A&A 424, 301–309 (2004)
DOI: 10.1051/0004-6361:20040372

Astronomy & Astrophysics

Impact of observational duty cycle on the measurement of low-ℓ solar p-mode frequencies⋆

W. J. Chaplin[1], Y. Elsworth[1], G. R. Isaak[1], B. A. Miller[1], R. New[2], and D. Salabert[1]

[1] School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
e-mail: salabert@bison.ph.bham.ac.uk
[2] School of Science and Mathematics, Sheffield Hallam University, Sheffield S1 1WB, UK

Received 2 March 2004 / Accepted 21 May 2004

Abstract. We investigate the impact of a multi-site, ground-based observational window function on the extraction of estimates of the frequencies of low-angular-degree (low-ℓ) solar p modes from decade-long datasets. To effect this study we have made use of some ≈10 yr of full-disc, "Sun-as-a-star" Doppler velocity data collected by the Birmingham-Solar Oscillations Network (BiSON). A coherent combination of observations made by all six BiSON sites provided the principal time series of data. This set was then modulated by a whole series of different window functions and the resulting sets analyzed. The windows were made from different combinations and numbers of BiSON stations. We find that for the majority of low-ℓ modes the bias in the frequencies given by the effects of the window function is not significant. However, for modes above ≈3300 μHz, and some $\ell = 2$ modes near ≈2500 μHz, the bias is important when the frequencies are extracted from long datasets.

Key words. methods: data analysis – Sun: helioseismology

1. Introduction

Studies of periodic and quasi-periodic astronomical phenomena can often benefit greatly from the availability of continuous observations, an obvious category being the study of various types of stellar oscillations. Here, we concentrate upon the demands offered by a single – but conspicuous – object, this being the Sun and its resonant p-mode oscillations. We do so for a ground-based observing strategy. Continuous, multi-decadal coverage of the Sun is desirable in order to investigate the 11-yr Schwabe activity cycle and to allow the accurate measurement of the frequencies and splittings of weak, but very long-lived, low-frequency modes. These serve to provide a stern test of stellar evolution theory.

If observations are made of the Sun from a single, mid-latitude, terrestrial site they will have long interruptions from the diurnal, day-night cycle. Quasi-random drop-outs in coverage will also arise from poor weather and occasional technical problems. Both effects serve to complicate analyses performed on the data in the frequency domain: quasi-random gaps increase the background level of "noise" across a range of frequencies; while the diurnal breaks create periodic structures, or sidebands, in the power spectrum.

One strategy for attempting to remove these effects is to observe from one of the terrestrial poles. However, this can present major logistical difficulties and regardless: continuity of coverage is limited at each pole to a maximum of approximately six months; and there is no redundancy to cope with poor weather at the observing site.

A more practical, long-term approach that offers the potential for multi-year continuity and the added advantage of some redundancy (through overlaps in coverage) is to deploy a network of observatories distributed widely in longitude. Two current examples that observe globally coherent solar eigenmodes are the Birmingham Solar-Oscillations Network (BiSON; Chaplin et al. 1996) and the Global Oscillations Network Group (GONG; Harvey et al. 1996). Even though each comprises six, well-separated sites, this is insufficient to give complete, 100-% coverage all year round. Recent annual duty cycles of useful data for both networks have typically been in the 70 to mid-80% range. While the artefacts are suppressed at these high levels of coverage, they nevertheless persist in the frequency domain at a level that cannot be ignored. Furthermore, an extended database of BiSON observations is available stretching back to the mid 1970s, but with somewhat reduced temporal coverage (i.e., because of limited single- or double-site summer campaigns, or latterly fewer permanent stations prior to the complete roll-out of the network). Here, the artefacts are rather more severe.

Our aim in this paper is to present a detailed study of the impact of a wide range of ground-based observational window functions on estimates of low-angular-degree (low-ℓ)

⋆ Appendix A is only available in electronic form at http://www.edpsciences.org

solar p-mode eigenfrequencies. We do so for data collected by Doppler velocity or intensity observations of the unresolved Sun, i.e., the "full-disc" or Sun-as-a-star BiSON data. Previous work in this area has concentrated on studies of the precision of the solar p-mode parameters (Toutain & Appourchaux 1994; Fierry-Fraillon et al. 1998), and on the evolution for different timescales of the measurement precision of the frequencies (Chaplin et al. 2002). Lazrek & Hill (1993) presented a first attempt to simulate the effects of the gaps on the measurement of oscillation parameters, but no one has investigated in detail the effects of temporal windows with different fractional fills on the measurement of p-mode frequencies.

A study of this type is important for several reasons. Ground-based networks making Sun-as-a-star observations – such as BiSON – have provided invaluable low-ℓ "core penetrating" mode data since the inception of the field, and will continue to underpin observations in this area over the coming solar cycle. As noted above, an historical low-ℓ database is available thanks to BiSON, but with reduced temporal coverage in its earlier years. We are currently in the process of recalibrating these older data by taking advantage of improvements in data preparation that have been made in recent years. Our intention will then be to study in detail the behaviour of the low-ℓ modes over an extended period of more than 25 yr, spanning Schwabe activity cycles 21 through 23. However, proper allowance will need to be made for the impact of reduced temporal fill (i.e., lower duty cycles) on the extraction of the mode parameters. Here, we investigate the effect of window functions with fractional fills that, while well below those encountered in post-1990 BiSON time series, will be typical of an extended 25-yr or pre-1990 set. While some gap-filling techniques appropriate to helioseismic data (i.e., many modes closely spaced in frequency) may prove useful (Fossat et al. 1999; Fierry Fraillon & Appourchaux 2001) we remain cautious regarding any advantage they may offer under conditions of very modest fill (e.g., less than 50 per cent). We leave a discussion of the impact of this to another paper.

A study of this type is also important in a stellar context. First, ground-based observations are sure to play an important rôle in observations of low-ℓ p modes on solar-like stars (Bedding & Kjeldson 2003), and as such similar analysis problems and issues will have to be confronted once extended time series of asteroseismic observations are commonplace. Second, having an accurate means of assessing the impact of different window functions on the precision with which the eigenfrequencies can be determined has important implications for studies of the diagnostic potential of the low-ℓ modes, e.g., for inversions of low-ℓ only data (Roxburgh & Vorontsov 2002; Basu 2003), or determining the near-surface properties of a star (Roxburgh & Vorontsov 2001; Verner et al. 2004).

2. Data and analysis

2.1. Data and window-function construction

Our study takes as its basis the analysis of data collected by the six-station, ground-based Birmingham Solar-Oscillations Network (BiSON) over the 3660-d period beginning 1991 January 1. This extensive epoch spans the declining phase of solar activity cycle 22, and the rising phase of cycle 23.

The instruments at each BiSON site make Sun-as-a-star observations of the Doppler shift of the potassium Fraunhofer line at 770 nm (e.g., Chaplin et al. 1996). Raw data were first processed in the manner described by Elsworth et al. (1995) to yield daily calibrated velocity residuals. We then combined coherently the resulting $\approx 2 \times 10^4$ individual daily sets from all six stations to yield the principal $\approx$10-yr time series of residuals, $v(t)$. The duty cycle (fractional fill) of this combined set was $\sim$74 per cent, with breaks in coverage largely the result of inclement weather.

In order to study the impact on the analysis of duty cycles lower than this, the principal series was modulated by a selection of window functions which were made from different numbers of stations.

