Astronomy & Astrophysics manuscript no. rn5 October 4, 2004
(DOI: will be inserted by hand later)

Research note

On detecting short-lived p modes in a stellar oscillation spectrum

T.Appourchaux

Institut d'Astrophysique Spatiale, Université Paris XI - C.N.R.S., Bâtiment 121, 91405, Orsay Cedex, France
e-mail: Thierry.Appourchaux@ias.u-psud.fr

Received 16 July 2004 / accepted 1 October 2004

Abstract. The false alarm probability for detecting peaks embedded in a power spectrum of noise was given by Scargle (1982). This test has been used in helioseismology to detect long-lived modes such as g modes (Appourchaux et al. 2000). With the development of asteroseismology, there is now a need to define a similar test but for short-lived p modes. In this article, I define a false alarm test for detecting short-lived p modes, and I give the probability of detecting such modes given their signal-to-noise ratio, their linewidths and the duration of observation.

Key words. Stars: oscillation – Methods: statistical

1. Introduction

The false alarm detection probability of a coherent periodic signal embedded in a power spectrum of noise was given by Scargle (1982). This test has been used in helioseismology to detect long-lived oscillation modes known as g modes (Appourchaux et al. 2000; Gabriel et al. 2002; Turck-Chièze et al. 2004). The probability as given by Scargle (1982) belongs to the H0 hypothesis class, i.e. what is observed is pure noise. The detection probability based on the H1 hypothesis has been given by Gabriel et al. (2002) based on assuming that a periodic signal is present with a given amplitude embedded in a power spectrum of noise. The H1 hypothesis requires that the inclusion of the periodic signal in the noise is subject to a given set of *assumptions*, e.g. how the signal (mode) excited is, or its amplitude. The detection probability derived from the H1 hypothesis does not test for the presence of the periodic signal but tests the probability that it can be detected given these *assumptions*.

The tests described above are now well understood and utilized when looking for long-lived modes or g modes. With the advent of asteroseismology, there is a growing need for a proper testing of what is going to be detected. Detection of solar-like p modes in α Cen A and B has been recently reported (Bouchy & Carrier 2002; Carrier & Bourban 2003). The modes detected are short-lived p modes for which the usual tests are irrelevant. Here I develop two tests based on the H0 and H1 hypothesis. The former will assess the probability that a peak of a given width is due to noise. The latter will give the detection probability of a p mode as a function of the linewidth, signal-to-noise ratio and length of observation.

2. Hypothesis testing

2.1. H0 hypothesis

The simplest way to detect short-lived p modes is to bin the power spectrum over a different number of bins. In the case of the H0 hypothesis, we assume that what is observed is pure noise, i.e. χ^2 with 2 degrees of freedom. When binning over p bins, the statistics becomes χ^2 with $2p$ degrees of freedom. The probability of having a peak above a given level s in the binned power spectrum is given by:

$$P(s' \geq s, p) = \int_s^{+\infty} \frac{1}{\Gamma(p)} \frac{u^{p-1}}{S^p} e^{-u/S} \mathrm{d}u \qquad (1)$$

where S is the mean of the power spectrum, p is the number of bins over which the spectrum is binned *without* normalization (See Appourchaux 2003), i.e. the mean in the binned power spectrum is Sp, $\Gamma(p)$ is the Gamma function, and u is a mute variable. It is interesting to compute how the binning affects the detection level. For a given confidence level, say 10%, we solve the following:

$$P(s' \geq s, p) = 10\% \qquad (2)$$

Figure 1 shows the normalized level s/p, derived from Eq. (2), for two confidence levels. Figure 1 shows that binning does lower the detection level, as expected.

