MODE EXTRACTION FROM TIME SERIES: FROM THE CHALLENGES OF COROT TO THOSE OF *EDDINGTON*

T. Appourchaux[1], O. Moreira[1], G. Berthomieu[2], and T. Toutain[2]

[1]Research and Space Science Department, European Space Agency, P.O. Box 299, 2200 AG Noordwijk, The Netherlands
[2]Département Cassini, UMR CNRS 6529, Observatoire de la Côte d'Azur, BP 4229, 06304 Nice Cedex 4, France

Abstract

With more than 30 years of experience in extraction of eigenmodes from power spectra of solar signals, we are now almost ready to apply this knowledge onto the fore-coming missions: COROT and *Eddington*. Howeverthe fitting task differs by 3 orders of magnitude; COROT will be able to get time series of stellar light for some 30 stars, while *Eddington* will be able to gather such data for about 50 000 stars. While for COROT, our current tools can be applied by hand, the case of *Eddington* is significantly more complex. We are looking forward having automatic fitting procedures that will allow to recover mode parameters for about 90% of the solar-like stars. Unfortunately, about 10% of these stars will require some more delicate attention that will cost time to take care of. We will use the example of the infamous HD 57006, known to be quite evolved with a difficult eigenmode spectrum, to explain how a star can evolve from an easy-to-fit target (90% of the solar-like stars) to a difficult-to-fit one (10% of the remaining stars). In the latter case, new techniques for detecting narrow peaks (g-mode like) out of broad peaks (p-mode like) has been devised in the context of the hare-and-hound exercise of COROT. This and other techniques will be used to implement the automatic fitting procedure for the remaining 10% of *Eddington* solar-like stars.

Key words: Stars: structure – Data analysis: time series – Data analysis: spectra fitting

1. Introduction

In the very near future, there shall be a fleet of space missions aiming at understanding the internal structure of the stars: MOST[1], COROT[2] and *Eddington*.

All these missions will observe global oscillations of the stars by measuring their tiny light fluctuations. The detection and identification of these modes of oscillation is the challenge of all these missions. This challenge is for most stars extremely difficult (e.g. Cepheids) but easier for solar-like stars. When the identification is achieved, peak bagging can be performed. The theory of mode identification and peak bagging has been reviewed by Appourchaux (2003); it is now believed to be well understood.

In the context of the COROT mission, the practice of mode identification and peak bagging is developed through the use of hare-and-hound exercise. During this exercise, it appears that a very challenging star (namely HD 57006) brought a new dimension to the usual challenge[3]. Here we would like to take this star as an example of the kind of difficulties that the *Eddington* mission may face. For that purpose, we will follow HD 57006 throughout its lifetime from the Zero Age Main Sequence (ZAMS) until its present evolutionary state.

2. Mode identification and peak bagging

The reader may wish to read the review on the subject by Appourchaux (2003), which summarize here for the sake of completeness.

The mode identification for solar-like stars is performed using the "Echelle diagram", based on the fact that the low-degree mode frequencies of successive orders (n) of a given degree (l) are spaced from each other by roughly the acoustic diameter of the star ($\Delta\nu_0$). The piling up of sections of the power spectrum (of the stellar time series) cut into piece of length $\Delta\nu_0$ produces 'ridges' of power along an ideal vertical line. The location of the ridges with respect to each other provides the means of tagging the ridges with a given l.

The next step is the peak bagging operation for which the theory can be found in Appourchaux et al. (1998) and Appourchaux (2003). It consists in fitting the power spectrum using Maximum Likelihood Estimators (MLE) and a model of the mode profile including parameters such as frequencies, linewidth, splittings, background noise and profile shape. Error bars of parameters can be derived as well as the significance of these parameters (Appourchaux et al. 1998).

[1] Microvariability and Oscillations of Stars, a Canadian mission to be launched in June 2003.

[2] COnvection and ROTation, a CNES mission to be launched in mid 2006.