To do this, we used time series constructed from coherent combinations of data collected by six (i.e., the principal series), five, four and three stations having duty cycles of 74, 68, 62 and 50-per-cent respectively, taking proper account of the longitudinal distribution of the stations. The strategy applied here for the window-function construction is very similar to that used by Chaplin et al. (2003a) to measure bias in the p-mode linewidths and heights caused by a ground-based window. Each set was first converted into a two-state format, $w(t)$, i.e.,

$$w(t) = \begin{cases} 0 \text{ for } v(t) = 0 \\ 1 \text{ for } v(t) \neq 0. \end{cases} \quad (1)$$

The four resulting "window" series were then divided into 88 partially overlapping (by 40 d) sub-series of length 122 d. Random 30-piece combinations of all available segments were then used to build 280 new 3660-d window functions, $w_i'(t)$ (with the index i running from 1 to 280). The multiplication of the principal velocity time series, $v(t)$ by each of these then gave final sets of data for analysis:

$$v_i'(t) = v(t) \cdot w_i'(t). \quad (2)$$

These had fractional duty cycles spanning the range $0.35 \leq D \leq 0.60$ (as shown in Fig. 1). Even though, strictly speaking, these windows are not fully independent, the effect of this on the results (e.g., error bars on the fits) is actually fairly minimal.

The main aim of our study is to assess the impact of these window functions on extracted estimates of the mode frequencies. Since we make use of real, and not artificial, observations we must be sure that frequency shifts introduced by the solar activity cycle do not bias the results. Some of the lowest-fill sets could, in principle, be preferentially weighted to different parts of the cycle and therefore different frequency levels. To check this we made five full sets of two-hundred and eighty 3660-d window functions. (The data used in the paper are from a single, 280-window set.) The 10.7-cm solar radio flux[1] was used as a proxy of the global level of activity, and we computed an

[1] The 10.7-cm solar radio flux used in this analysis was downloaded through the website `http://spidr.ngdc.noaa.gov/spidr/`

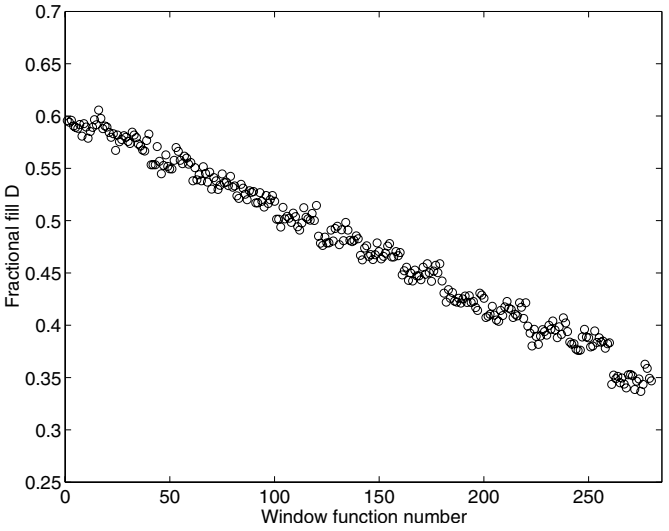

Fig. 1. Fractional fill D of the 280 BiSON window functions. Each was made from the random combination of shorter pieces taken from four ≈10-yr series; these were in turn constructed from coherent combinations of data collected from different numbers of BiSON sites (ranging from three to six).

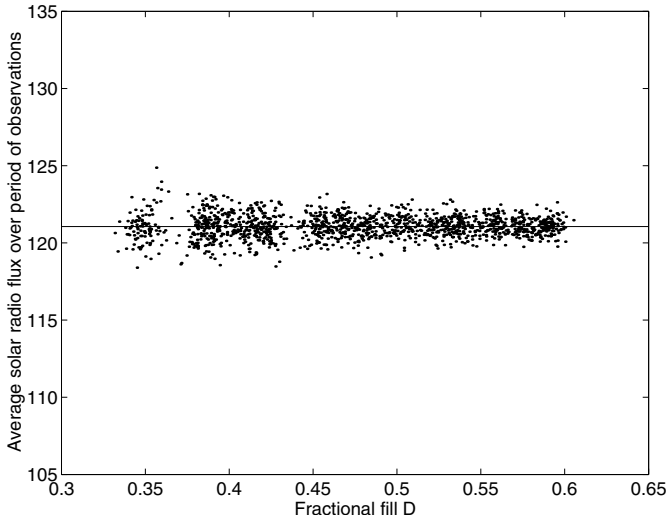

Fig. 2. Average 10.7-cm solar radio flux (in units of 10^{-22} W m^{-2} Hz^{-1}) over the period of observations, for 5 series of 280 simulated time-series as a function of the corresponding fractional fill D. The solid line represents the mean value over all windows.

average of this over the period of each window when non-zero data were present.

Figure 2 shows the mean flux value over all five series of window functions, plotted against the corresponding fractional fill, D. The average value, represented by the solid line, is slightly higher than 121 radio flux units. As can be seen, the trend is flat with no significant evidence for the introduction of any bias. The scatter on the average flux level does increase at lower D, reflecting the larger variance expected from fewer non-zero data. However, the full spread, of at most 5 flux units, would give rise to a frequency shift of only ∼0.01 μHz for a mode at ∼3000 μHz. This is somewhat smaller than the frequency uncertainty associated with modes in this part of the spectrum. The same is true for the other frequencies fitted in our study. We conclude, therefore, that our results should not be biased by any cycle-dependent contribution and that any additional scatter introduced in the fitted frequencies will be modest.

2.2. Mode parameter extraction

The power spectrum of each time series, $v_i'(t)$, was fitted to yield estimates of the mode parameters over the range $1800 \leq \nu \leq 3700$ μHz for each of the different duty cycles. Because of their close proximity in frequency the modes were fitted in pairs (i.e., $\ell = 2/0$ and $3/1$) – following the basic prescription of Chaplin et al. (1999) – to a model where all resonant components were represented explicitly in the form of the asymmetric function of Nigam & Kosovichev (1998). An additional offset was also included to describe the background in the fitting window. We used a multidimensional direction-set minimization algorithm (Press et al. 1992) to perform the fitting, maximizing an appropriate log-likelihood function. The natural logarithm of the height, width and background terms were varied – not the parameter values themselves – in order to give quasi-normal fitting distributions. Formal uncertainties on the fitted values were then derived from the Hessian matrix of each fit in the usual manner.

3. Results

In what follows we have chosen to investigate the variation of the fitted frequencies as a function of the fractional fill, D. As alternative independent variables we could have instead chosen to use the fitted first sideband height, which reflects the impact of the diurnal content of the window function, or the fitted background in the vicinity of the modes, which reflects the remaining, quasi-random content. Both do however show a clear functional dependence on the fill (albeit with some realization scatter), and as such we have settled on using D only. As an example of this dependence, Fig. 3 shows the fitted first fractional sideband height (left-hand panel) and background (right-hand panel) as a function of D for each of the 280 spectra at $\ell = 0$, $n = 18$.

3.1. Dependence of fitted frequency on duty cycle D

Figure 4 shows the fitted $\ell = 0$ frequencies at $n = 18$ and 25. They are plotted as a function of the fractional duty cycle, D, of each of the 280 sets. The single, representative error bar corresponds in each case to the mean formal uncertainty returned by the mode-fitting procedure.

The plots reveal that there is some systematic bias, i.e., a change in fitted frequency with fill. Interestingly, the sign of this is reversed for some radial orders (e.g., for the two modes shown here). A series of χ^2 tests for different polynomial orders indicated that in all cases a simple linear model provides an adequate description of any uncovered trend. The solid line in each panel shows such a straight-line fit.

