties of the core can be inferred from the global behaviour of the inversion on the models of the set. The ratios are influenced by the overshooting, the radiative opacities, and the chemical composition. It is, therefore, not possible to directly constrain one of these quantities, but favoured and forbidden regions can be highlighted in the parameter space. We stress that if the ratios



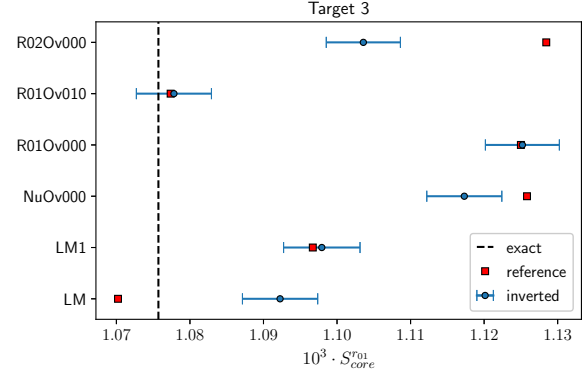
(a)



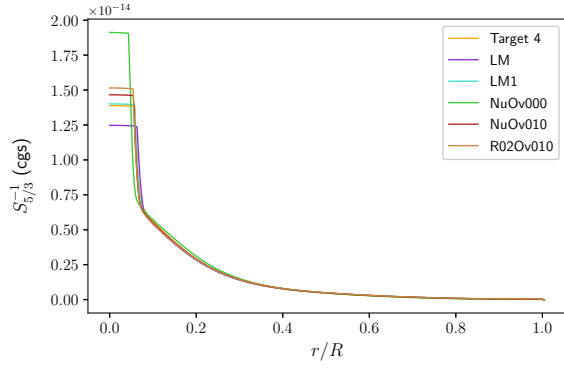
(b)



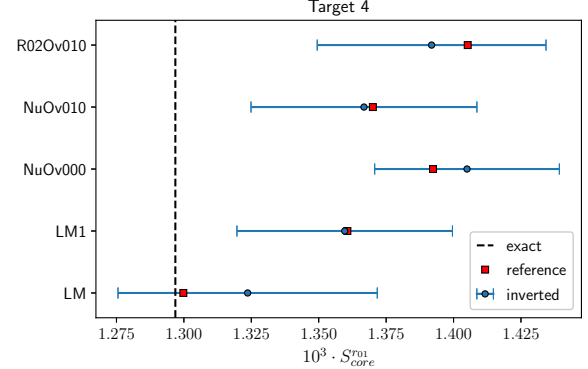
(c)



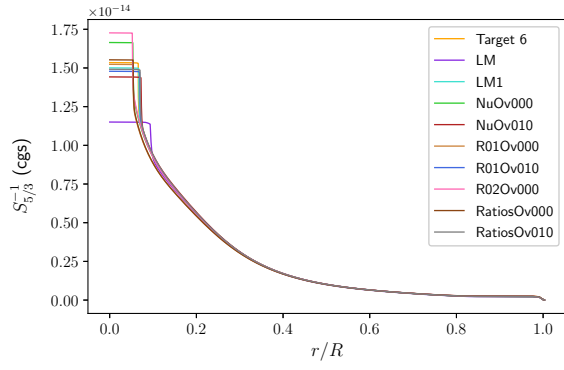
(d)



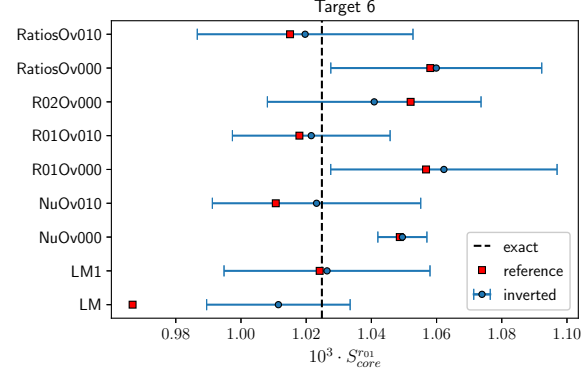
(e)



(f)



(g)



(h)

Fig. 8. Entropy profiles and hare and hounds results of targets 1, 3, 4, and 6. The entropy proxy $S_{5/3}$ is in cgs units ($\frac{cm^4}{g^{-2/3}s^{-2}}$). (a) Entropy profiles of target 1. (b) Hare and hounds results for target 1. (c) Entropy profiles of target 3. (d) Hare and hounds results for target 3. (e) Entropy profiles of target 4. (f) Hare and hounds results for target 4. (g) Entropy profiles of target 6. (h) Hare and hounds results for target 6.

are well reproduced, an inversion is meaningless due to the resolution limit of the method.

