A&A 416, 341–351 (2004)
DOI: 10.1051/0004-6361:20034465

Astronomy & Astrophysics

Solar p-mode frequencies at $\ell = 2$: What do analyses of unresolved observations actually measure?

W. J. Chaplin[1], T. Appourchaux[2], Y. Elsworth[1], G. R. Isaak[1], B. A. Miller[1], R. New[3], and T. Toutain[4,5]

[1] School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[2] ESA Research and Science Support Department, ESTEC, 2200 AG Noordwijk, The Netherlands
[3] School of Science & Mathematics, Sheffield Hallam University, Sheffield S1 1WB
[4] Département Cassini, URA CNRS 1362, Observatoire de la Côte d'Azur, 06304 Nice, France
[5] Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029, 0315 Oslo, Norway

Received 7 October 2003 / Accepted 3 December 2003

Abstract. We have studied in detail the extraction of estimates of $\ell = 2$ p-mode frequencies from unresolved observations of the visible disc of the Sun. Examples of data of this type include ground-based observations made by the Birmingham Solar-Oscillations Network (BiSON), and space-borne observations made by the GOLF and VIRGO/SPM instruments on board the ESA/NASA *SOHO* satellite. The fitting of the modes is complicated in practice by the asymmetric arrangement in frequency of the three components ($m = -2$, 0 and 2) that are prominent in such data. In order to investigate the effect of this we used a series of 10-yr artificial datasets into which varying degrees of asymmetry were introduced. The sets were designed to mimic the characteristics of the BiSON and GOLF data, and were analyzed both with and without the BiSON window function from the period 1992 through 2001. Since reliable estimates of the asymmetry have only recently been extracted from unresolved observations (Chaplin et al. 2003a) it has for a long time been standard practice to fit the $\ell = 2$ modes to a model that assumes a symmetrically arranged multiplet. We have tested the impact of this on the accuracy of the extracted frequencies. Furthermore, we demonstrate that asymmetric models can be successfully applied, provided the data are of sufficient length and quality. We also discuss the implications of our simulations for analyses of real solar data.

Key words. methods: data analysis – Sun: helioseismology – Sun: magnetic fields

1. Introduction

Doppler velocity or intensity observations of the unresolved Sun (i.e., "full-disc" or Sun-as-a-star data) have the requisite spatial sensitivity and stability to allow studies of globally coherent, low angular-degree (low-ℓ) p modes to be made. An invaluable database of observations of this type has been accumulated by ground-based networks (e.g., BiSON and IRIS) and space-borne instruments (e.g., GOLF and VIRGO/SPM on the ESA/NASA *SOHO* satellite) and this has played both a formative, and key on-going, rôle in studies of the deep radiative interior of the Sun.

The net mode signal observed from a given angle of inclination, i, depends upon the average over the visible disc of its spatial displacement eigenfunction, this being a spherical harmonic described by an angular degree and azimuthal order, m. When observations are made in or near the ecliptic plane (i.e., for $i \sim 90°$) only those components for which $\ell + m$ is even are clearly visible in unresolved data. This is of course the perspective offered from ground-based networks and the *SOHO* satellite at the inner Lagrangian point. Furthermore, the sensitivity of these data to different azimuthal orders is such that the outer components are observed to be the most prominent. (The zonal, $m = 0$ mode is the only component present at $\ell = 0$, whilst at $\ell = 1$ only $|m| = 1$ is seen.)

Once made, a set of unresolved solar observations gives a single time series, the frequency spectrum of which consists of many closely spaced resonant peaks. Due to their proximity in frequency the individual m of a multiplet at $\ell \geq 1$ cannot be analyzed independently; suitable models that seek to describe the characteristics of the m present must therefore be fitted to the components simultaneously. Furthermore, when $\ell \geq 2$ the relative arrangement in frequency of the visible components within each multiplet will show *asymmetry*, i.e., an asymmetric arrangement – or non-uniform spacing – of frequencies. This is because of the differing response of individual m to the distribution of activity over the solar surface. The asymmetry will also vary over time in response to changes in activity over the solar cycle.