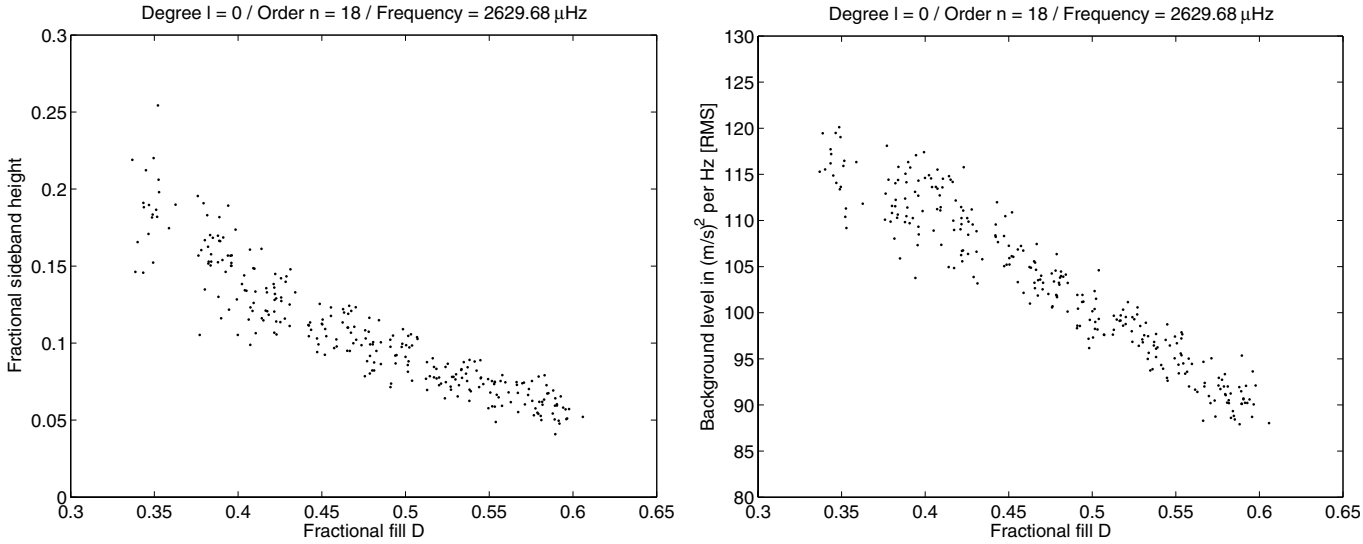

Fig. 3. Fitted first fractional sideband height (left-hand panel) and background level (right-hand panel) from the analysis of the 280 sets at $\ell = 0$, at $n = 18$. Both are plotted as a function of the fractional duty cycle, D, of each set.

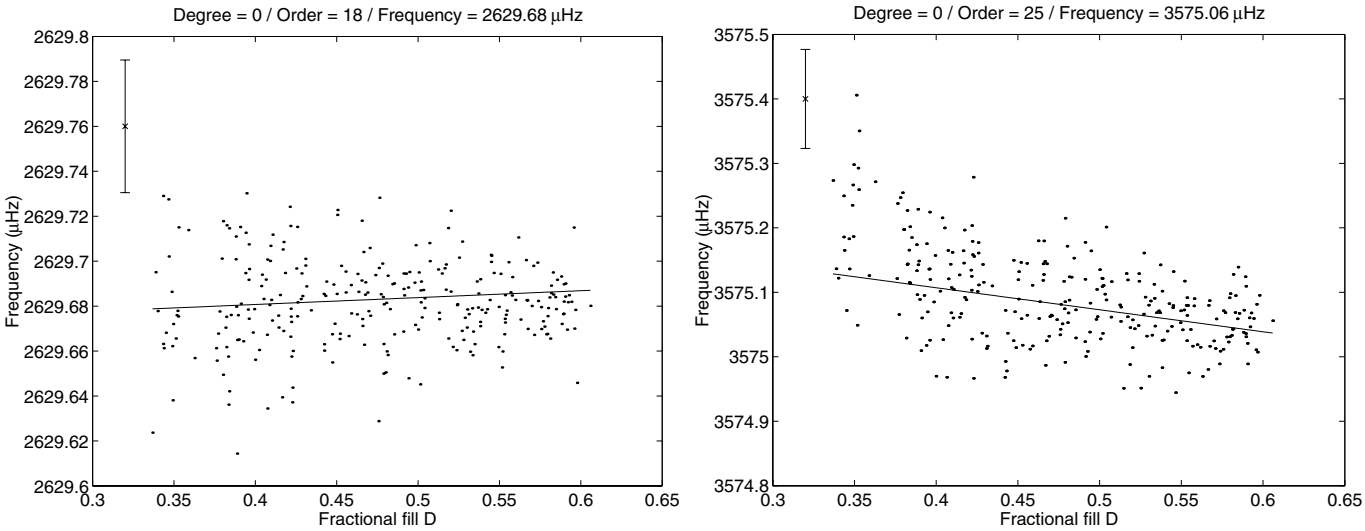

Fig. 4. Fitted $\ell = 0$ frequencies from the analysis of the 280 sets at $n = 18$ and 25, plotted as a function of the fractional duty cycle, D, of each set. The representative error bar in each panel corresponds to the mean formal uncertainty returned by the mode-fitting procedure. Also shown is a best straight-line fit to each cohort of fitted frequencies.

In order to quantify the extent to which the frequencies are "pulled" we fitted the frequency estimates extracted at each (n, ℓ) to the linear model:

$$\nu_{n\ell} = a_{n\ell} \cdot (1 - D) + b_{n\ell}, \tag{3}$$

where $a_{n\ell}$ and $b_{n\ell}$ are the gradient (or sensitivity) and the intercept respectively. We chose to fit against $1 - D$ so that the intercept corresponds to the frequency expected for $D = 1$ (assuming the linear relation persists at fills above the fitted-data range). For the $\ell = 0$ examples shown in Fig. 4, the fitted $a_{n\ell}$ is negative at $n = 18$, but positive at $n = 25$.

Figure 5 plots the fitted sensitivities, $a_{n\ell}$, for the complete range of modes covered by the analysis. The data have been paired off according to the ℓ with which they were fitted, i.e., $\ell = 0$ with 2 (left-hand panel) and $\ell = 1$ with 3 (right-hand panel). The sensitivity (ordinate) has been scaled in both panels to show the frequency change expected for each 10-per-cent *decrease* in fill, i.e., we show:

$$\delta\nu_{n\ell} = \delta D \cdot a_{n\ell} = 0.1 \cdot a_{n\ell}. \tag{4}$$

The implied systematic changes in frequency are, for the most part, quite small at mode frequencies up to ∼3300 μHz. The nature of the patterns in $a_{n\ell}$ reflects the complicated interaction of the window-function-related background and sideband structure with the varying characteristics (signal-to-noise ratio, linewidth) of the resonant structure. That the $\ell = 3$ coefficients appear more scattered than their lower-ℓ counterparts can be explained by the fact that these are by far the least