In a broader modelling context, the inversion provides quasi-model-independent constraints on core properties that can help to point out the limitations of the reference models and, for example, be the sign of an incorrect or missing physical process. The inversion can also help to exclude some of the reference models, thus improving the accuracy and precision of stellar fundamental parameters such as mass, radius, and age, a crucial objective for the PLATO mission.

Acknowledgements. J.B and G.B. acknowledge funding from the SNF AMBIZIONE grant No 185805 (Seismic inversions and modelling of transport processes in stars).

References

- Adelberger, E. G., García, A., Robertson, R. G. H., et al. 2011, *Rev. Mod. Phys.*, **83**, 195
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, *ARA&A*, **47**, 481
- Backus, G., & Gilbert, F. 1968, *Geophys. J.*, **16**, 169
- Backus, G., & Gilbert, F. 1970, *Philos. Trans. R. Soc. London Ser. A*, **266**, 123
- Baglin, A., Auvergne, M., Barge, P., et al. 2009, in *Transiting Planets*, eds. F. Pont, D. Sasselov, & M. J. Holman, *IAU Symp.*, **253**, 71
- Bellinger, E. P., Basu, S., Hekker, S., & Ball, W. H. 2017, *ApJ*, **851**, 80
- Bellinger, E. P., Basu, S., Hekker, S., & Christensen-Dalsgaard, J. 2019, *ApJ*, **885**, 143
- Bellinger, E. P., Basu, S., Hekker, S., Christensen-Dalsgaard, J., & Ball, W. H. 2021, *ApJ*, **915**, 100
- Bétrisey, J., Pezzotti, C., Buldgen, G., et al. 2022, *A&A*, **659**, A56
- Borucki, W. J., Koch, D., Basri, G., et al. 2010, *Science*, **327**, 977
- Buldgen, G., Reese, D. R., Dupret, M. A., & Samadi, R. 2015a, *A&A*, **574**, A42
- Buldgen, G., Reese, D. R., & Dupret, M. A. 2015b, *A&A*, **583**, A62
- Buldgen, G., Reese, D. R., & Dupret, M. A. 2016a, *A&A*, **585**, A109
- Buldgen, G., Salmon, S. J. A. J., Reese, D. R., & Dupret, M. A. 2016b, *A&A*, **596**, A73
- Buldgen, G., Reese, D., & Dupret, M.-A. 2017a, *Eur. Phys. J. Web Conf.*, **160**, 03005
- Buldgen, G., Reese, D. R., & Dupret, M. A. 2017b, *A&A*, **598**, A21
- Buldgen, G., Salmon, S. J. A. J., Noels, A., et al. 2017c, *A&A*, **607**, A58
- Buldgen, G., Reese, D. R., & Dupret, M. A. 2018, *A&A*, **609**, A95
- Buldgen, G., Salmon, S., & Noels, A. 2019a, *Front. Astron. Space Sci.*, **6**, 42
- Buldgen, G., Rendle, B., Sonoï, T., et al. 2019b, *MNRAS*, **482**, 2305
- Buldgen, G., Farnir, M., Pezzotti, C., et al. 2019c, *A&A*, **630**, A126
- Buldgen, G., Farnir, M., Eggenberger, P., et al. 2022, *A&A*, **661**, A143
- Canuto, V. M., Goldman, I., & Mazzitelli, I. 1996, *ApJ*, **473**, 550
- Chandrasekhar, S. 1964, *ApJ*, **139**, 664
- Chandrasekhar, S., & Lebovitz, N. R. 1964, *ApJ*, **140**, 1517