The influence of nearby ℓ cannot be ignored either, particularly so at higher mode frequencies where the widths of the resonant peaks are comparable to the separation in frequency of adjacent multiplets. This means that the mode structure must also be fitted over much of the spectrum in pairs (i.e., monopole modes with quadrupole modes, and dipole modes with octupole modes).

Send offprint requests to: W. J. Chaplin,
e-mail: wjc@bison.ph.bham.ac.uk

These primary dataset characteristics raise the important question of how estimates of the eigenfrequencies are affected by having to accurately model, and then fit, the various components within a given multiplet; and how the fitted frequencies actually relate to those of the underlying components. Here, we seek to investigate these issues for $\ell = 2$ multiplets in unresolved data. It is at this angular degree that the impact of frequency asymmetry becomes an issue, and our work here represents the first in-depth study of the consequences of this for low-ℓ modes.

An attempt to quantify the impact of frequency asymmetry is important in both a solar and stellar context. Estimates of the $\ell = 2$ p-mode frequencies and splittings provide important input to studies of the solar core and deep radiative interior. We also stand on the threshold of having a wealth of data of this type from both ground-based and satellite observations of solar-like stars. This of course raises the question of how many years it might be before the complications studied here will need to be considered in stellar data. Given the current high quality of ground-based observations (e.g., the recent α Cen A data of Butler et al. 2003) the answer is probably after the accumulation of only a few years of data (lengths at which one can attempt to constrain the asymmetric structure of the $\ell = 2$ multiplet in unresolved solar data). Clearly, the particular problems posed by a given star will depend upon the angle of inclination, i, it offers to us and the characteristics of its rotation and surface activity. Nevertheless, the results here serve to illustrate some of the subtle analysis issues that will need to be addressed once high-quality stellar time series become a commonplace data product.

The layout of the rest of the paper is as follows. In Sect. 2 we give an overview of the particular problems posed by the $\ell = 2$ mode structure. To test the impact of this on fits to the modes we have generated and analyzed artificial p-mode-like data, whose characteristics – in particular the arrangement in frequency of the $\ell = 2$ components – are known a priori. Two cohorts of data were generated, as described in Sect. 3, and then fitted with the strategies outlined in Sect. 4.

In the first set of simulations we introduced different levels of asymmetry at each radial order, n. Analyses of these data (Sect. 5) were used to map out the response of the fits to input asymmetries covering a wide dynamic range. With this behaviour characterised fully, a second cohort of data were then constructed to mimic the asymmetries expected at each n in *real* solar data spanning almost a decade. These allowed us to test the influence of, and the importance of allowing for, the asymmetry in analyses of the extensive database of solar observations now available (Sect. 6). The results are drawn together and summarized in Sect. 7.

2. The particular case of $\ell = 2$

Unresolved observations of the Sun with $i \sim 90^\circ$ show prominent quadrupole-mode structure at $m = -2$, 0 and 2 only. The spatial eigenfunctions of the outer, $|m| = 2$ modes are sensitive predominantly to the activity at lower latitudes, where the most conspicuous solar-cycle changes occur. Here, large, strong-field structures contribute most of the observed magnetic flux (e.g., De Toma et al. 2000). In contrast, the $m = 0$ mode shows a greater relative response to activity at higher latitudes where strong-field magnetic structures are almost completely absent, and only weak-component flux is present. It is therefore expected to show a smaller minimum-to-maximum shift than its $|m| = 2$ neighbours if one assumes that this change is the result of the influence of magnetic field structures (be that directly, via the Lorentz force, or indirectly, as a result of changes in stratification). Asymmetry in frequency is therefore expected.