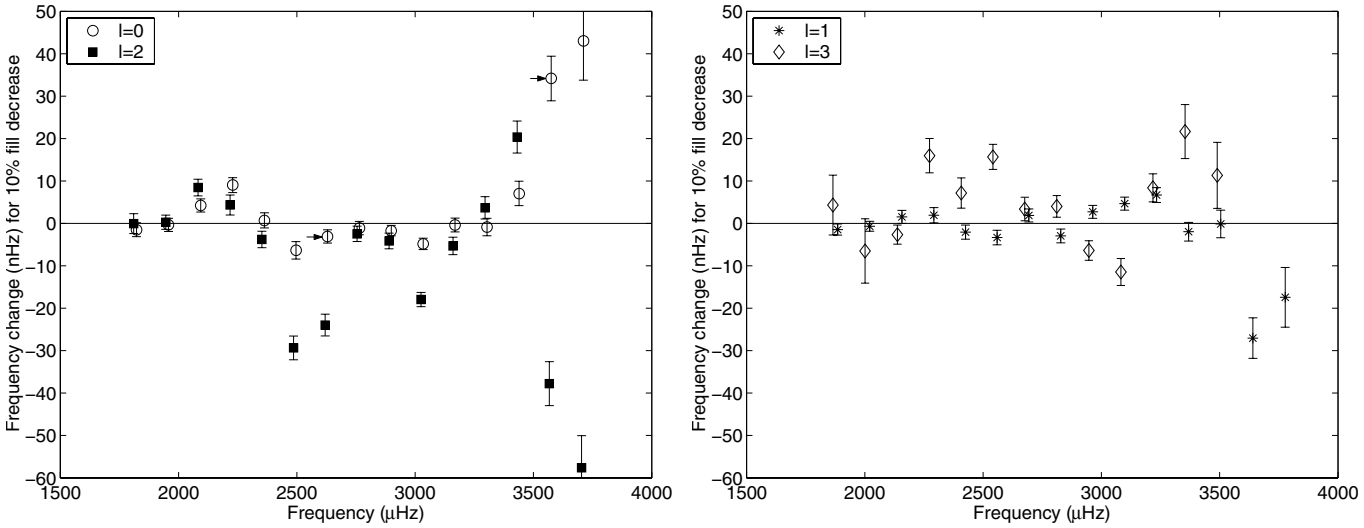

Fig. 5. Change in fitted frequency, $a_{n\ell}$ (in nHz), for a 10 per cent change in fill: left-hand panel – for $\ell = 0$ (○) and 2 (■); right-hand panel – for $\ell = 1$ (∗) and 3 (◊). (The two modes used in Fig. 4 are indicated by the arrows in the left-hand panel.)

prominent modes in BiSON-like, Sun-as-a-star data[2]. The uncovered "trend line" is therefore less well constrained.

Some of the systematic changes uncovered at $\ell = 2$ – in particular those at $n = 16$ (≈2486 μHz) and 17 (≈2620 μHz) – are well out of line with the rest of the data (larger in magnitude than their neighbouring points). In order to seek a possible explanation we have checked both the quality of the estimated frequencies used to generate these coefficients, and any obvious spacings in frequency that might account for them, e.g., the influence of the first diurnal sideband of the nearby $\ell = 0$. (The second sidebands lie too far from the principle peaks to affect the splitting.) In this part of the spectrum, the first sideband lies "over" one of the outer components of the $\ell = 2$. In spite of this we found that the fitting procedure is still able to extract robust estimates of the height of the sidebands, and that as a result the $\ell = 2$ frequencies are not "pulled" by an inadequate modelling of them. There were also no clear outliers in the fitted frequencies at $n = 16$ and 17.

Whilst we are therefore at somewhat of a loss for an explanation of the out-of-line coefficients, it is perhaps worth noting two points. First, we fitted the $\ell = 2$ modes assuming a symmetric arrangement of the components within each multiplet. One would expect the observed m to be unevenly distributed in frequency on account of the influence of the inhomogeneous distribution of magnetic activity over the surface of the Sun. However, this effect should be very small for most of the modes present in these 10-yr data (Chaplin et al. 2003b, 2004). Therefore, even though the fitting model used does not account for this, the effect is most likely not a cause for concern. Second, the linewidths of the modes have a local maximum (point of inflexion) at ∼2600 μHz, and it is over the central part of the spectrum that a clear tendency is revealed for the fitted widths to be overestimated in lower fill data (see Chaplin et al. 2003a). However, it is unclear why the $\ell = 2$ modes might be affected whilst at the same time the other ℓ are not.

At frequencies above ∼3300 μHz there is a clear increase observed in the size of the $a_{n\ell}$. This corresponds to that part of the spectrum where the linewidths of the resonant peaks begin to increase markedly, thereby giving rise to a total blending of adjacent azimuthal orders, m, within a given multiplet and blending between adjacent ℓ. The $\ell = 2$ and 0 modes are both very prominent and therefore have a strong influence upon one another, giving rise to the large increase in magnitude at the highest frequencies (left-hand panel of Fig. 5). In contrast, the weaker $\ell = 3$ modes have a less pronounced impact on their $\ell = 1$ neighbours and so the dipole frequencies are not pulled as strongly. The $\ell = 3$ modes become difficult to fit above ∼3500 μHz in low-fill data and so we have not analyzed these data at frequencies high enough to see the strong influence of the nearby $\ell = 1$ (right-hand panel of figure).

In order to place the sensitivities from Fig. 5 in full context they need to be compared to the typical precision with which the frequencies can be determined. If the fill-dependent frequency changes are found to be larger than the mode-frequency uncertainties the effect of this bias cannot then be ignored. As such, we look next at the behaviour of the errors.

3.2. Dependence of frequency uncertainty on duty cycle D

In the absence of any source of background noise, the precision with which the frequency of a mode can be determined is dependent upon its lifetime, $\tau_{n\ell}$, and the length of time for which the resonance has been observed, T. If $\Gamma_{n\ell}$ is the linewidth of the resonance in the frequency domain, we define the lifetime according to:

$$\tau_{n\ell} = 1/\pi\Gamma_{n\ell}. \quad (5)$$

This corresponds to the e-folding time for the amplitude of the signal. If the mode is observed for a duration corresponding to

[2] The most prominent, outer peaks of each $\ell = 3$ multiplet (those with $|m| = 3$) have peak heights (maximum power densities) in the frequency domain that are ≈20 per cent of their nearest, $\ell = 1$, $|m| = 1$ neighbours.

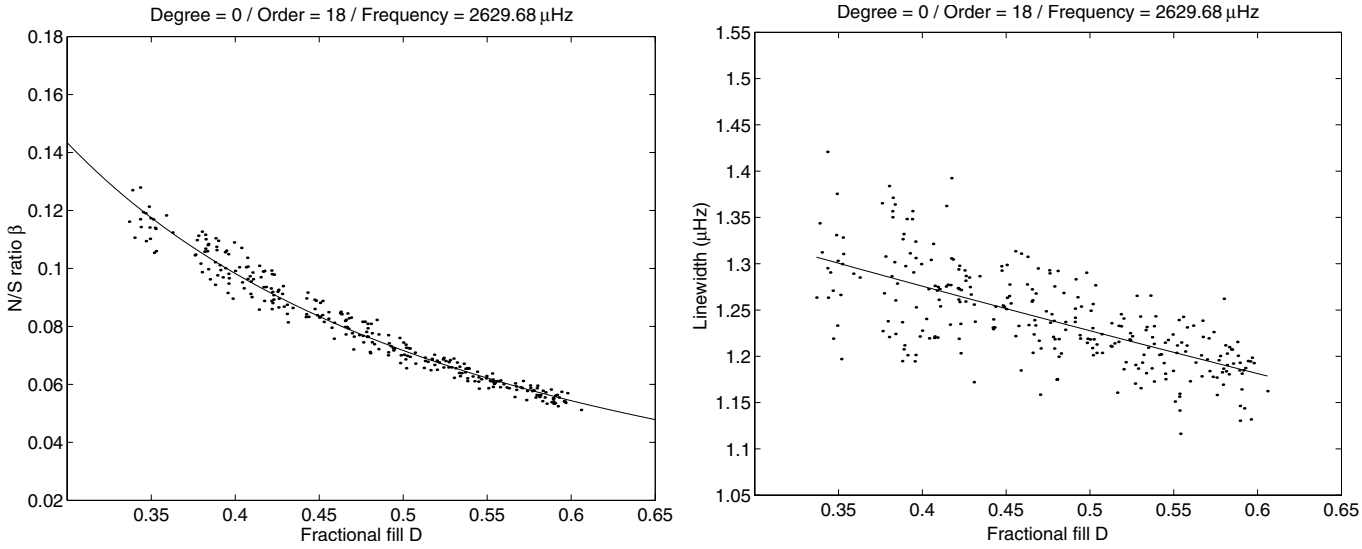

Fig. 6. Left-hand panel – fitted background-to-signal ratio, $\beta_{n\ell}$ at $\ell = 0$, $n = 18$ (scattered points), as a function of fill D; and the best fit to these data of the Chaplin et al. (2002) model described by Eq. (11) (solid line). Right-hand panel – fitted linewidths for the same mode (scattered data) and the best linear fit of these to the fractional fill, D (solid line).