- Christensen-Dalsgaard, J. 2021, *Liv. Rev. Sol. Phys.*, **18**, 2
- Clement, M. J. 1964, *ApJ*, **140**, 1045
- Colgan, J., Kilcrease, D. P., Magee, N. H., et al. 2016, *ApJ*, **817**, 116
- Cox, J. P., & Giuli, R. T. 1968, *Principles of Stellar Structure* (New York: Gordon and Breach)
- Deal, M., Alecian, G., Lebreton, Y., et al. 2018, *A&A*, **618**, A10
- Deheuvels, S., Brandão, I., Silva Aguirre, V., et al. 2016, *A&A*, **589**, A93
- Dziembowski, W. A., Pamyatnykh, A. A., & Sienkiewicz, R. 1990, *MNRAS*, **244**, 542
- Eddington, A. S. 1959, *The Internal Constitution of the Stars* (New York: Dover Publication)
- Elliott, J. R. 1996, *MNRAS*, **280**, 1244
- Farnir, M., Dupret, M. A., Salmon, S. J. A. J., Noels, A., & Buldgen, G. 2019, *A&A*, **622**, A98
- Ferguson, J. W., Alexander, D. R., Allard, F., et al. 2005, *ApJ*, **623**, 585
- Grevesse, N., & Noels, A. 1993, in *Origin and Evolution of the Elements*, eds. N. Prantzos, E. Vangioni-Flam, & M. Casse, 15
- Iglesias, C. A., & Rogers, F. J. 1996, *ApJ*, **464**, 943
- Irwin, A. W. 2012, *FreeEOS: Equation of State for Stellar Interiors Calculations*
- Kosovichev, A. G. 2011, in *Lecture Notes in Physics*, eds. J. P. Rozelot, & C. Neiner (Berlin: Springer Verlag), 832, 3
- Lynden-Bell, D., & Ostriker, J. P. 1967, *MNRAS*, **136**, 293
- Otí Floranes, H., Christensen-Dalsgaard, J., & Thompson, M. J. 2005, *MNRAS*, **356**, 671
- Paquette, C., Pelletier, C., Fontaine, G., & Michaud, G. 1986, *ApJS*, **61**, 177
- Pijpers, F. P., & Thompson, M. J. 1994, *A&A*, **281**, 231
- Potekhin, A. Y., Baiko, D. A., Haensel, P., & Yakovlev, D. G. 1999, *A&A*, **346**, 345
- Rauer, H., Catala, C., Aerts, C., et al. 2014, *Exp. Astron.*, **38**, 249
- Reese, D. R., Marques, J. P., Goupil, M. J., Thompson, M. J., & Deheuvels, S. 2012, *A&A*, **539**, A63
- Rendle, B. M., Buldgen, G., Miglio, A., et al. 2019, *MNRAS*, **484**, 771
- Ricker, G. R., Winn, J. N., Vanderspek, R., et al. 2015, *J. Astron. Telescopes Instrum. Syst.*, **1**
- Roweis, S. 1996, *Levenberg-marquardt optimization*, Notes, University Of Toronto
- Roxburgh, I. W., & Vorontsov, S. V. 2003, *A&A*, **411**, 215
- Schlattl, H. 2002, *A&A*, **395**, 85
- Scuflaire, R., Théado, S., Montalbán, J., et al. 2008a, *AP&SS*, **316**, 83
- Scuflaire, R., Montalbán, J., Théado, S., et al. 2008b, *AP&SS*, **316**, 149
- Sonoï, T., Samadi, R., Belkacem, K., et al. 2015, *A&A*, **583**, A112
- Thoul, A. A., Bahcall, J. N., & Loeb, A. 1994, *ApJ*, **421**, 828
- Turcotte, S., Richer, J., Michaud, G., Iglesias, C. A., & Rogers, F. J. 1998a, *ApJ*, **504**, 539
- Turcotte, S., Richer, J., & Michaud, G. 1998b, *ApJ*, **504**, 559
- Vernazza, J. E., Avrett, E. H., & Loeser, R. 1981, *APJS*, **45**, 635