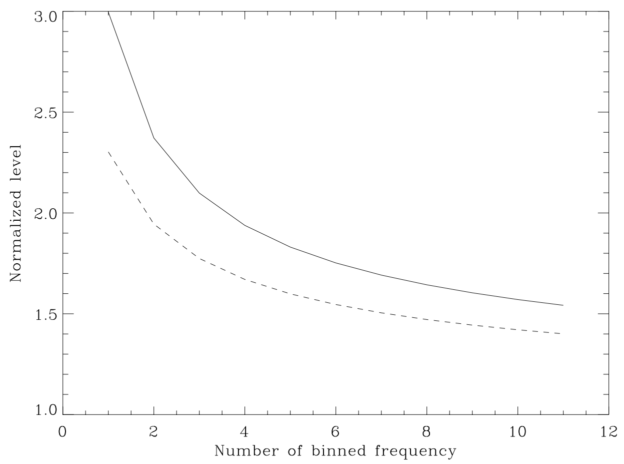

Fig. 1. Detection level as a function of the number of binned frequencies p, for a probability of 10% (dash line) and 5% (solid line). For $p=1$ we have the regular value of 2.3 σ and 3 σ values.

2.2. H1 hypothesis

Here I will calculate the probability that one can detect a mode given its characteristics (amplitude, linewidth), the background noise and the duration of observation. The detection is made by binning the power spectrum over p bins. The binning requires us to compute the statistic of the partial power of a mode profile. Hereafter, I provide the analytical derivation of that statistic and a more useful approximation that facilitates the computation of the probability in practice.

2.2.1. Analytical derivation

In this case we assume that the observed p modes can be described in the power spectrum as:

$$f(\nu) = \frac{r}{1+\frac{4(\nu-\nu_0)^2}{\Gamma^2}} + 1 \tag{3}$$

where r is the signal-to-noise ratio ($=A/B$) ratio of mode amplitude to background noise, Γ is the linewidth, ν is the frequency, ν_0 is the mode frequency. [1] Here we would like to assess the probability of detecting such a mode in the power spectrum when that spectrum is smoothed over p bins. The statistics of each bin is a χ^2 with 2 d.o.f with a mean given by Eq. (2) at $\nu=\nu_i$, $f_i = f(\nu_i)$. The statistic is computed using the characteristics function. That function for the partial power in a mode profile can be derived from Gabriel (1994) as:

$$F(v) = \prod_{i=1}^{i=p} \frac{1}{(1-vf_i)} \tag{4}$$

[1] Here I would like to point out that the concept of the signal-to-noise ratio was introduced by Libbrecht (1992) as stated above. It does not make sense to express the the signal-to-noise ratio as $(A+B)/B$ as one would naively think. It would mean that the signal-to-noise ratio would be 1 when there is no mode.

After decomposing that equation onto simple elements, the probability density function (pdf) of the partial power $\mathcal{S}$ is given as:

$$p(\mathcal{S}) = \sum_{i=1}^{i=p} \frac{e^{-\mathcal{S}/f_i}}{f_i} \frac{1}{\prod_{k\neq i}(1-f_k/f_i)} \tag{5}$$

Equation (5) becomes singular when the profile is summed symmetrically around ν_0. In this case, I calculate the characteristic function of the product of χ^2 with 4 d.o.f as given by:

$$F(v) = \prod_{i=1}^{i=p/2} \frac{1}{(1-vf_i)^2} \tag{6}$$

Here p is even by construction. We use the previous simple element decomposition to find that the pdf is given by:

$$\begin{aligned} p(\mathcal{S}) = & \sum_{i=1}^{i=p/2} \frac{\mathcal{S}e^{-\mathcal{S}/f_i}}{f_i^2} \frac{1}{\prod_{k\neq i}(1-f_k/f_i)^2} \\ & + \sum_{i=1}^{i=p/2} \frac{2e^{-\mathcal{S}/f_i}}{f_i} \frac{1}{\prod_{k\neq i}(1-f_k/f_i)} \\ & \cdot \sum_{j\neq i} \frac{1}{(1-f_j/f_i)} \frac{1}{\prod_{j\neq k}(1-f_k/f_j)} \end{aligned} \tag{7}$$

Unfortunately, Eqs. (5) and (7) become highly divergent when the smoothing is performed over a large number of bins, typically greater than 8. The divergence appears for very small values of $\mathcal{S}$ while it should give a value close to zero.