[3] See papers of the Third COROT week in `http://www.astro.ulg.ac.be/orientation/asterosis/week3`

3. The challenges in theory

3.1. COROT vs. *Eddington*

The COROT mission shall be able to observe up to ten primary seismological targets for which there will be a high signal-to-noise ratio sufficient to perform proper mode identification and peak bagging. The COROT secondary targets will amount to less than a hundred. The *Eddington* mission shall be able to observe more than 50 000 stars.

For COROT, the primary targets can be analyzed by a single scientist. The mode identification and fitting can be carefully analyzed. Modes out of the main stream can be fitted by hand, and each fitted mode can be assessed for its validity. One could imagine that the secondary targets are also analyzed by the same scientists; the task may start to be somewhat difficult to handle.

For *Eddington*, this hand crafted work is to be abandoned. Automated ridge identification, mode fitting and fit rejection has to be implemented. At the time of writing, the first task is still done by hand using the Echelle diagram and adjusting the large separation as to have vertical ridges. We could envisage a step where after computing the power spectrum, one identifies the excess power, filter the excess power and compute the Fourier transform for extracting the large separation (or as a matter of fact $\Delta\nu_0/2$); this technique was used by Gelly et al. (1986) on the α-Cen data. The next step is to compute the *Echelle diagram* using the derived large separation. As for the proper ridge identification, one could try to fit[4] 2 pairs of peaks ($l = 0 - 2$ and $l = 1 - 3$) over $\Delta\nu_0$; the amplitude of the fitted peaks and their location would automatically provide the degree tagging. The mode fitting can then be done as usual. The last step is the validation of the fit. It is envisaged that each mode be assessed for its significance using the likelihood ratio test (Appourchaux et al. 1998).

3.2. The stellar evolution

But the real challenges of either COROT or *Eddington* may lie beyond these mundane details. The identification and fitting steps described above can be applied to well-behaved solar-like stars; that is about 90% of the solar-like stars to be observed by *Eddington.* The remaining 10% are evolved solar-like stars for which the automated techniques fall apart. Here we should outline that we do not really know the proportion of evolved stars to that of good stars. Even if we were to have only 0.1% of evolved solar-like stars in the *Eddington* mission, we would have the same challenging difficulties. As we will outline, this part of the challenge is not to be neglected.

Here we would like to follow the evolution of a star, to scan the many stars of *Eddington.* We start with a star at the ZAMS, and study how the Echelle diagram evolves with the star, and how the difficulties evolve.

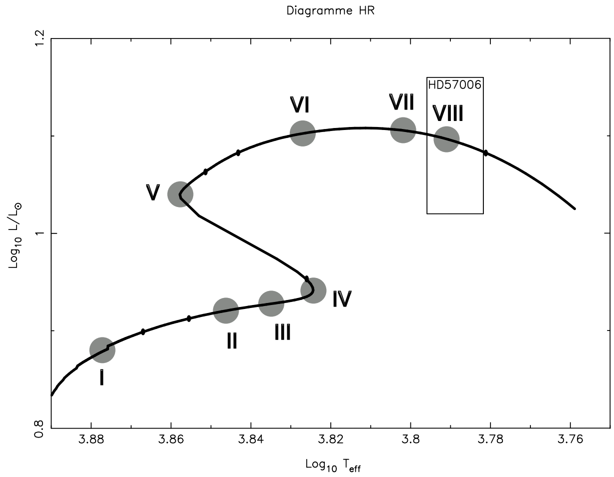


Figure 1. The Hertzsprung-Russell diagram of HD 57006. We computed the mode frequencies of models labeled from I to VIII located at different positions on the evolutionary track of the H-R diagram. The first four locations follow the evolution of the star until the central hydrogen content is about 5%. The last four locations follow the evolution of the star during the shell hydrogen burning. The transition from model IV to V is fast by stellar standard but likely long enough by human standards for being observed.

Here we chose the COROT primary target HD 57006 for which the scientific value is being assessed by the Seismology Working Group of the eponymous mission. The assessment is being performed in the frame work of the third hare-and-hound exercise of COROT (Appourchaux 2003).