Appendix A: The r_{10} and r_{02} kernels

For the r_{02} ratios, we have:

$$\frac{\delta r_{02}}{r_{02}}(n) = \frac{\delta d_{02}(n)}{d_{02}(n)} - \frac{\delta \Delta_1(n)}{\Delta_1(n)}, \quad (\text{A.1})$$

$$= \int_0^R \left(K_{a,b}^{r_{02}}(n) \frac{\delta a}{a} + K_{b,a}^{r_{02}}(n) \frac{\delta b}{b} \right) dr + \mathcal{O}(\delta^2), \quad (\text{A.2})$$

where:

$$K_{a,b}^{r_{02}}(n) = \frac{\nu_{n,0} K_{a,b}^{n,0} - \nu_{n-1,2} K_{a,b}^{n-1,2}}{\nu_{n,0} - \nu_{n-1,2}} - \frac{\nu_{n,1} K_{a,b}^{n,1} - \nu_{n-1,1} K_{a,b}^{n-1,1}}{\nu_{n,1} - \nu_{n-1,1}}. \quad (\text{A.3})$$

The kernel $K_{b,a}^{r_{02}}(n)$ is found by switching a and b in the previous expression.

For the r_{10} ratios, we have:

$$\frac{\delta r_{10}}{r_{10}}(n) = \frac{\delta d_{10}(n)}{d_{10}(n)} - \frac{\delta \Delta_0(n+1)}{\Delta_0(n+1)}, \quad (\text{A.4})$$

$$= \int_0^R \left(K_{a,b}^{r_{10}}(n) \frac{\delta a}{a} + K_{b,a}^{r_{10}}(n) \frac{\delta b}{b} \right) dr + \mathcal{O}(\delta^2), \quad (\text{A.5})$$

where:

$$K_{a,b}^{r_{10}}(n) = \frac{\nu_{n-1,1} K_{a,b}^{n-1,1} - 4\nu_{n,0} K_{a,b}^{n,0} + 6\nu_{n,1} K_{a,b}^{n,1} - 4\nu_{n+1,0} K_{a,b}^{n+1,0} + \nu_{n+1,1} K_{a,b}^{n+1,1}}{\nu_{n-1,1} - 4\nu_{n,0} + 6\nu_{n,1} - 4\nu_{n+1,0} + \nu_{n+1,1}} - \frac{\nu_{n+1,0} K_{a,b}^{n+1,0} - \nu_{n,0} K_{a,b}^{n,0}}{\nu_{n+1,0} - \nu_{n,0}}. \quad (\text{A.6})$$

Again, the kernel $K_{b,a}^{r_{10}}$ is found by switching a and b in the previous expression.

Appendix B: Correlation matrix for the ratios

Let X_i be N random variables, and X_i^0 be their respective mean. We consider the random variable Z which is the general form of a frequency ratio:

$$Z = \frac{\sum_{i=1}^N a_i X_i}{\sum_{i=1}^N b_i X_i}. \quad (\text{B.1})$$

In order to derive the error propagation, we use the first order Taylor expansion around the mean,

$$Z(X_1, \dots, X_n) \simeq Z(X_1^0, \dots, X_n^0) + \sum_{i=1}^N \frac{\partial Z}{\partial X_i}(X_1^0, \dots, X_n^0) (X_i - X_i^0), \quad (\text{B.2})$$

$$= Z(X_1^0, \dots, X_n^0) + \sum_{i=1}^N \frac{(a_i - Z(X_1^0, \dots, X_n^0) b_i) (X_i - X_i^0)}{\sum_k b_k X_k^0}. \quad (\text{B.3})$$

In the following, we introduce the short notation $Z^0 \equiv Z(X_1^0, \dots, X_n^0)$. At the first order, the correlation matrix $\text{Cov}(Z, \hat{Z})$ is equal to

$$\text{Cov} \left(Z^0 + \sum_{i=1}^N \frac{(a_i - Z^0 b_i) (X_i - X_i^0)}{\sum_k b_k X_k^0}, \hat{Z}^0 + \sum_{j=1}^N \frac{(\hat{a}_j - \hat{Z}^0 \hat{b}_j) (X_j - X_j^0)}{\sum_k \hat{b}_k X_k^0} \right), \quad (\text{B.4})$$

$$= \text{Cov} \left(\sum_{i=1}^N \frac{a_i - Z^0 b_i}{\sum_k b_k X_k^0} X_i, \sum_{j=1}^N \frac{\hat{a}_j - \hat{Z}^0 \hat{b}_j}{\sum_k \hat{b}_k X_k^0} X_j \right), \quad (\text{B.5})$$

$$= Z^0 \hat{Z}^0 \sum_{i=1}^N \sum_{j=1}^N \left(\frac{a_i}{\sum_k a_k X_k^0} - \frac{b_i}{\sum_k b_k X_k^0} \right) \left(\frac{\hat{a}_j}{\sum_k \hat{a}_k X_k^0} - \frac{\hat{b}_j}{\sum_k \hat{b}_k X_k^0} \right) \text{Cov}(X_i, X_j), \quad (\text{B.6})$$

where we used the linearity properties of the covariance matrix, and that the covariance of a random variable with a constant is null.