Observational evidence for such behaviour has been found in unresolved data. Chaplin et al. (2003a) demonstrated that it was possible to detect changes in the arrangement of quadrupole multiplet structure in BiSON and GOLF data. They found that the frequency asymmetry

$$d_n = \frac{1}{2}\left(\nu_{n,m=+2} + \nu_{n,m=-2}\right) - \nu_{n,m=0}, \quad (1)$$

was indistinguishable from zero (at the level of precision of the observations) during an epoch of low activity, but increased to a marginally significant positive value at high activity (in a manner dependent upon the radial order, n). The results were found to be in reasonable agreement with a range of predictions (Dziembowski & Goode 1997; Dziembowski et al. 2000; Moreno-Insertis & Solanki 2000). From these, d_n for a mode at 3000 μHz is expected to be approximately $\approx$300 nHz at the epoch of maximum activity. This is to be compared with the average separation between the visible m of $\sim$800 nHz. (Here, we use d_n to identify the frequency asymmetry, rather than a_n used by Chaplin et al. (2003a). This is to avoid confusion with the a coefficients used to describe the multiplet structure in resolved-Sun data. Examples of this latter class of observations – for which the solar surface is resolved into many pixels – include the ground-based GONG and ECHO networks, and the SOI/MDI and LOI/VIRGO instruments on board *SOHO*. Appendix A shows how the d_n and resolved a coefficients are related.)

In making their predictions, Dziembowski & Goode suggested the need for fitting models to include some representation of the frequency asymmetry. The advance of Chaplin et al. does indeed suggest that the common practice of fitting symmetric multiplet models to the data should at last be dropped in favour of one that always allows for a potential, non-zero d_n. Doing so by attempting to *fit* for the asymmetry may not, however, always be a wise course.

The small magnitude of the effect is such that high resolution is required in the frequency domain in order to extract reasonably well-constrained estimates of the asymmetric arrangement of the components. This demands the use of long data sets of size $\gtrsim$2 to 3 yr. However, the lengthier the basic dataset cadence, the smaller the observed change in the d_n over the solar cycle. For example, the asymmetry present in data covering a full 11-yr cycle will be much smaller than in a shorter-length piece at or near the epoch of activity maximum. The effect is therefore likely to be small in those long datasets which have sufficient resolution to in principle allow the d_n to be fitted. Unfortunately, while it can potentially be much larger in short datasets (e.g., the 108-d cadence commonly used for studies

of p-mode variations) these have insufficient resolution to give well-constrained estimates of the asymmetry.

This suggests that fits to long datasets must be used to constrain the possible size of the effect. In addition, guidance can also be offered by resolved observations (e.g., see Dziembowski & Goode 1997). If data from higher ℓ are used this approach is only valid if the agent responsible for the introduction of the asymmetry is confined in the near-surface layers (i.e., there is no significant contribution from magnetic-field structures at the base of the convection zone or deeper). It is also worth noting that for the foreseeable future this option will not be available in the stellar case.

In order to study the impact of the multiplet asymmetry on fits of BiSON-like data we have constructed, and then analyzed, a series of artificial 10-yr datasets into which varying degrees of asymmetry have been introduced. In the first cohort of simulations the asymmetry was held at a constant level for all modes over the duration of each simulated set. Different levels were introduced into different sets to allow us to assess and then calibrate any bias present in the various analyses. For the second cohort of data we chose to simulate observations made over the period 1992 July 25 through 2002 January 9, i.e., covering the declining phase of cycle 22 and all of the rising phase of cycle 23. To achieve this the asymmetry was introduced as a time-varying parameter designed to mimic the impact of variations in activity over this epoch. As such, this allowed us to judge the impact of the fitting problems on data that are expected to be similar to actual BiSON and GOLF observations.

Analyses of both cohorts were performed with and without the BiSON window function from the chosen 10-yr period. Those results extracted without the window should therefore be regarded a guide for analyses of the near-continuous GOLF dataset (whose characteristics the artificial data also match closely); while those with the window provide, of course, direct guidance for the analysis of BiSON observations.

3. Artificial data

3.1. Overview of construction

The Laplace transform solution of the equation of a forced, damped harmonic oscillator was used to generate artificial p-mode-like data on a 40-s cadence in the time domain (as described at length in Chaplin et al. 1997). All m components to which the full-disc observations are sensitive were represented explicitly in each complete time series set as an oscillation signal generated by a single "run" of the model. Each component was re-excited continually by small "kicks" at each 40-s sample and we also adhered to a stochastic description that assumed each mode was excited independently. This was realized in practice by randomly phasing, in time, all generated signals.