The star HD 57006 has been represented by a 1.65 $M_\odot$ stellar model, i.e. still a solar-like star when on the main sequence, according to the definition of Appourchaux (2003). It is nowadays sufficiently evolved as to have entered the phase of burning hydrogen in shells, i.e. it has an helium core. This produces a large peak in the Brunt-Vaïsala frequency at the core of the star. It leads to the existence of the so-called mixed modes that have a p-mode character in the outer stellar regions, and a g-mode character in the stellar core. The mixed modes, if detected, are powerful tools (like the g modes) for understanding the internal structure of the star. Unfortunately, their mixed character make them difficult to detect for they do not follow the asymptotic relationship given by Tassoul (1980); in other words they do not line up for making ridges.

In order to see how the Echelle diagram changes with the evolution of HD 57006, we computed frequencies of low-degree modes along the evolution track of the star as outlined in Fig. 1. The *Echelle diagrams* for each evolution stage in the central hydrogen burning phase are shown in Fig. 2; and for the shell hydrogen burning phase in Fig. 3.

[4] where the peak power is maximum

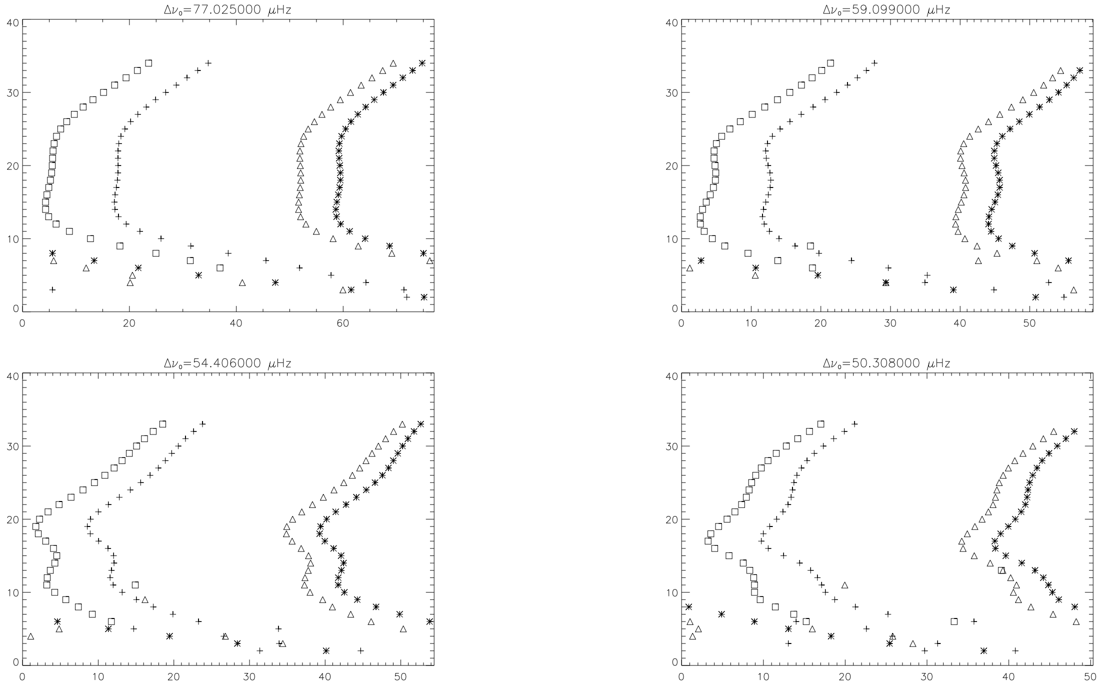


Figure 2. Echelle diagram for the first four stages of the evolution of HD 57006 during the core hydrogen burning phase; I (left, top), II (right, top), III (left, bottom), IV (right, bottom). The echelle diagrams display the order as a function of frequency; the large separation is annotated atop the diagram. The stars and triangles are the $l = 0$ and $l = 2$ modes; the plusses and squares are the $l = 1$ and $l = 3$ modes.