2.2.2. Approximation

I found that Eq. (5) (and thereby Eq. (7) could be approximated by a Gamma law given by:

$$p(\mathcal{S}) = \frac{\lambda^\nu}{\Gamma(\nu)} \mathcal{S}^{\nu-1} e^{-\lambda\mathcal{S}} \tag{8}$$

The mean and σ are given by:

$$\begin{aligned} E[\mathcal{S}] &= \frac{\nu}{\lambda} \\ \sigma^2 &= \frac{\nu}{\lambda^2} \end{aligned} \tag{9}$$

λ and ν are then derived from the mean and σ as:

$$\begin{aligned} \lambda &= \frac{E[\mathcal{S}]}{\sigma^2} \\ \nu &= \frac{E[\mathcal{S}]^2}{\sigma^2} \end{aligned} \tag{10}$$

In our case the mean $E[\mathcal{S}]$ and σ are given by:

$$\begin{aligned} E[\mathcal{S}] &= \sum_{i=1}^{i=p} f_i \\ \sigma &= \sqrt{\sum_{i=1}^{i=p} f_i^2} \end{aligned} \tag{11}$$

2.2.3. Detection probability

The detection probability is then given by:

$$P(s' \geq s) = \int_{s}^{+\infty} p(u)\mathrm{d}u \qquad (12)$$

where $p(v)$ is given by Eq (5) or (7) for the analytical derivation and by Eq (8) for the approximation, and the cutting level s is given by Eq (1) for a predetermined confidence level (e.g. 5%).

Given the numerical difficulties encountered with the analytical formulation, I checked that the approximation given above provides accurate results when compared with the analytical formula, even for low binning values. Figure 2 gives the relative probability difference between the analytical formula and the approximation for an observing time of 100 days and a linewidth of 1 μHz. The largest differences occur for low signal-to-noise ratio and large binning values. These differences are due to the divergence of the analytical formula, and not the result of the approximation used. In most cases, the relative error on the probability will not be larger than 0.001.

Figure 3 gives the maximum detection probability using the density probability given by Eq. (8) with λ and ν derived from Eq. (10) and (11). Of course, the maximum detection probability is reached when smoothing symmetrically over a mode profile.

3. Discussion

The results provided by the analytical calculation and by the approximation are quite consistent with expectation. The detection probability increases with the number of smoothing bins until the number of bins becomes larger than the linewidth of the mode (visible in the upper left diagram of Fig. 3). For a given linewidth, the detection probability will increase with the duration of observation. The detection probability only depends upon the number of bins in a profile. Therefore, the probability only depends upon the product of the linewidth and the observation time, i.e. the probability is the same for a linewidth of 0.5 μHz and an observation time of 200 days as for a linewidth of 1 μHz and an observation of 100 days.

The detection probability ranges from 20% for narrow modes up to 80% for larger modes for a signal-to-noise ratio of 1. A typical 50% figure is a good order of magnitude for the mean detection probability. This means that over 10 detectable modes, an average of 5 will detected.

4. Conclusion

I showed that one can easily compute the maximum detection probability of a p mode using Eq. (8) with λ and ν derived from Eqs. (10) and (11). This detection probability can be used in asteroseismology by the future space mission COROT or by the ground-based spectrometer HARPS.

Acknowledgements. This article is the result of a casual discussion between Eric Michel and the author. Thanks to Patrick Boumier for pointing out the inconsistencies of an earlier version. I am grateful to the referee for making constructive comments.

References

Appourchaux, T. 2003, A&A, 412, 903
Appourchaux, T., Fröhlich, C., Andersen, B., et al. 2000, ApJ, **538**, 401
Bouchy, F. & Carrier, F. 2002, A&A, **390**, 205
Carrier, F. & Bourban, G. 2003, A&A, **406**, L23
Gabriel, A. H., Baudin, F., Boumier, P., et al. 2002, A&A, 390, 1119
Gabriel, M. 1994, A&A, **287**, 685
Libbrecht, K. G. 1992, ApJ, **387**, 712
Scargle, J. D. 1982, ApJ, 263, 835
Turck-Chièze, S., García, R. A., Couvidat, S., et al. 2004, ApJ, 604, 455