The model gives rise to a limit spectrum that is a Lorentzian in shape for each component. As such, it provides no representation of the small peak asymmetries that have now been uncovered at low ℓ (e.g., Toutain et al. 1998; Chaplin et al. 1999; Chaplin & Appourchaux 1999; Thiery et al. 2000; Bertello et al. 2000). While we accept this is a shortcoming of the simulations, the small size of the asymmetries is such that other factors (e.g., the blending of adjacent mode components) impose more severe constraints on the fitting problem posed here.

Each time series was made with a full cohort of simulated low-ℓ modes, covering the ranges $0 \leq \ell \leq 5$ and $1000 \leq \nu \leq 5200$ μHz. A "pink" background noise component was added also in the time domain, the size of which increased at lower frequencies in order to give realistic signal-to-noise ratios. We used a database of mode frequency, power and linewidth estimates obtained from analyses of BiSON spectra in order to fix the characteristics of each oscillator component (i.e., its resonant frequency and characteristic damping constant and power). It is worth adding that the variation of the signal-to-noise ratio in the modes is found also to mimic excellently that in GOLF power spectra (e.g., see Chaplin et al. 2003b).

Of interest here, the heights of the zonal $\ell = 2$ components were fixed at 55% of their adjacent $|m| = \ell$ neighbours to match the values found in BiSON data by Chaplin et al. (2001a). Visibility levels for the barely-detectable $\ell = 4$ and 5 modes were fixed at the levels calculated by Christensen-Dalsgaard (1989). At the extreme ends of the modeled spectra – where we do not have reliable, fitted estimates for the parameters – we made appropriate, low-order extrapolations of the parameter values from the adjacent interior frequency bands.

3.2. Introduction of asymmetry

Artificial sets of length 3456 d (~10 yr) were constructed, into which $\ell = 2$ multiplet frequency asymmetries were introduced in one of two ways. In our first set of simulations, multiplets at all n were constructed with a fixed, time-invariant d_n. We made eleven 3456-d sets, each with their own input d_n, ranging from -0.5 to 0.5 μHz in steps of 0.1 μHz. The separation between the outer $|m| = 2$ components was chosen to be 1.2 μHz to match approximately values of the splittings extracted from fitting real data.

In the second cohort of simulations a time-dependent variation of the natural oscillator frequencies was introduced. This was achieved as follows. We fitted a smoothed record of the 10.7-cm radio flux – covering the 10-yr period from 1992 July 25 through 2002 January 9 – to a sixth-order polynomial in time. After multiplying by a suitable calibration factor (see below) to reflect the observed relationship between the radio flux and mode frequencies, the resulting values gave the required, time-dependent frequency changes. In order to mimic the activity cycle, the basic database frequencies used to construct each mode (i.e., the underlying frequency of the excited oscillator) were varied, sample-by-sample, by the linear addition of this modelled solar-cycle contribution.

The accepted quiet-Sun level of the 10.7-cm radio flux is 64 RF units (64×10^{-22} W m^{-2} Hz^{-1}; see Tapping & DeTracey 1990). If $A(t)$ is the polynomial-fitted value of the radio flux observed at time t, and $A_{\rm quiet}$ its quiet-Sun magnitude, the frequency shift induced was modelled in full as:

$$\delta\nu_{n\ell m}(t) = g_{\ell m} \cdot \mathcal{F}(\nu) \cdot [A(t) - A_{\rm quiet}]. \quad (2)$$

In the above $g_{\ell m}$ calibrates the size of the shift for different ℓ and m. The function $\mathcal{F}(\nu)$ reflects the frequency dependence in the shifts. It measures the fractional amount by which the