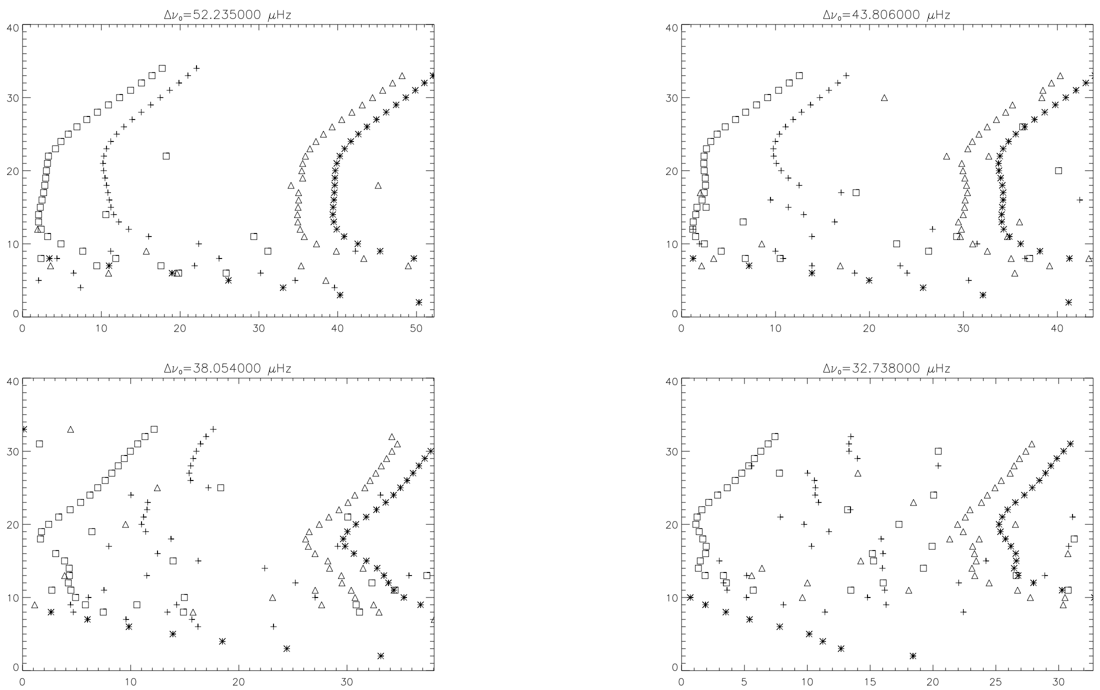


Figure 3. Echelle diagram for the first four stages of the evolution of HD 57006 during the shell hydrogen burning phase; V (left, top), VI (right, top), VII (left, bottom), VIII (right, bottom). The echelle diagrams displays the order as a function of frequency; the large separation is annotated atop the diagram. The stars and triangles are the $l = 0$ and $l = 2$ modes; the plusses and squares are the $l = 1$ and $l = 3$ modes.

In the early evolution stages, the mode identification and fitting would be rather classic. The large separation decreases due to the increase in the stellar radius. The ridges get closer from each other; the $l = 0$ and $l = 2$ mode-ridge separation decreases from 8 μHz to 4 μHz in this phase. This stage of the evolution is likely to be similar to that of 90% of *Eddington*'s solar-like stars.

In the later evolution stages, the mode identification and fitting gets more difficult and challenging. The separation between the $l = 0$ and $l = 2$ mode ridge decreases even more to a mere 2 μHz. The tagging of each ridge could be made extremely difficult if the linewidth mode is about 1 μHz. Nevertheless, it is possible to fit the mode pair as a single mode giving very limited information about the internal structure of the star. Mixed modes start to appear especially for $l = 1$. The $l = 1$ mode ridge is progressively destroyed and disappears completely, rendering their tagging almost impossible. The $l = 1$ modes start to appear everywhere and are even sometimes right next to the other regular p modes. Due to their mixed character, the $l = 1$ modes are likely to be long-lived modes. Due to their erratic location in the Echelle diagram the mixed $l = 1$ modes are bound to have mode frequencies close to the short-lived modes; in which the identification becomes even more difficult. This stage of the evolution is likely to be similar to that of the remaining solar-like stars of *Eddington*.