a single lifetime, the error in the determination of the frequency will be of the order of (Chaplin et al. 2002):

$$\sigma_{\nu_{n\ell}} \sim \frac{\Gamma_{n\ell}}{2} \sim \frac{1}{2\pi\tau_{n\ell}}. \tag{6}$$

When $T \gg \tau_{n\ell}$, the observations will extend over a whole series of independent realizations of the mode. The mode will be resolved in the frequency domain and the precision with which its frequency can be determined will be increased by a factor $[T/\tau_{n\ell}]^{1/2}$. The uncertainty will therefore now be:

$$\sigma_{\nu_{n\ell}} \sim \frac{1}{2\pi\tau_{n\ell}} \cdot \left[\frac{\tau_{n\ell}}{T}\right]^{1/2} \sim \left[\frac{\Gamma_{n\ell}}{4\pi T}\right]^{1/2}. \tag{7}$$

In practice, the modes are observed against a background level, B. If $H_{n\ell}$ is the maximum power density (height) of the mode in the frequency domain, the prominence of its peak is dependent upon the background-to-signal ratio

$$\beta_{n\ell} = B/H_{n\ell}. \tag{8}$$

The presence of background affects the size of the uncertainty, and this can be modelled as a non-linear function of $\beta_{n\ell}$. Equation (7) is then modified to (Libbrecht 1992; Toutain & Appourchaux 1994):

$$\sigma_{\nu_{n\ell}} = \left[\frac{f(\beta_{n\ell})\,\Gamma_{n\ell}}{4\pi T}\right]^{1/2}, \tag{9}$$

in which the added function of $\beta_{n\ell}$ takes the form:

$$f(\beta_{n\ell}) = (1+\beta_{n\ell})^{1/2}\left[(1+\beta_{n\ell})^{1/2} + \beta_{n\ell}^{1/2}\right]^3. \tag{10}$$

For fixed T and $\Gamma_{n\ell}$ the uncertainty therefore depends upon $\beta_{n\ell}$. (Formulae for multi-component modes, i.e., for $\ell \geq 1$, are given in Appendix A.)

Chaplin et al. (2002) have shown that the background-to-signal ratio can be modelled as a monotonic function of the fractional duty cycle, D, according to:

$$\beta_{n\ell} = \frac{\Gamma_{n\ell} e_{n\ell}(1-D) + \beta^1_{n\ell}}{D}. \tag{11}$$

In the above, $\beta^1_{n\ell}$ is the background-to-signal ratio for $D = 1$ and $e_{n\ell}$ is a free parameter, characteristic of a given instrumental dataset, with a value at the centre of the p-mode spectrum of around 0.03 μHz^{-1}.

The left-hand panel of Fig. 6 shows how the fitted $\beta_{n\ell}$ varies at $\ell = 0$, $n = 18$ (scattered points). The observed increase with decreasing fill is representative of the trend uncovered for all modes. The solid line is the best fit of the Chaplin et al. (2002) model in Eq. (11) and can be used, together with the fitted linewidth, to predict $\sigma_{\nu_{n\ell}}$ for different D (assuming T = 3660 d).

The two panels of Fig. 7 show the formal $\ell = 0$ frequency uncertainties, $\sigma_{\nu_{n\ell}}$ (scattered points), compared to those predicted using Eqs. (9) and (11) (dashed line). The predictions were made using the fitted linewidths from the principal, 74-%-fill dataset. The radial orders chosen are the same as in Fig. 4. Again, the trend uncovered is typical of all ℓ and n we study: a decrease in precision at lower fill which is largely the result of the increase in the background-to-signal ratio.

Except at lower radial orders, where the match is good, the prediction is seen to underestimate the observed uncertainty at $n = 18$ and above. This discrepancy can be explained partly by the fact that no account has been taken in the above of the impact of the window function on the fitted linewidth $\Gamma_{n\ell}$. Chaplin et al. (2003a) performed an in-depth study of this and found a clear tendency for the linewidth to be overestimated in data with low fill. This is illustrated clearly in the right-hand panel of Fig. 6, where we plot the fitted linewidths for the $\ell = 0$, $n = 18$ mode. The solid line shows the best linear fit to these data (which, given the observed uncertainties in the width, provides an adequate description of the uncovered trend). When corresponding linear fits are used to parameterize the fitted $\Gamma_{n\ell}$ at each n in Eq. (9), the solid-line predictions of the uncertainties in Fig. 7 are given. These are seen to reduce the discrepancy between the formal and modelled errors. That which remains is strongly dependent upon mode frequency, with the

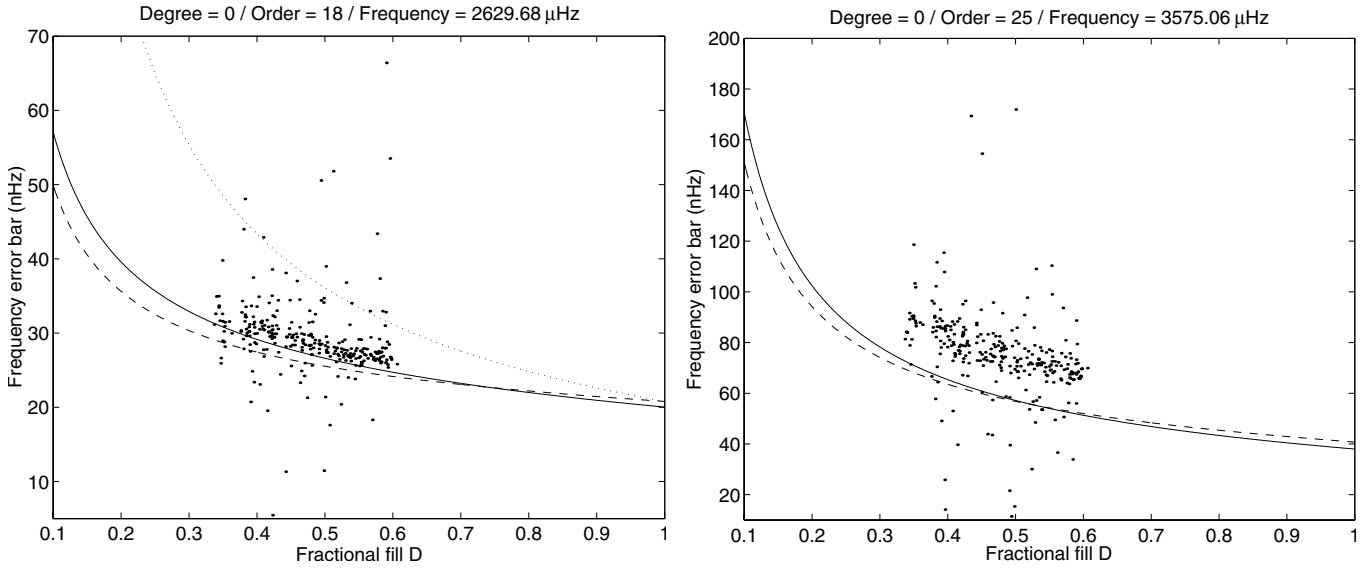

Fig. 7. Frequency uncertainties, $\sigma_{\nu_{n\ell}}$ (points), returned by the mode-fitting procedure as a function of fractional fill D for the same radial ($\ell = 0$) modes shown in Fig. 4. The different linestyles correspond to three predictions based upon the formula of Libbrecht (1992): dashed line – prediction using fitted linewidth from principal 74-per-cent-fill BiSON timeseries; solid line – prediction using linear function to describe fitted linewidth as a function of fill, D (see right-hand panel, Fig. 6); dot-dashed line in upper left-hand panel – modified model prediction of Gelly et al. (1997) and Fierry-Fraillon et al. (1998).

size of the formal errors increasingly exceeding the model at higher n (higher frequency). This can be explained by the sharp increase in linewidth in this part of the p-mode spectrum which gives rise to peak blending and the introduction of effects over and above those modelled in Eq. (9). The influence, and eventual blending, of components adjacent in frequency then makes it harder to constrain reliably the location of the mode peaks. This gives rise to a corresponding increase in the uncertainty (recall that the model assumes a peak in complete isolation).