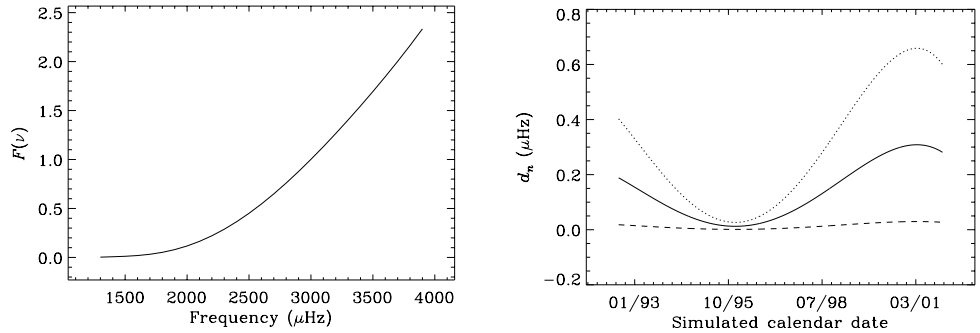

Fig. 1. Left-hand panel – plot of $\mathcal{F}(\nu)$ (see Eq. (2)) as a function of frequency. It measures the fractional amount by which the shift of a low-ℓ mode of a given frequency either exceeds or falls short of that expected at 3000 μHz. Right-hand panel – time-dependent values of the multiplet frequency asymmetry, d_n, introduced into the second cohort of artificial data. The three curves show d_n for the $\ell = 2$ overtones at $n = 13$ (dashed), $n = 20$ (solid) and $n = 26$ (dotted).

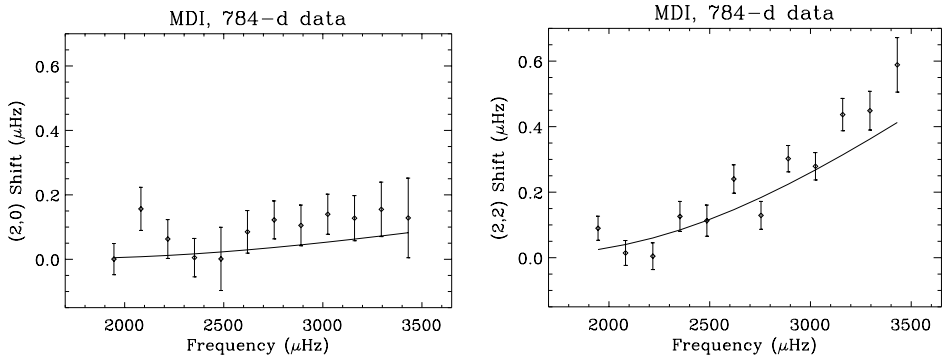

Fig. 2. Frequency shift of the (2, 0) (left-hand panel) and (2, 2) (right-hand panel) components as measured by fits to two 784-d spectra from resolved observations made by the MDI instrument on the ESA/NASA *SOHO* satellite. The solid line is the predicted shift based on the parameterization outlined in Sect. 3.2 (cf. Eq. (2)), which also gives the modelled time-dependent asymmetries in the right-hand panel of Fig. 1.

shift of a notional low-ℓ mode of that frequency either exceeds or falls short of that expected at 3000 μHz. The function, which is plotted in Fig. 1, was derived from GONG, BiSON and LOI/VIRGO observations by Chaplin et al. (2001b).

Chaplin et al. (2003c) used BiSON observations to extract estimates of the $g_{\ell m}$ for $|m| = \ell$ modes. Here, we use their value of $g_{2,2} = 3.05\,\mathrm{nHz\,RF^{-1}}$ to describe the shift given to the (2, 2) components. Our choice of $g_{2,0} = 0.60\,\mathrm{nHz\,RF^{-1}}$ for the inner, zonal components was fixed in an effort to match the frequency-shift predictions of Dziembowski et al. (2000). They made use of frequency shifts observed in higher-ℓ data, collected by the MDI resolved-Sun instrument on the *SOHO* satellite, as the basis for these.

Since the sensitivity of the simulated $m = 0$ components to the changing "activity" is smaller than at $|m| = 2$, the artificial multiplets show varying degrees of positive asymmetry over the duration of the modelled epoch. The time-varying d_n introduced are shown in the right-hand panel of Fig. 1 for radial orders $n = 13$ (1957.4 μHz; dashed), $n = 20$ (3024.7 μHz; solid) and $n = 26$ (3839.7 μHz; dotted).