3.3. Stellar rotation

An additional difficulty to the mode identification is the influence of the rotation upon the determination of the mode frequencies. When the small separation (δ_{02}, δ_{13}) is about 2–3 times the rotational splitting, the proper derivation of the frequencies of the $l = 0 - 2$ and $l = 1 - 3$ modes get ambiguous. It leads to the small separation being negative. This ambiguity is likely to happen for solar-like stars with a rotational splitting ranging from 1 to 5 μHz. In this case, it is required to fit the mode parameters not locally around the $l = 0 - 2$ or $l = 1 - 3$ pair but globally over the spectrum as already implemented by Jiménez et al. (2002) and Gelly et al. (2002) – see Neiner & Apporchaux (2003).

4. HD 57006 as an evolved star

The star HD 57006 is one of the candidate primary target of COROT. In the framework of the third hare-and-hound exercise of COROT Seismology working group, a 150-day long time series was generated by one of the authors (TT) with frequencies provided by another author (GB). The task of the data fitter (the 2 other authors: OM and TA) was to derive the frequencies of the detected modes.

Figure 4 shows the power spectrum of the time series. The distribution of the modes in the Echelle diagram in Fig. 4 does not permit to take any conclusion about any l but $l = 0$, as for higher orders it starts to be scrambled. As a matter of fact we realize a posteriori that the single ridge attributed to the $l = 0$ modes was mixed with the $l = 2$ modes (See Fig. 3). The detected modes fall in 2 categories:

- short-lived modes (l tagged or unknown l)
- long-lived modes (unknown l)

A specific strategy for each category is described hereafter.

4.1. Short-lived mode detection

The $l = 0$ modes were fitted as single modes using MLE. The other detected short-lived modes for which there was no possible l tagging were also fitted in the same manner.

4.2. Long-lived mode detection

As mentioned in the previous section, long-lived modes appear because of the mixed character of the modes. They can be seen in Fig. 4 as sharp peaks. There are two main cases to be considered:

- long-lived mode alone
- long-lived mode embedded in short-lived mode

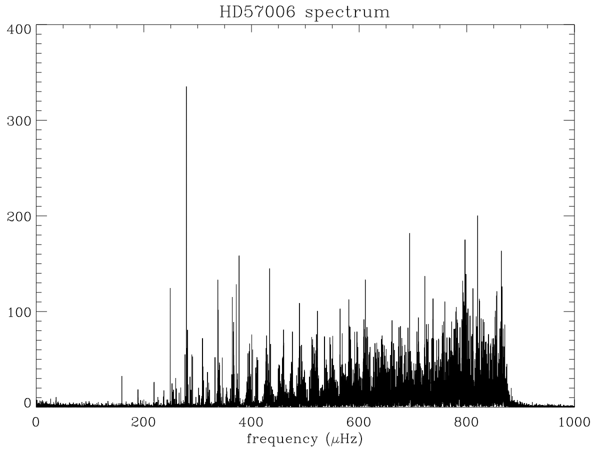
The first case happens at low order (low frequency below 300 μHz). The second case happens at higher frequency (typically order 13).

Narrow peaks alone: If one considers a pure noise signal with a χ^2 statistical distribution, the probability that the power within one bin is greater than m times the mean of the noise power, σ, is:

$$\mathcal{P}_N(m) \approx N e^{-m} \quad (1)$$

By setting a given value for $\mathcal{P}_N(m)$, for instance 10% (which means 10% probability that a peak due to noise is above m), choosing a window range in our spectrum that contains N bins, and estimating σ, one can derive using the equation (1) the correct value for m. This way, we have a statistical test for detecting the peaks that can be considered as having a low probability of being due to noise. This classical test was used by Appourchaux et al. (2000) for detecting long-lived p modes and g modes in the SOHO data.