For a set of N_f frequencies ν_i , whose uncertainties are described by the covariance matrix E_{ij} , and i being a short notation for the identification pair (n, l) , a frequency ratio can be expressed as:

$$r = \frac{\sum_{i=1}^{N_f} a_i \nu_i}{\sum_{i=1}^{N_f} b_i \nu_i}. \quad (\text{B.7})$$

For example, for the $r_{02}(n)$ ratio, the coefficients are:

$$\mathbf{a} = [0_{(n_{\min}, 0)}, \dots, 0_{(n-1, 0)}, 1_{(n, 0)}, 0_{(n+1, 0)}, \dots, 0_{(n-2, 2)}, -1_{(n-1, 2)}, 0_{(n, 2)}, \dots, 0_{(n_{\max}, l_{\max})}], \quad (\text{B.8})$$

$$\mathbf{b} = [0_{(n_{\min}, 0)}, \dots, 0_{(n-2, 1)}, -1_{(n-1, 1)}, 1_{(n, 1)}, 0_{(n+1, 1)}, \dots, 0_{(n_{\max}, l_{\max})}], \quad (\text{B.9})$$

where the index indicates the identification pair of the corresponding frequency.

The covariance between two ratios is given by

$$\text{Cov}(r, \hat{r}) = r \hat{r} \sum_{i=1}^{N_f} \sum_{j=1}^{N_f} \left(\frac{a_i}{\sum_k a_k \nu_k} - \frac{b_i}{\sum_k b_k \nu_k} \right) \left(\frac{\hat{a}_j}{\sum_k \hat{a}_k \nu_k} - \frac{\hat{b}_j}{\sum_k \hat{b}_k \nu_k} \right) E_{ij}. \quad (\text{B.10})$$

If the frequencies are uncorrelated, which is the typical assumption, one has $E_{ij} = \sigma_i^2 \delta_{ij}$, where σ_i is the uncertainty of the frequency ν_i . In this case, the covariance simplifies to

$$\text{Cov}(r, \hat{r}) = r \hat{r} \sum_{i=1}^{N_f} \left(\frac{a_i}{\sum_k a_k \nu_k} - \frac{b_i}{\sum_k b_k \nu_k} \right) \left(\frac{\hat{a}_i}{\sum_k \hat{a}_k \nu_k} - \frac{\hat{b}_i}{\sum_k \hat{b}_k \nu_k} \right) \sigma_i^2, \quad (\text{B.11})$$

and the 1σ error bar on r is

$$\sigma_r^2 = r^2 \sum_{i=1}^{N_f} \left(\frac{a_i}{\sum_k a_k \nu_k} - \frac{b_i}{\sum_k b_k \nu_k} \right)^2 \sigma_i^2. \quad (\text{B.12})$$

In the SOLA method, we use relative differences of ratios,

$$\frac{\delta r}{r} = \frac{r_{\text{obs}} - r_{\text{mod}}}{r_{\text{mod}}}, \quad (\text{B.13})$$

where *obs* stands for observed and *mod* for modelled. Eq. (B.11) therefore becomes

$$\text{Cov} \left(\frac{\delta r}{r}, \frac{\delta \hat{r}}{\hat{r}} \right) = \frac{r_{\text{obs}}}{r_{\text{mod}}} \frac{\hat{r}_{\text{obs}}}{\hat{r}_{\text{mod}}} \sum_{i=1}^{N_f} \left(\frac{a_i}{\sum_k a_k \nu_k} - \frac{b_i}{\sum_k b_k \nu_k} \right) \left(\frac{\hat{a}_i}{\sum_k \hat{a}_k \nu_k} - \frac{\hat{b}_i}{\sum_k \hat{b}_k \nu_k} \right) \sigma_i^2. \quad (\text{B.14})$$

The error term in the SOLA cost function is then

$$\tan \theta \sum_{p=1}^{N_r} \sum_{q=1}^{N_r} \frac{c_p c_q \Sigma_{pq}}{\langle \Sigma^2 \rangle}, \quad (\text{B.15})$$

where $\Sigma_{pq} = \text{Cov} \left(\left(\frac{\delta r}{r} \right)_p, \left(\frac{\delta \hat{r}}{\hat{r}} \right)_q \right)$, $\langle \Sigma^2 \rangle = \frac{1}{N_r^2} \|\Sigma\|_F^2$, $\|\Sigma\|_F^2$ is the square of the Frobenius norm of the covariance matrix of the ratios, and N_r is the number of observed ratios.