In order to test our artificial parameterization – in particular our decision to calibrate against Dziembowski & Goode's prediction based on their use of shifts from *higher* ℓ – we made use of resolved $\ell = 2$ data collected by MDI. The imaged nature of these observations means that individual m can be isolated thereby removing the substantial blending of adjacent m that is present in the BiSON/GOLF data. Here, we applied optimal masks (Toutain & Kosovichev 2000) to two 784-d time series collected over the periods 1996 May through 1998 June (low activity) and 1999 March and 2001 April (high activity). A standard maximum-likelihood fitting of each (ℓ, m) component was then made giving estimates of the (2, 0) and (2, 2) frequencies. Differences between the high and low-activity sets are plotted in Fig. 2. The solid line in each panel is the predicted shift based on the parameterization in Eq. (2), using the mean 10.7-cm fluxes observed over each of the two periods.

The observed shifts – while seen to be in reasonable agreement with the model – are nevertheless slightly larger in magnitude. However, given the uncertainties on the measured data this difference is significant at only $\lesssim 2\sigma$. We therefore chose to proceed with the model as formulated. It is worth noting that the differences are sufficiently small that were we to have modified our simulations to match the MDI observations this would have had an insignificant impact on the results.

Finally, we note that analyses were performed on the 10-yr artificial time series both with and without the application of the 3456-d BiSON window function. The fractional duty cycle of the full BiSON window is $D = 0.76$. In addition,

data were also subdivided into 108-d sets to test the analysis on short stretches. The resulting 108-d windows ranged in fill from $D = 0.63$ to 0.87.

4. Fitting strategies

The modes were fitted in the power spectrum of each time series in pairs (i.e., $\ell = 2/0$ and 3/1; see Chaplin et al. 1999) as a superposition of components described by the polynomial fitting formalism of Nigam & Kosovichev (1998), with an additional flat background term. All components within a given ℓ were assumed to take the same width. The Nigam & Kosovichev Lorentzian-like model includes a free parameter to allow for the asymmetric appearance of the resonant peaks. Although, as we have discussed already, the artificial resonant profiles are Lorentzian and therefore not asymmetric, we opted to vary the same number of free parameters that would have been fitted to real, asymmetric data.

Different parameter combinations were used to describe the frequency arrangement of the $\ell = 2$ multiplet. The basic strategy adopted the commonly-used practice of assuming a symmetric configuration. If $\boldsymbol{p}_i$ are the free parameters to be fitted, two must be used to describe the symmetric multiplet, i.e.,

$$\text{Symmetric strategy} \begin{cases} \nu_{n,2,+2} = \boldsymbol{p}_1 + 2\boldsymbol{p}_2 \\ \nu_{n,2,0} = \boldsymbol{p}_1 \\ \nu_{n,2,-2} = \boldsymbol{p}_1 - 2\boldsymbol{p}_2. \end{cases} \quad (3)$$

Here, $\boldsymbol{p}_1$ is the fitted frequency of the multiplet – referred to hereafter as $\nu^{\rm sym}_{n,2}$ – and $\boldsymbol{p}_2$ gives a measure of the component separation for $\Delta|m| = 1$. While this description provides no direct information regarding the size of any non-zero d_n present, the frequency parameter $\boldsymbol{p}_1$ will nevertheless be influenced by the presence of asymmetry in the multiplet.

Next, we sought to test the ability of the fitting to extract estimates of the locations of all three observed components. There are several ways in which one might choose to parameterize an asymmetric multiplet. Irrespective of the formulation adopted, the small expected size of the asymmetry compared to the component separations and widths means it is extremely difficult for any fitting to constrain the location of the central component and therefore the magnitude of any asymmetry present. The interaction of the various components also leads to strong correlations between the underlying frequency parameters.

Here, we have chosen to define separate models to fit, and extract estimates of, the components we wish to study[1]. To do so we used two "asymmetric" strategies. In the first we attempted to extract directly the (2, 0) frequency (with parameter $\boldsymbol{p}_1$), with two additional parameters used to fix the locations of the outer components:

$$\text{Asymmetric strategy \#1} \begin{cases} \nu_{n,2,+2} = \boldsymbol{p}_1 + 2\boldsymbol{p}_2 \\ \nu_{n,2,0} = \boldsymbol{p}_1 \\ \nu_{n,2,-2} = \boldsymbol{p}_1 - 2\boldsymbol{p}_3. \end{cases} \quad (4)$$

In what follows we refer to the fitted estimate of the (2, 0) frequency as $\nu^{\rm ass}_{n,2,0}$.