Narrow peaks mixed with short-lived modes: Above 300 μHz the peaks that we wanted to analyze are among broad modes (See Fig. 4). Therefore, the application of the aforementioned test directly to the power spectrum is not very useful because it detects peaks that are part of a single broad mode. In this case we cannot assume that the detected peaks are all individual modes. Furthermore in some cases the detected peaks are very close to the broad modes; it becomes impossible to determine reliably if they are sharp modes or if they are just part of the broad mode.



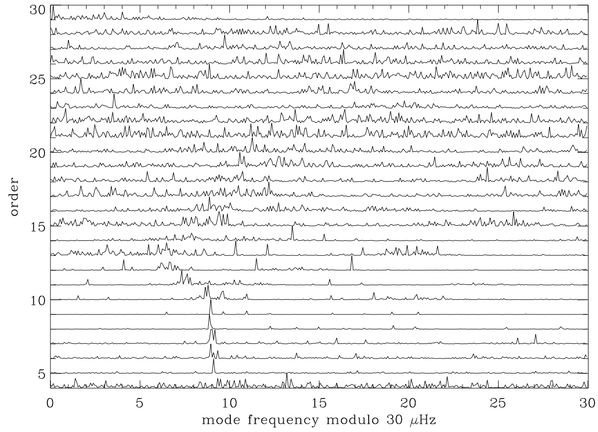


Figure 4. Left panel: the spectrum of the star HD 57006 for model VIII. Right panel: the corresponding Echelle diagram with $\Delta\nu_0 = 30\,\mu Hz$. The many peaks visible for higher orders are due to the larger linewidth of the modes at higher frequencies.

We need to find a test that can distinguish the broad modes from the sharp modes.

In order to solve this problem, we devised a technique that:

1. Fits the short-lived modes using MLE
2. Corrects the spectrum for the fitted model
3. Applies the aforementioned test as if we had only narrow peaks

Step 1: Assuming that p modes are stochastically excited oscillators, one can derive that the power spectrum of p mode oscillators is distributed around a mean Lorentzian profiles with a χ^2 probability distribution (Anderson et al. 1990; Appourchaux et al. 1998). The power spectrum of the p modes can be described as:

$$P(\nu) = M(\nu)F(\nu), \tag{2}$$

where $F(\nu)$ is a random function with a χ^2 statistical distribution with 2 d.o.f, and $M(\nu)$ is the model of the fitted mode made of a single Lorentzian profile plus noise. One can fit this model to the observed power spectra using the Maximum Likelihood Estimators technique (Anderson et al. 1990; Toutain and Appourchaux 1994); this is the classic and well-known approach used for short-lived modes.

Step 2: After having done the fitting, one can divide the power spectrum by the fitted profile $M'(\nu)$, obtaining:

$$P'(\nu) = \frac{P(\nu)}{M'(\nu)} \sim F(\nu) \tag{3}$$

In a first approximation $P'(\nu)$ has a χ^2 statistical distribution with 2 d.o.f. This is an approximation because $M'(\nu)$ is derived from data and has also a statistical distribution that should be taken into account. We have performed a Monte-Carlo analysis confirming that indeed $P'(\nu)$ has such a statistical distribution. This way we solved the problem of the mixing between the sharp modes and the broad modes.

Step 3: Applying the χ^2 test to HD 57006 spectrum, we set $\mathcal{P}_N(m) = 10\%$, $\sigma = 1$, and a window size of 30 μHz (corresponding to 389 bins). This last step is the same as for the long-lived modes alone.

For each window we use the stepped approach. The result of this can be seen in Fig. 5.

The comparison with the original frequencies[5] showed that only about 1 peak was misidentified over the band 200–500μHz. On average we should statistically have had 1 peak (±1) due to noise. It validates the approach taken for detecting long-lived modes either isolated or embedded in short-lived modes.