The dotted line in the left-hand panel of Fig. 7 ($n = 18$) is the prediction based upon the formula of Gelly et al. (1997) and Fierry-Fraillon et al. (1998), who modified the factor of $T^{-1/2}$ in Eq. (9) to $(TD)^{-1/2}$. This is seen to result in a large overestimation of the fitted uncertainty, and indicates that the effect of the window function is adequately represented in the basic equation by the fitted $\beta_{n\ell}$ and $\Gamma_{n\ell}$.

3.3. Significance of the uncovered frequency bias

We are now in a position to assess the significance of the frequency bias present at each (n, ℓ). Figure 8 plots the magnitude of the bias, together with the fitted frequency uncertainties, as a function of fill D for the same $\ell = 0$ modes as in Figs. 4 and 7. We have chosen to display the *absolute* values of the bias so that they can be overlayed directly on the frequency uncertainties. (Note that the bias lines appear curved because of the logarithmic scaling on each ordinate.) We also show results for an additional $\ell = 2$ mode ($n = 16$). This is of interest since it shows an anomalously large bias (left-hand panel Fig. 5).

The magnitude of frequency bias in each case comes from the product $|(1 - D) \cdot a_{n\ell}|$ (cf. Eq. (4)). Even though the $a_{n\ell}$ were extracted from analyses covering $0.35 \leq D \leq 0.60$ only, we have assumed they are also valid outside this range. Tests with artificial data indicate this is a robust assumption.

The fitted uncertainties are plotted as points, with the bold line marked "3660 days" corresponding to the uncertainty predicted with Eq. (9) (for which we used the mean fitted linewidth and background-to-signal ratio for the mode in question, and $T = 3660$ d). The equation must be modified for the multiplet $\ell = 2$ mode, and we used Eq. (A.1) to generate the required predicted uncertainties (see full discussion of errors at $\ell \geq 1$ in Appendix A).

We have also plotted predicted uncertainties for three other dataset lengths. These were determined by scaling the 3660-d values according to the square root of ratio of the lengths[3]. The chosen values correspond to:

- 100 d, as often used in studies of the solar cycle (e.g., Komm et al. 2000; Chaplin et al. 2001; Gelly et al. 2002);
- 800 d, as required to study more subtle effects in the spectrum, e.g., $\ell = 2$ multiplet frequency asymmetry in high-activity data (Chaplin et al. 2003b, 2004); and
- 25 yr, which extends over two 11-yr Schwabe activity cycles, or one full 22-yr magnetic cycle and is of course our target length for analyses of the complete BiSON database.

Here, we assume that the bias is independent of the length of the dataset in order to give a first-order estimate of the problems involved for different lengths. Figure 8 shows that for the $n = 18$ radial mode the bias is smaller than the uncertainty for all lengths, T. This is representative of what is observed

[3] This assumes that the background-to-signal ratio $\beta_{n\ell}$ remains unaltered for the various T. As has been shown by Chaplin et al. 2003a this is indeed the case provided the mode is well resolved (i.e., $T \gg \tau$). This is satisfied for the range of modes and T considered in our analyses.

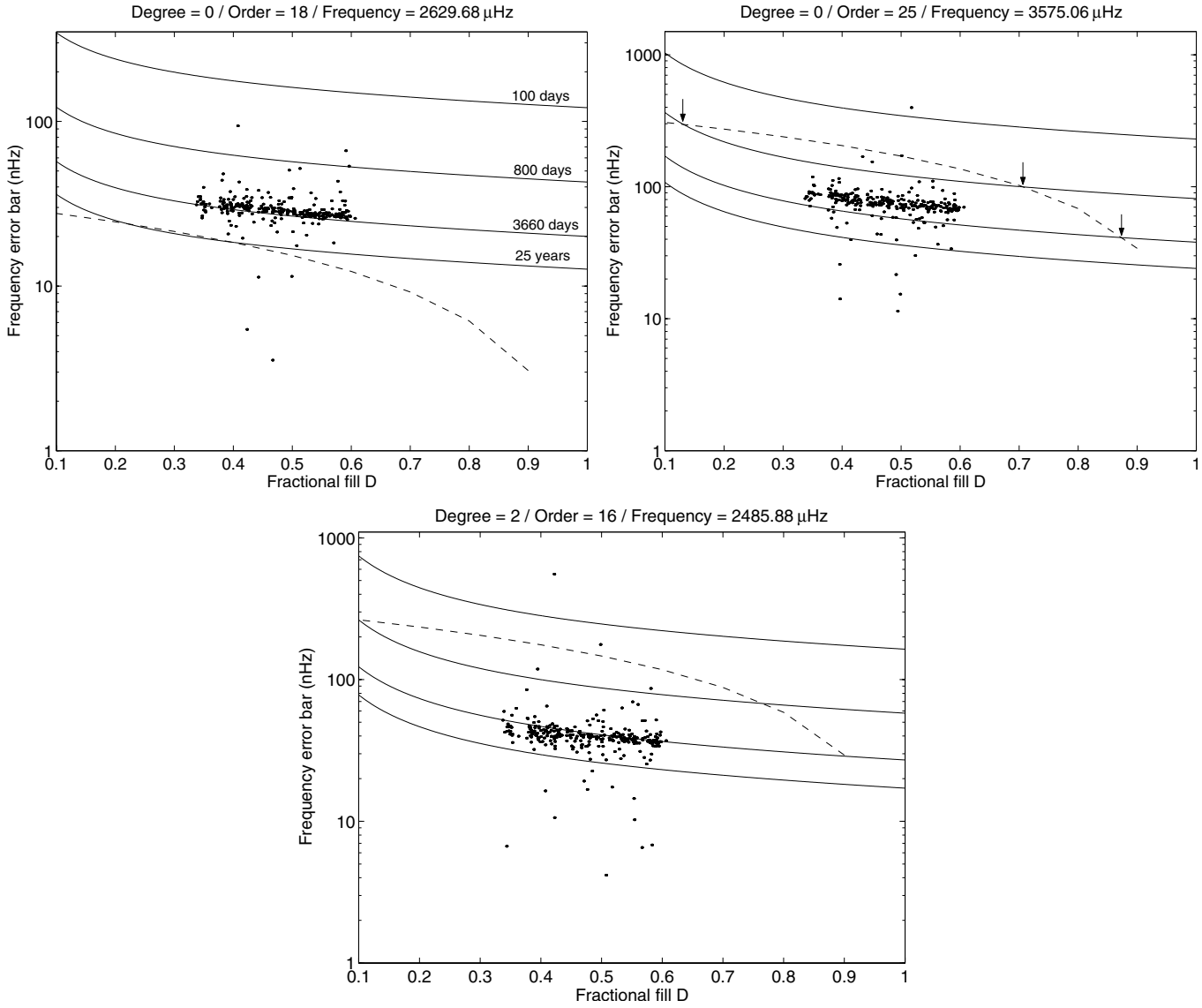

Fig. 8. Frequency uncertainties $\sigma_{\nu_{n\ell}}$ (scattered points) and absolute frequency bias, $|(1-D)\cdot a_{n\ell}|$ (dashed line), as a function of fractional fill D for two different radial ($\ell = 0$) modes (upper two panels) and one quadrupole ($\ell = 2$) mode (bottom panel). Also shown are the uncertainties predicted from Eq. (9) (Eq. (A.1) for $\ell = 2$) at the fitted 3660-d length and also extrapolated to lengths of 100 and 800 d and 25 yr.

for most of the modes below ≈3300 μHz, and implies that any bias introduced is not important at the level of precision of the low-ℓ data.