In the second, we instead defined the basic multiplet frequency parameter, $\boldsymbol{p}_1$, to be that of the mean of the outer (2, 2) and (2, −2) components, $\nu^{\rm ass}_{n,2,|2|}$:

$$\text{Asymmetric strategy \#2} \begin{cases} \nu_{n,2,+2} = \boldsymbol{p}_1 + 2\boldsymbol{p}_2 \\ \nu_{n,2,0} = \boldsymbol{p}_1 - \boldsymbol{p}_3 \\ \nu_{n,2,-2} = \boldsymbol{p}_1 - 2\boldsymbol{p}_2. \end{cases} \quad (5)$$

Here, $\boldsymbol{p}_3$ gave a direct estimate of d_n.

5. Results from fixed d_n simulations

5.1. Symmetric fits

Figure 3 shows a selection of results from fitting the symmetric-component model to the first cohort of artificial 3456-d sets. The $\ell = 2$ modes within these were all given the same frequency asymmetry, d_n. Plotted is the difference between the fitted frequency, $\nu^{\rm sym}_{n,2}$, and that of the underlying central (2, 0) component present, for three different input values of d_n (value in μHz given as title to each figure).

When the multiplet is asymmetric it is to be expected that the fitted frequency will differ from that of the central component, by an amount corresponding to $\approx d_n$. This is because the symmetric-fitted frequency, $\nu^{\rm sym}_{n,2}$, is determined largely by the stronger $|m| = 2$ components, whose mean frequency corresponds to:

$$\nu_{n,2,|2|} = \frac{1}{2}\left(\nu_{n,m=+2} + \nu_{n,m=-2}\right).$$
$$= \nu_{n,2,0} + d_n.$$

Were the asymmetrically aligned $m = 0$ mode to have no influence on the fitted frequency, the extracted $\nu^{\rm sym}_{n,2}$ would then correspond to the mean of the (2, 2) components. As such, it would be offset from the frequency of the central (2, 0) component by precisely d_n. The horizontal, dotted line in each panel of Fig. 3 shows the differences one would expect if this were indeed the case. That the actual differences depart from this suggests that the presence of the (2, 0) component "pulls" the fitted frequency, with the magnitude of the bias introduced depending upon both the size of the asymmetry present. Visual inspection of the plots in Fig. 3 suggests the bias depends also on the linewidth of the mode, $\Delta\nu_n$, i.e., the curves mimic the observed increase of linewidth with frequency. The dashed line in each panel is a simple model of the difference given by the scaled product of d_n and $\Delta\nu_n$. The fitted frequency is seen to be pulled in the direction of the centroid of the multiplet, e.g., for positive d_n the (2, 0) component lies closer to the (2, −2) mode and so, as a result, does the fitted $\nu^{\rm sym}_{n,2}$.

[1] It is worth noting that the various frequencies and splittings could of course be made from the parameters extracted by just one of the adopted asymmetric strategies. However, because of parameter correlations proper allowance must be made in computations of the associated uncertainties, e.g., by using the correlation information present in the Hessian matrix generated by each fit. We have instead chosen the more transparent route of defining models to extract the parameters we wish to study.

