

## Discussion document: A peak-bagging package for Kepler

W. J. Chaplin Dec 16 2007

### Abstract

It was agreed at the First Kepler Asteroseismic Science Consortium (KASC) Workshop that we – being interested members of the asteroFLAG group – would provide a peak-bagging package for analysis of the Kepler photometry data on Sun-like oscillators. This package will be used to extract peak-bagging estimates of the principle p-mode parameters. Here, I raise in brief discussion some of the important issues we will need to consider in putting this together. The main challenge will, I believe, be one of introducing automation in the procedures, and of trusting the outputs that are thereby generated. To get round this problem, I propose an interactive strategy, involving “quicklook” data products, which may be examined quickly to check procedures are generating sensible output, and to enable choices to be made where this is not the case, on a timescale that is practicable.

This document hopefully has some usable ideas within it; and no doubt also some bad ones. The aim is to provide a document off which we may hang discussion on how to proceed.

## 1 Introduction

The Danish AsteroSeismology Centre (DASC), which is co-ordinating the KASC activities, will provide a core pipeline for analysis of the p-mode data. This is the Kepler Asteroseismic Science Operations Centre (KASOC) pipeline (Christensen-Dalsgaard et al. 2007; Kjeldsen et al. 2007). The KASOC pipeline will include packages for all stages of analysis, from initial extraction of mode parameters through to using the extracted parameters to make inference on the properties and structures of the stars.

Of interest to us, the packages will include ones for identification (“tagging”) of modes, and for p-mode parameter extraction. The methods that will be used to extract the parameters are based on methods developed for analysis of ground-based data on Sun-like stars (e.g., see Kjeldsen et al. 2005, 2007).

We have offered to provide a pipeline for p-mode parameter extraction, based on the “peak bagging” approach applied to Sun-as-a-star data and to the CoRoT data on Sun-like stars (Appourchaux et al. 2006a, b)

The first release of data to the KASC will comprise timeseries data on 512 short-cadence targets, with those timeseries being between 2 and 3 months in length. Not all, but the vast majority, of these targets will be Sun-like oscillators. In turn, target selection should have ensured that most of the Sun-like oscillators will be main sequence or subgiant stars. These are the targets that are of interest to us.

Our first objective should be to provide a pipeline to extract p-mode parameters from timeseries on Sun-like stars of length 2 to 3 months. This should be the baseline criterion for the first package(s) we deliver, because the first data we will be confronted with will have these lengths. As time passes, and some of the stars are monitored continually, timeseries of length 1 year and longer will come available for analysis. The packages should be flexible enough to cope with timeseries analysis of

longer datasets.

## 2 Procedural Steps

There are several stages of processing needed to extract estimates of the mode parameters:

1. Mode detection: identify prominent features (peaks and spikes) in the (power or amplitude) frequency spectrum that are candidate modes;
2. Tag the angular degrees, measure dominant frequency spacings, determination of S/N in modes;
3. Generate set of first-guess parameters for mode peaks;
4. Decide on strategy for peak-bagging implementation (or strategies for more than several stages or types of fit);
5. Fit the mode spectrum to estimate the mode parameters;
6. Quality check on results before they are “passed as o.k.” for further science processing.

Between us, we will hopefully be able to provide all steps above, to give a complete complementary pipeline for extraction of the mode parameters. Discussion in this document concentrates on steps 3 through 6 inclusive, which the peak-bagging package will need to accomplish.

The peak-bagging package will need to be configured to accept as inputs estimates of frequency,  $\nu$ , and angular degree,  $l$ , for each candidate mode in the power (amplitude) frequency spectrum. It will therefore need to interface with elements of the KASOC pipeline, which accomplish steps 1 and 2 above, and alternative packages, provided by folk like us, which are also designed to accomplish steps 1 and 2.

The peak-bagging algorithms will then have to make decisions on fitting strategy, or strategies, based on the information presented to them. This is one stage where I believe the proposed “quicklook” strategy will be useful. By “quicklook”, I mean easily digestible graphical outputs of results that summarize the salient points regarding quality and robustness of results, and any automated decisions made concerning rejection of poor data<sup>1</sup>

So, the idea would be to present easy-to-digest graphical and data outputs on each target star. At this stage – before the fitting is performed – the graphical outputs could include plots of the echelle diagram, folded appropriately at the estimated large frequency spacing; plots of first-guess frequencies for each  $l$  as a function of frequency or overtone number (where the overtone numbers for each  $l$  may still have arbitrary offsets at this point); and plots of large and small frequency spacings as a function of frequency or overtone number.