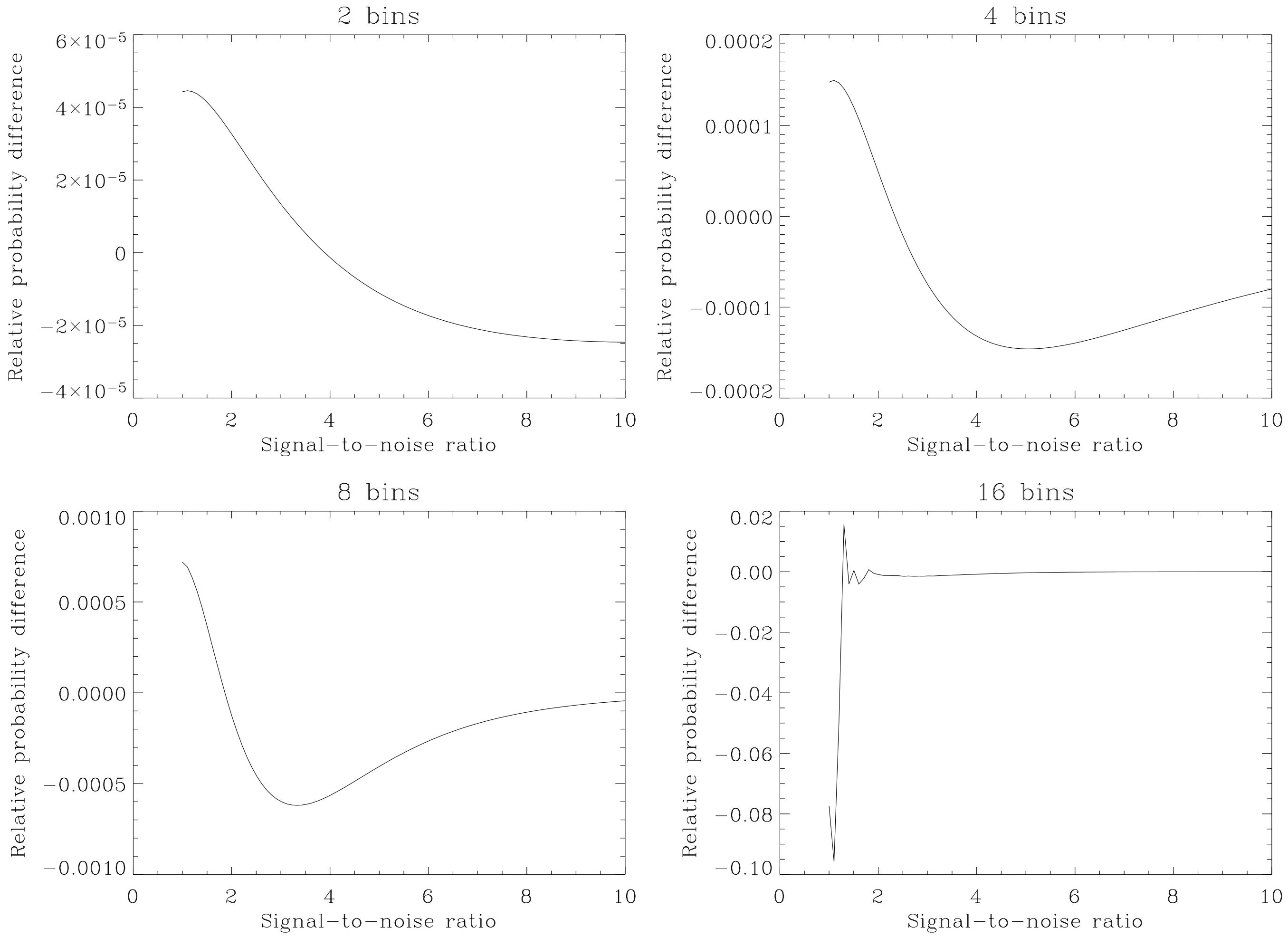

Fig. 2. Relative probability difference (analytical-approximation) as a function of the signal-to-noise ratio for various numbers of smoothing bins for a 100-day observing time and a 1-μHz linewidth. The predetermined confidence level used for computing the cutting level s is 5%. The probability density are given by Eq (5) or (7) (analytical) and by Eq. (8) (approximation).