4.3. l TAGGING OF THE FREQUENCIES

The l tagging of the frequencies of the modes could not be derived from the Echelle diagram except for the $l = 0$ modes. As a matter of fact, due to the width of these p modes, we may have fitted the mean location of the $l = 0 - 2$ ridge.

The l tagging for the non-$l = 0$ p modes would need to be derived from the splitting. After having derived the frequencies of the modes for the 2 categories, we performed a correlation analysis on the extracted frequencies for getting the signature of a possible splitting. Since we did not find any splitting signature, we could not further tag the identified modes. At least we were able to derive the frequency of the modes for the 2 categories as explained above.

Our inability to properly tag the frequencies of the modes has serious consequences for the usefulness of such a star for understanding its internal structure. We could

[5] Comparison performed a posteriori.

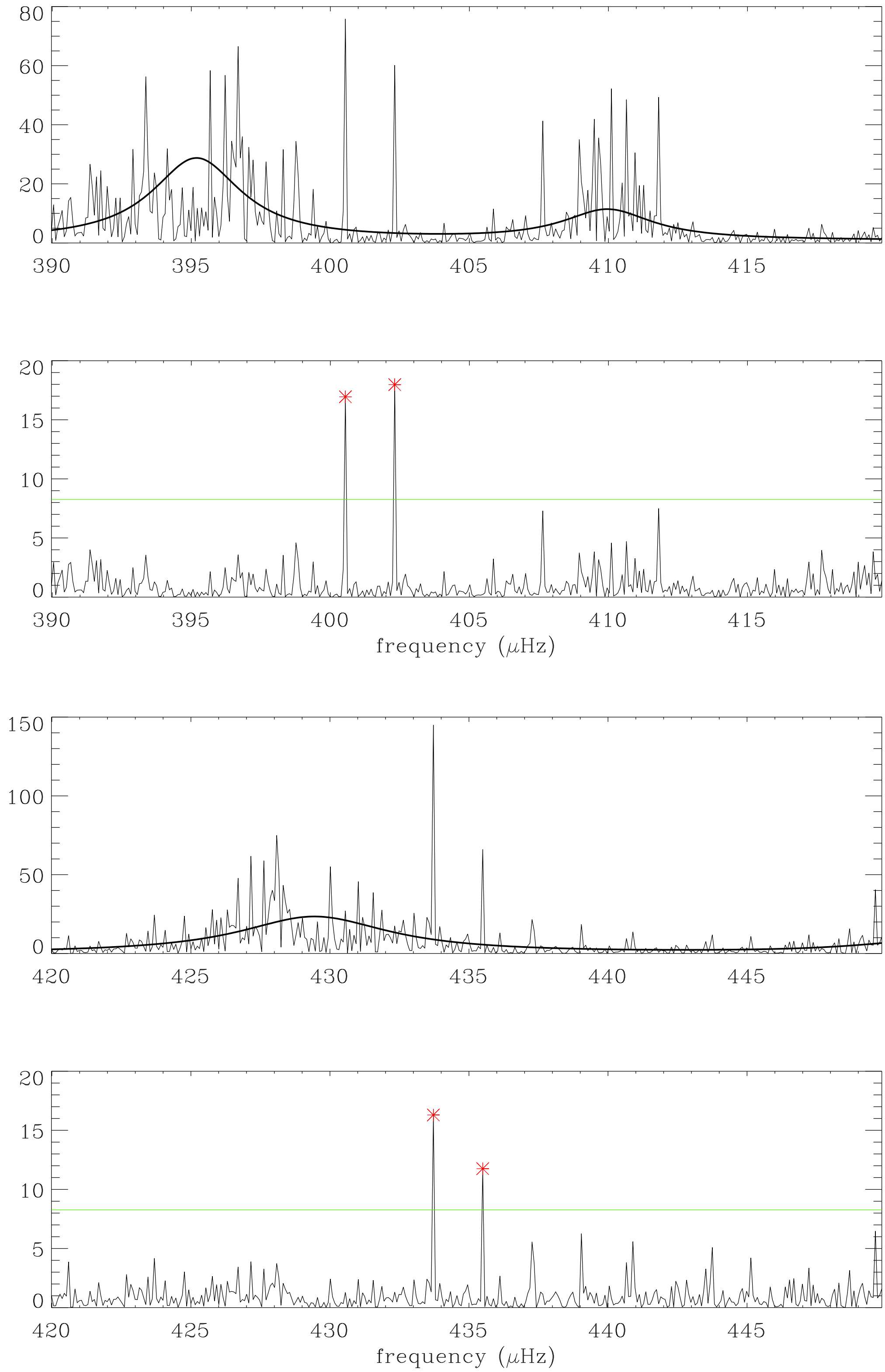


Figure 5. Result of the χ^2 test on the power spectrum discussed in Sect. 4.2 for two different frequency ranges. The first and third panels from the top show the fits, while the second and fourth panel show the power spectrum divided by the fitted model, as well as the bins above the 10% probability level.

imagine a scheme where we could search for the stellar model for which mode frequencies would match the fitted mode frequencies. Given the erratic behaviour of the $l = 1$ mode frequencies, this exercise could be 'easily' achieved when comparing input frequencies and output frequencies of a stellar model coming out of a single routine (e.g. the CESAM code). This is only of pure academic interests as the stars do not follow the CESAM code...

The only way out would be to use color information for tagging the frequencies of the modes. This method is not used for solar-like stars but is very useful for, say, Cepheids. It will involve proper computing of the radiative transfer which will likely not to be easy. Nevertheless, a simple approximation to the problem is well known Toutain and Gouttebroze (1993). This additional color information is likely to be useful for COROT, and should become a must for *Eddington*.

5. Conclusion