First-guess parameter data will have already been prepared at this point for input to the peak-bagging algorithm. However, inspection of the quicklook products will allow a trivial flagging of those targets where we believe the fitting may have problems, e.g., for stars with high rotational splittings, or large peak linewidths, giving rise to potential overlap of components in  $l = 2/0$  pairs; or the presence of mixed modes giving similar, component-blending-related problems; or the first-guess frequency estimates may look far-from smooth as a function of frequency, yet we may not be expecting mixed modes.

With regard to estimation of first-guess parameters such as the peak heights, peak widths and mode frequency splittings, it behoves not to have too complicated a procedure, and too accurate a set of first guesses. This is because we want to avoid premature convergence on the first-guess values

---

<sup>1</sup>The term “quicklook” usually refers to “quick” and straightforward analysis tools that allow a “quick look” at subsets of what may be large volumes of data to help make decisions regarding further, higher-level processing. Here, we instead mean a “quick look” at the outputs of the fitting and quality measures associated with those outputs.

(which may of course fool us into thinking we have robust parameter estimates, when we do not). It may well be worth building into the automation repetition of the fitting with randomly tweaked first-guess parameters, to test robustness of the convergence (see Section 3 below).

While the quicklook products will anyway flag potential problem cases, it may be useful to have an *a priori* guide to indicate which parts of the HR diagram are likely to throw up particular complications for parameter extraction. When a target is initially selected for parameter extraction, warning flags could therefore have already been set, based on the position of the star in the HR diagram, and initial estimates of the frequency spacings of its spectrum. Of particular importance will be flagging regions where we might expect stars to have evolved sufficiently to display mixed modes. Stars that “just” have mixed modes may be the trickiest to deal with. If we use the inertias of such mixed modes as guides to the amounts by which the amplitudes may be suppressed, compared to the neighbouring p modes, we find the answer is: not by much. So we could have prominent peaks that the fitting procedures cannot ignore, but whose locations may give problems (e.g., component overlap in frequency) for the fitting itself. I am already working on issues related to such an *a priori* guide.

Clearly, we need to use the asteroFLAG and Aarhus artificial timeseries data to test some of the difficult scenarios. Experience with real data tells us there will be problem stars, where the procedures we have throw up non-sensible results. The aim of the parameter extraction packages should be to deal as smoothly as possible with the other more straightforward targets, so we have more time to look at the problem cases individually. I am sure there will be no getting round our having to devote individual time to problem stars.

### 3 Some comments on fitting strategies

Some thoughts on the fitting strategies to be adopted:

- The CoRoT Data Analysis Team (DAT) [e.g., see Appourchaux et al. 2006a, b] has of course already gone through the process of developing a recipe for peak-bagging analysis of Sun-like stars observed by CoRoT. Timeseries lengths are up to 6 months or so, i.e., not too dissimilar to the initial requirements for Kepler data. The data quality (i.e., S/N in the modes) is also pretty similar (see Appendix A): so the CoRoT experience will be very relevant to analysis of Kepler timeseries of length up to 6 months or so. The CoRoT peak-bagging recipe may be a useful default option for the initial Kepler dataset releases (see Appourchaux et al., 2006a).
- One of the big problems will be estimation of the inclination. Without a well-constrained, and reliable, estimate of the inclination we will not be able to extract reliable estimates of the mode frequency splittings. Clearly, we need to use the artificial asteroFLAG data to test fitting for the inclination. There are also now several detailed studies in the literature, which are devoted to this topic (e.g., Gizon & Solanki 2003; Ballot, García & Lambert 2006; Ballot et al. 2008a, b). These demonstrate numerous pitfalls, e.g., a propensity for the fitting to “lock onto” an inclination of 90 degrees. Some problems can be partially alleviated by fitting neighbouring  $l = 0, 1$  and  $2$  modes together. However, we will probably find that, from the first few releases of Kepler data, there will be only a small number of stars for which we can estimate the inclination.
- With regard to the inclination problem, the suggestion is that we perform several fits, with different strategies, e.g., attempt to fit for inclination and splitting; and another for product of splitting and sine of inclination (see Ballot et al. references); perform a fit with a fixed inclination; and perhaps another where the non-radial modes are simply modelled as single peaks. We then sift through the multiple results at the “quality-control” stage (see comments after itemized list).