In our previous discussion in Sect. 3.1 we noted the larger bias at higher frequencies. Here, the $n = 25$ panel demonstrates that the bias does indeed exceed the uncertainty, $\sigma_{\nu_{n\ell}}$, for some combinations of T and D (when $T = 25$ yr, it does so for all D). Only at $T = 100$ d is the bias smaller in size than the observed uncertainty.

The $\ell = 2$ data shown (lower panel of Fig. 8) are for one of the modes whose extracted bias was much larger than, and therefore out of line with, its neighbouring points. Here, we again see that the bias is for some D significantly larger than $\sigma_{\nu_{n\ell}}$, but again only in the longer datasets.

Our results demonstrate that for the majority of low-ℓ modes the bias in the frequencies given by the effects of the window is not significant. However, for modes above ≈3300 μHz, and some $\ell = 2$ modes near ≈2500 μHz, they become important when the frequencies are extracted from long datasets. As such allowance then needs to be made to avoid misinterpreting any results.

Acknowledgements. We would like to take this opportunity to acknowledge the significant technical contribution made over many years to BiSON by our recently deceased colleague, H. K. Williams. We are indebted to J. Allison, R. Bryan and B. Jackson for their technical and analysis support in Birmingham and to former colleagues, in particular C. P. McLeod, J. Litherland and R. Lines. We also thank P. Fourie at SAAO; the Carnegie Institution of Washington; the Australia Telescope National Facility (CSIRO); E. J. Rhodes (Mt.Wilson, California); and members (past and present) of the IAC, Tenerife. BiSON is funded by the UK Particle Physics and Astronomy Research Council.

References

Bedding, T. R., & Kjeldson, H. 2003, PASA, 20, 203

Basu, S. 2003, Ap&SS, 284, 153

Chaplin, W. J., Elsworth, Y., Howe, R., et al. 1996, Sol. Phys., 168, 1

Chaplin, W. J., Elsworth, Y., Isaak, G. R., Miller, B. A., & New, R. 1999, MNRAS, 308, 424

Chaplin, W. J., Appourchaux, T., Elsworth, Y., Isaak, G. R., & New, R. 2001, MNRAS, 324, 910

Chaplin, W. J., Elsworth, Y., Isaak, G. R., Miller, B. A., & New, R. 2002, MNRAS, 330, 731

Chaplin, W. J., Elsworth, Y., Isaak, G. R., et al. 2003a, A&A, 398, 305

Chaplin, W. J., Elsworth, Y., Isaak, G. R., et al. 2003b, MNRAS, 343, 343

Chaplin, W. J., Appourchaux, T., Elsworth, Y., et al. 2004, A&A, 416, 341

Elsworth, Y., Howe, R., Isaak, G. R., et al. 1995, A&AS, 113, 379

Fierry Fraillon, D., Gelly, B., Schmider, F. X., et al. 1998, A&A, 333, 362

Fierry Fraillon, D., & Appourchaux, T. 2001, MNRAS, 324, 1159

Fossat, E., Kholikov, S., Gelly, B., et al. 1999, A&A, 343, 608

Gelly, B., Fierry Fraillon, D., Fossat, E., et al. 1997, A&A, 323, 235

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

Harvey, J. W., Hill, F., Hubbard, R., et al. 1996, Science, 272, 1284

Komm, R. W., Howe, R., & Hill, F. 2000, ApJ, 531, 1094

Lazrek, M., & Hill, F. 1993, A&A, 280, 704

Libbrecht, K. G. 1992, ApJ, 387, 712

Nigam, R., & Kosovichev, A. G. 1998, ApJ, 505, L51

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, in Numerical recipes in Fortran, 2nd ed. (Cambridge: University Press), 407

Roxburgh, I. W., & Vorontsov, S. V. 2001, MNRAS, 322, 85

Roxburgh, I. W., & Vorontsov, S. V. 2002, in Proc. of the First Eddington workshop on stellar structure and habitable planet finding, ed. F. Favata, I. W. Roxburgh, & D. Galadi, ESA SP-485, Noordwijk, 349

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

Verner, G. A., Chaplin, W. J., & Elsworth, Y. 2004, MNRAS, in press

Online Material

Appendix A: Dependence of frequency uncertainty on angular degree ℓ

The analysis of the frequency uncertainties presented in Sect. 3.2 must be extended when $\ell \geq 1$. This is to allow for the fact that these multiplets contain more than one peak.

The various panels of Fig. A.1 show the mean formal frequency uncertainties returned at $\ell = 0$, 1, 2 and 3 by the mode-fitting procedure (symbols joined by unbroken line). The dashed line is in each case the single-peak prediction of Eq. (9) and is meant to serve as an eye guide. These predictions were made using the average fitted linewidth and background-to-signal ratio for each mode. That there is a mis-match between the single-peak-modelled and fitted errors is not surprising. This then begs the question of how Eq. (9) needs to be modified when $\ell \geq 1$.

The nature of any modification depends upon the number of peaks that are observed, and the separation in frequency between them. When Sun-as-a-star observations are made in or close to the ecliptic plane (e.g., for BiSON and GOLF) only those components for which $l + m$ is even are clearly visible. Furthermore, the outer $m = \pm\ell$ peaks are by far the most prominent. Provided the separation of the outer components is much larger than their inherent linewidth the mode can to first order be modelled as two independent peaks.

If $\nu_{s_{n\ell}}$ is the splitting between adjacent m, this demands that the "reduced splitting", $[2\ell\nu_{s_{n\ell}}/\Gamma_{n\ell}]$, be much greater in size than unity. Provided this condition is met the frequency uncertainty will to good order be given by:

$$\begin{aligned}\sigma_{\nu_{n\ell}} &= \frac{1}{\sqrt{2}} \cdot \left[\frac{f(\beta_{n\ell})\,\Gamma_{n\ell}}{4\pi T}\right]^{1/2}, \\ &= \frac{1}{\sqrt{g_{n\ell}}} \cdot \left[\frac{f(\beta_{n\ell})\,\Gamma_{n\ell}}{4\pi T}\right]^{1/2}. \end{aligned} \tag{A.1}$$

In the above we have followed the convention of Toutain & Appourchaux (1994) and defined the multiplicative correction factor, $g_{n\ell}$, as shown. For the idealized case of two independent peaks, $g_{n\ell} = 2$. When the assumption of independence is no longer valid, i.e., when $2\ell\nu_{s_{n\ell}}/\Gamma_{n\ell} \not\gg 1$, the peaks begin to merge and the fitted uncertainty increases, i.e., the correction $g_{n\ell}$ decreases as the fit becomes less-well constrained. This is illustrated in the left-hand panel of Fig. A.2 where the formal uncertainties from the fitting procedure are plotted as a function of the reduced splitting. The errors are seen to increase noticeably when $[2\ell\nu_{s_{n\ell}}/\Gamma_{n\ell}] \lesssim 2$.