In this paper, we concentrated mainly on the fitting challenges that the COROT and *Eddington* missions are going to face. For each stage of the stellar evolution there are two challenges to be solved:

- Challenge I: power spectra fitting for solar-like stars in their hydrogen burning phase
- Challenge II: power spectra fitting for solar-like stars in their hydrogen-shell burning phase

Challenge I provides no major problem for fitting most of star spectra; this can even be automated.

Challenge II is mainly related to the proper identification of the degree of the modes. This challenge is much more difficult and could cost a lot of time as it is only solvable by hand at the time of writing. For this purpose, we devised an automated method for identifying long-lived modes embedded in short-lived modes. So we have been able to solve part of challenge II. The bulk of this challenge may only be solved by covering a given evolutionary track with stars of similar masses; thereby trying to follow with different stars how the mixed modes could appear. Additional information related to color could be brought in for the degree identification. This could make the automation feasible. The automation of the mode identification and fitting for challenge II is still in its infancy but 2 research tracks have been laid out. They shall be tested with COROT and perfected for *Eddington*.

Acknowledgements

This paper was kindly presented by Jesper Schou, whose courtesy the authors acknowledge. This work benefited from the collaboration with the Data Analysis Team of the Seismology Working Group of COROT. Thanks to Ian Roxburgh for interesting discussions, and to Coralie Neiner for commenting on the manuscript.

References

Anderson E. R., Duvall T. L., Jefferies S. M. 1990, ApJ 364, 699

Appourchaux T. 2003, Ap&SS 284, 109

Appourchaux T., Fröhlich C., Andersen B. et al. 2000, ApJ 538, 401

Appourchaux T., Gizon L., Rabello-Soares M. C. 1998, Ap&SS 132, 107

Gelly B., Grec G., Fossat E. 1986, A&A 164, 383

Gelly B., Lazrek M., Grec G. et al. 2002, A&A 394, 285

Jiménez A., Roca Cortés T., Jiménez-Reyes S. J. 2002, Sol. Phys.209 247

Neiner C., Apporchaux T. 2003, this volume.

Tassoul M. 1980. ApJS 43, 469

Toutain T., Appourchaux T. 1994, A&A 289, 649

Toutain T., Gouttebroze P. 1993, A&A 268, 309