- We must bear in mind that mode visibility in the more massive targets may be such as to allow fitting of some  $l = 3$  modes. This will modify the fitting models.
- Peak widths of the non-radial modes should be tied to those of the  $l = 0$  modes as the default option, at least for the first shorter dataset releases. There are of course circumstances where this may be a poor approximation to make, i.e., when the surface activity concentrated in the active latitudes is significantly stronger than on the Sun. We also know that should such stars also have rapid stellar cycles, their mode peaks will likely be distorted compared to the expected Lorentzian form (see Chaplin et al. 2007b). This is something we will need to bear in mind, for while the distortion has little impact on extraction of the frequencies, it can result in the return of poor estimates for the heights and widths etc.
- We should expect mode peaks of the Sun-like stars to show asymmetry, like their solar counterparts. For each target, we should perform fits with asymmetry as a free parameter, and fits with the asymmetry fixed to zero. A judgement on the robustness of the asymmetric strategy results may then be made later, at the quality control stage, by applying the likelihood-ratio test.
- We know from fitting the Sun-as-a-star data that power from modes outside the target windows – the latter containing single pairs, or a pair of odd- $l$  and a pair of even- $l$  modes – can affect the fits, since power from these outlying modes is not usually included in the fitting models. This will be an even bigger issue for the higher-mass main sequence stars. This is because the large spacing is smaller in the higher-mass stars; and these stars will also show in the power frequency spectrum some of the highest mode-height-to-background ratios, and possibly also some of the highest peak linewidths (Chaplin et al. 2007a). We should therefore consider applying full-spectrum (e.g., see Appourchaux et al. 2003) and pseudo-full-spectrum (e.g., see Gelly et al. 2002; Fletcher et al. 2008) fitting codes to the Kepler data, which take account of “outlying” power. We might consider running fits with the usual approach, and also with the full-spectrum approach. We would then inspect results at the quality-control stage.
- As noted previously, we may also wish to build into the procedures the option to repeat fits with randomly tweaked first-guess parameters. This may anyway be useful when we have fits that fail to converge, but which may have done so – and returned useful results – had the first-guess parameters been slightly different.

How do we deal with all the choices above? One option is the following: We set up the automation to do multiple fitting of each timeseries, e.g., fits for inclination, fits with fixed inclination, fits with and without peak asymmetry etc. This way we cover all plausible options for the fitting. That way we do not need to have in place a complicated algorithm that would have to make intelligent decisions regarding what strategy is best for fitting the target in question.

This of course leaves us with multiple sets of parameters estimates from the different fittings. It is at the next stage of the processing, which we might term “quality control”, that we sift through the results to find out which fits worked, and which didn’t (in the sense of having stable convergence, and giving sensible results).

Clearly, we need to keep things manageable, so we don’t want too many fitting options. We must remember also that as datasets get progressively longer for some stars (i.e., as we stay on the same targets), fitting strategies will change, and we will be able to start fitting for the inclination, and looking at more subtle features such as asymmetric frequency splittings etc. With the suggested strategy there is no need to have to switch to a fitting different algorithm, or to have to decide which is “best”. We just continue to run all options, but are able to pick out reliable parameter data at the quality control stage.

We go on now to make some comments about the quality control procedures.