We have estimated $g_{n\ell}$ for each mode at $\ell \geq 1$ using the fitted data. Rearrangement of Eq. (A.1) gives:

$$g_{n\ell} = \frac{f(\beta_{n\ell})\,\Gamma_{n\ell}}{4\pi T \sigma^2_{\nu_{n\ell}}}. \tag{A.2}$$

The value of $g_{n\ell}$ for a given mode then follows from the mean fitted $\beta_{n\ell}$, $\Gamma_{n\ell}$ and $\sigma_{\nu_{n\ell}}$. The resulting estimates are shown in the right-hand panel of Fig. A.2 as a function of the reduced splitting. Here, we have plotted the square root of the parameter, rather than $g_{n\ell}$ itself, since it is on this that the frequency uncertainty is linearly dependent. Toutain & Appourchaux (1994) plotted their model predicted $g_{n\ell}$ against the same reduced splitting parameter (their Fig. 1) for $\ell = 1$ and we find similar observed values here. As one would expect, for larger values of the reduced splitting, i.e., a larger separation of the outer modal components, $(g_{n\ell})$ does indeed tend to the expected value of 2.

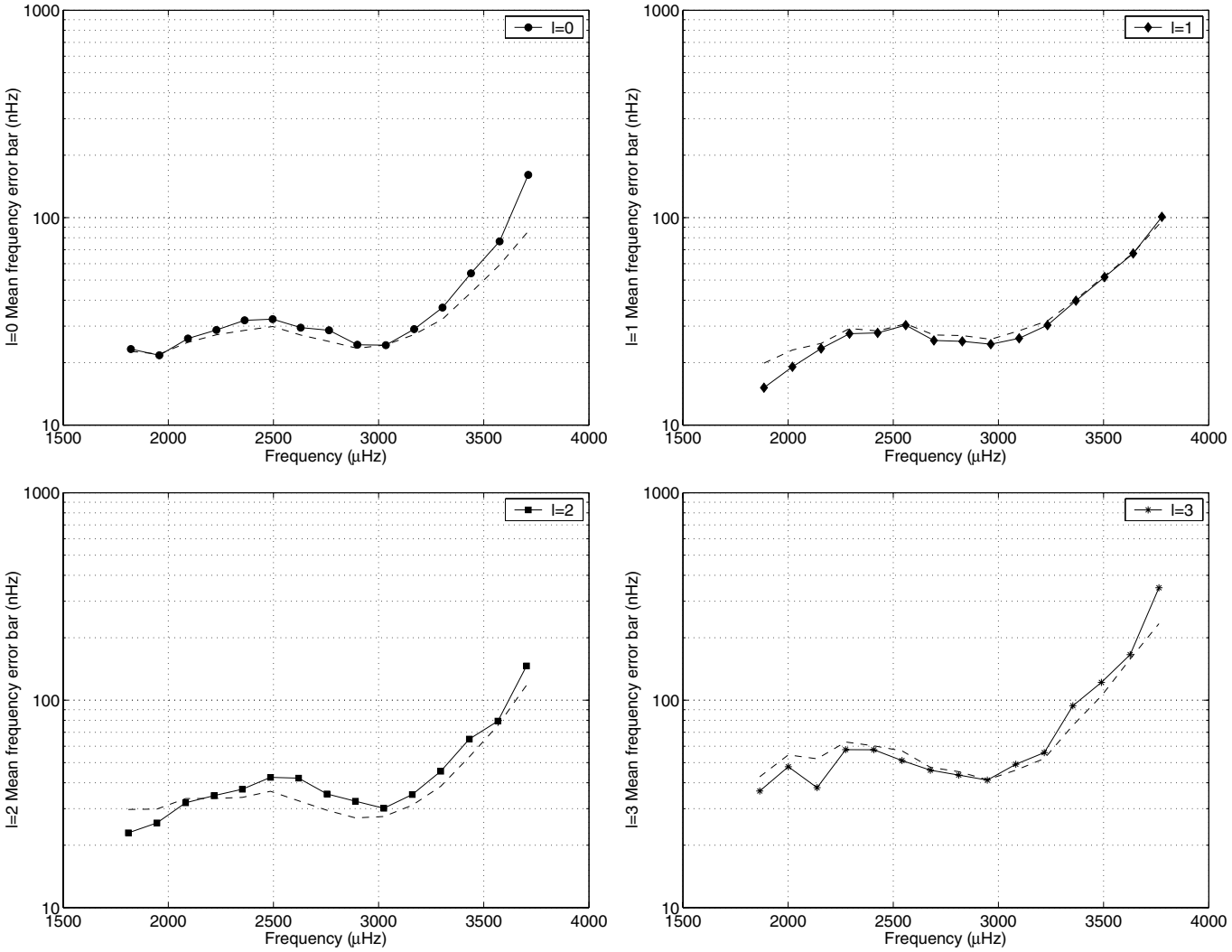

Fig. A.1. Mean formal frequency uncertainties (as returned by the mode-fitting procedure) for ℓ = 0, 1, 2 and 3 modes. The dashed lines are error bars predicted with the single-peak formula (Eq. (9)) using the average fitted linewidth and background-to-signal ratio for each mode.

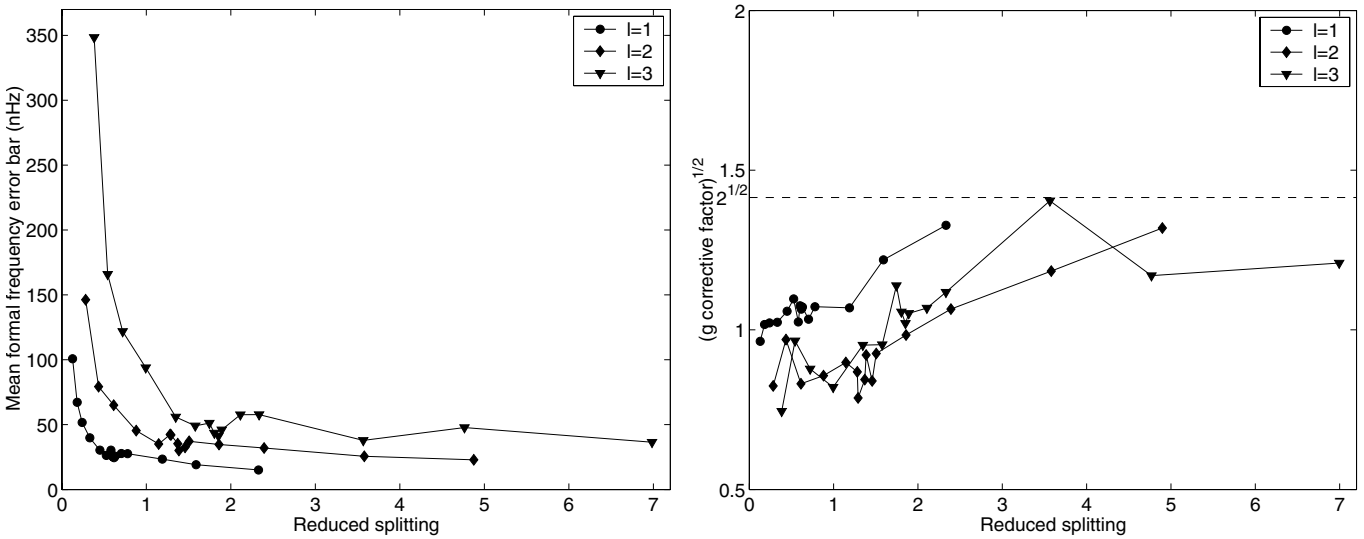

Fig. A.2. Left-hand panel – mean formal frequency uncertainties at ℓ = 0, 1, 2 and 3, plotted as a function of the splitting-linewidth ratio (reduced splitting) $2\ell\nu_{s_{n\ell}}/\Gamma_{n\ell}$. Right-hand panel – square root of the modelled uncertainty correction factor, $(g_{n\ell})^{1/2}$, also as a function of the reduced splitting.