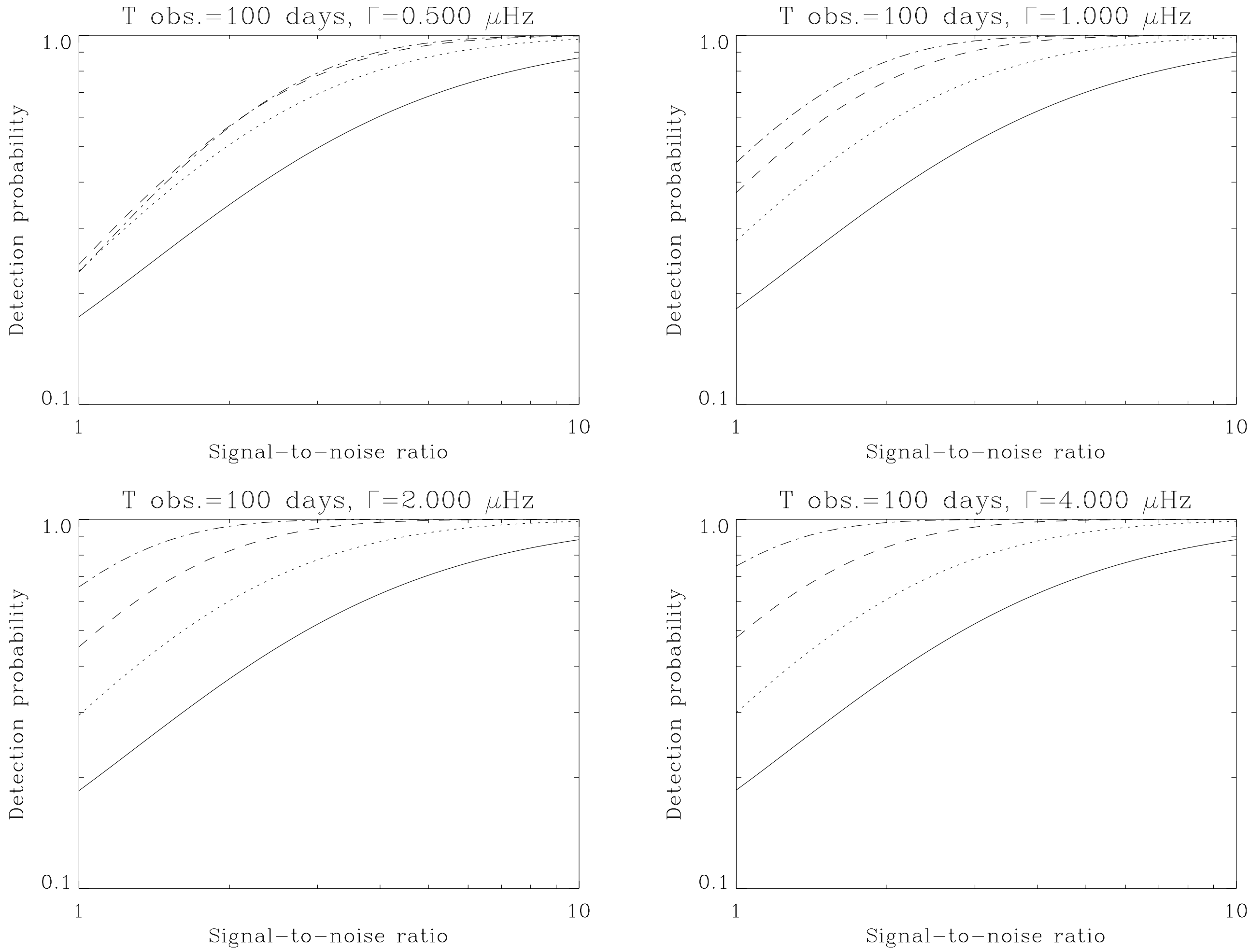

Fig. 3. Maximum detection probability as a function of the signal-to-noise ratio for various numbers of smoothing bins (2, continuous line; 4 dotted line; 8 dashed line; 16 dashed-dotted line) for various linewidths and observing times. The predetermined confidence level used for computing the cutting level s is 5%. The probability density is an approximation as given by Eq. (8).