## 4 Quality control of raw fitted parameter data

The quality control procedures should involve some automated post-processing of the raw fitted parameters, and graphical output quicklook products for manual inspection. We need to be able to judge the quality of fits and the output parameters of particular fitting strategies, and of one strategy against another:

- The automated procedures should flag poor fits for rejection; compare quantitatively results of one strategy with another; and compute other useful indicators of fit quality. The outcomes of these procedures would then be summarized for visual inspection in the graphical quicklook outputs.
- When we add in extra free parameters to fit, for example peak asymmetry, we may very readily compute the significance of freeing up that extra parameter, by performing likelihood ratio tests on the fitting results (with and without the free parameter). The automated software could then make a choice, based on a pre-specified rejection threshold: if adding in the extra parameter does not give a significant improvement in the fit likelihood, the fit with the free parameter should be flagged as being poorer than its counterpart. This information would be summarized visually in the quicklook outputs.
- We should record, for each fit, belt-and-braces measures of the quality of the fit, e.g., the likelihood at convergence and the number of iterations to converge; the ratio of the estimated variance of the data to the parent variance; and we could test for the presence of statistically significant spikes or features in the fitting residuals (i.e., after division of the raw spectrum by the best-fitting model). Measures like these may again be included in graphical outputs and summary information.
- We may look at the “smoothness” of the parameters of a given  $l$  against frequency or overtone number – e.g., using first or second differences – as a means of judging when fits give estimates of parameters that lie significantly beyond the underlying frequency trends. (We of course need to apply such procedures with caution: some stars may show sharp seismic features in frequency.)
- Problem stars will get flagged at this stage too (if we have let them through) as a result of giving poor quality outputs from the fitting (e.g., in worst-case scenarios we may get results from no fits at all!).

We will then need a set of guidelines to decide on which fitting data are passed as o.k. for subsequent higher-level analysis. The simplest choice may be to say that anything that gets through the objective, automated quality-control procedures is fit for release. Inspection of the quicklook products, where these choices will be summarized, will allow a sanity check that the automated procedures have done what they are supposed to do.

For some stars there may be multiple datasets that are passed through. Provided we are clear on what fitting strategy was applied to extract each set of data, data passed as o.k. could be assigned as appropriate for particular science problems, e.g., some datasets may be expected to have smaller bias in the excitation and damping parameters than others. One thing we do know from the Sun-as-a-star case is that the frequencies are pretty resilient to changes in strategy, which should keep things simple where this parameter is concerned.

## 5 Final comments

If we are to run with the proposed overall philosophy, the next stages in the development of the packages will involve:

Table 1: Scientist hours of effort for quicklook inspection

<i>Number of stars</i>	<i>Average quicklook time per star (mins)</i>	<i>Scientist time required</i>		
		<i>Total (hours)</i>	<i>Between 3 (hours)</i>	<i>Between 5 (hours)</i>
512	10	85	28	17
512	20	171	57	34
400	10	67	22	13
400	20	133	44	27
250	10	42	14	8
250	20	83	28	17
150	10	25	8	5
150	20	50	17	10

- Testing fitting strategies with the artificial asteroFLAG data;
- Making decisions on what fitting strategies to implement as a matter of course;
- Making decisions on quicklook products, at the pre- and post-fitting stages, and the automated post-processing procedures, data rejection rules etc.;
- What to do with problem stars, e.g., having been flagged, and allowed to run through the normal procedures, do we have a subsequent set of rules and procedures, or do we accept we may have to deal with each problem target on its own, individual merits?

Is this all practicable, from a scientist-effort perspective? Provided we are smart about what the quicklook outputs contain, and what guidelines we have in place for dealing with results, I think this is do-able. I think we would all agree that the alternative, which would be to completely trust the automated outputs, could be a recipe for creating problems. Table 1 gives some idea of manual inspection/vetting effort involved. The hardest job may come with the first release of data, where the vast majority of the 512 short-cadence targets will be tasked to what we think are Sun-like, main sequence stars. But not all these stars will have prominent enough modes in their power frequency spectra to allow fits to be made to individual modes.

## References

- [1] Appourchaux T., 2003, Ap&SS, 284, 109
- [2] Appourchaux T., Berthomieu G., Michel E. et al, 2006a, in The CoRoT Mission, A. Baglin, J. Lochard, M. Fridlund, L. Conroy eds., Publication ESA SP-1306, p. 429
- [3] Appourchaux T., Berthomieu G., Michel E. et al, 2006b, in The CoRoT Mission, A. Baglin, J. Lochard, M. Fridlund, L. Conroy eds. Publication ESA SP-1306, p. 377
- [4] Ballot J., García R. A., Lambert P., 2006, MNRAS, 369, 1281
- [5] Ballot J., Appourchaux T., Toutain T., Guittet M., 2008a, AN, submitted
- [6] Ballot J., Appourchaux T., Toutain T., Guittet M., 2008b, MNRAS, in preparation
- [7] Chaplin W. J., Elsworth Y., Houdek G., New R., 2007a, MNRAS, 377, 17

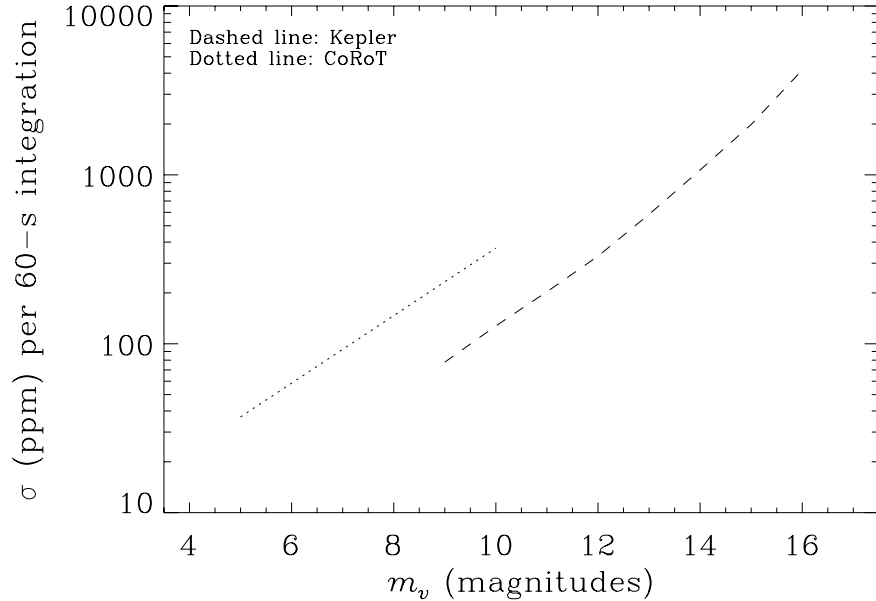


Figure 1: Expected noise levels, per 60-sec integration, for Kepler (dashed line) and CoRoT (dotted line). Data are plotted over the approximate apparent magnitude target ranges of both instruments.

- [8] Chaplin W. J., Elsworth Y., New R., Toutain T., 2007b, MNRAS, in the press
- [9] Christensen-Dalsgaard J., Arentoft T., Brown T. M., Gilliland R. L., Kjeldsen H., Borucki W. J., Koch D., 2007, CoAst, 150, 350
- [10] Fletcher S. T., Chaplin W. J., Elsworth Y., New R., 2008, AN, submitted
- [11] Gelly B., et al., 2002, A&A, 394, 285
- [12] Gizon L., Solanki S. K., 2003, ApJ, 589, 1009
- [13] Kjeldsen H., Christensen-Dalsgaard J., Gilliland R., Arentoft T., 2007, ‘The KASOC Pipeline’, DASC, University of Aarhus, Denmark
- [14] Michel E., Baglin A., Samadi R., Baudin F., Auvergne M., 2007, CoAst, 150, 341

## A Noise levels for Kepler and CoRoT

It turns out that expected noise levels for the apparent magnitude target ranges of CoRoT and Kepler have similar sizes. For Kepler, the default range covers approximately  $9 \leq m_v \leq 15$  (e.g., Christensen-Dalsgaard et al. 2007). The corresponding range for CoRoT is approximately  $5 \leq m_v \leq 9$ . Expected noise levels per 60-sec integration are plotted in Fig. 1 for the respective target ranges. The Kepler data, which are shown by the dashed line, come from calculations made by Borucki and colleagues (Kjeldsen, private communication). The CoRoT data, which are shown by the dotted line, come from photon shot-noise estimates given in Michel et al. (2007).

The plots show that when  $m_v \sim 6.6$  for CoRoT we should expect to get similar noise levels in Kepler when  $m_v \sim 9$ . This correspondence persists to the point where  $m_v \sim 9$  for CoRoT and  $m_v \sim 11.5$  for Kepler.