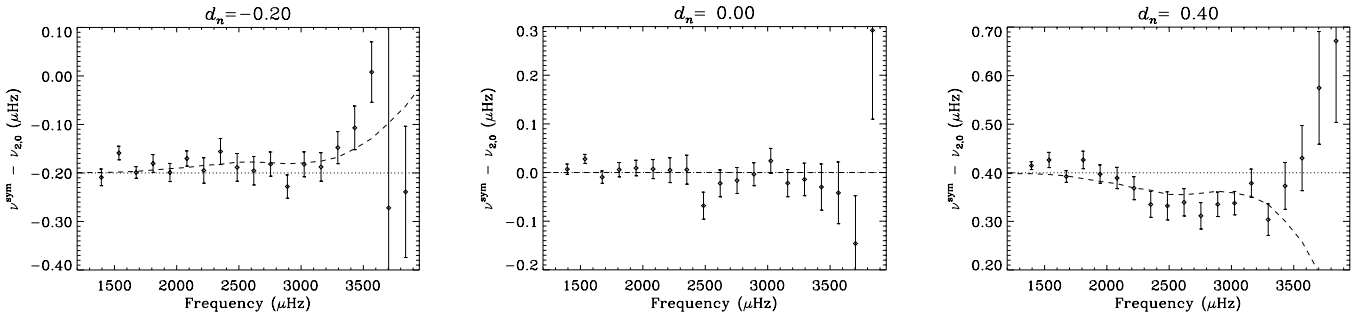

Fig. 3. Results from fitting the symmetric-component model to the first cohort of artificial 3456-d sets in which all $\ell = 2$ modes were given the same frequency asymmetry, d_n. Plotted is the difference between the fitted frequency and that of the underlying central (2, 0) component present for three different input values of d_n (see title of each figure). Also shown is the residual that would result were the $m = 0$ component to have *no* influence on the fitted frequency (dotted line); this corresponds to the input d_n. The dashed line is a prediction of the difference that actually arises because the fitted frequency is pulled by the $m = 0$ mode; it is proportional to the product of the input d_n and the linewidth, $\Delta\nu_n$.

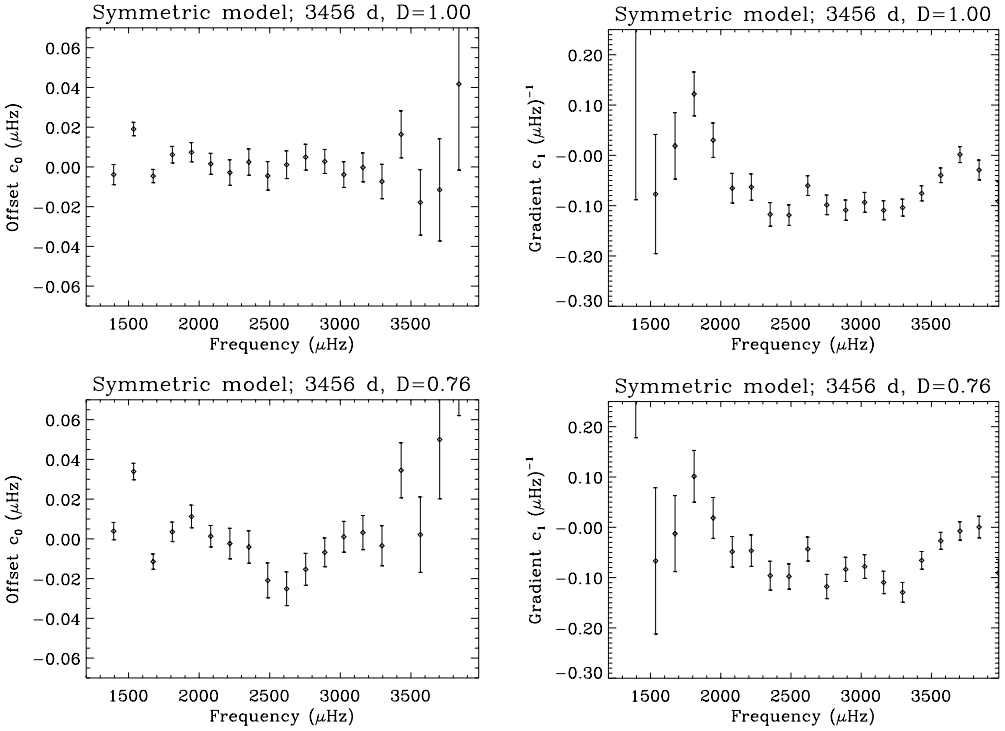

Fig. 4. The zeroth and first-order coefficients, c_0 and c_1, describing the linear relation between the difference in frequency $\nu^{sym}_{n,2} - \nu_{n,2,0}$, and the product of the input asymmetry, d_n, and mode linewidth, $\Delta\nu_n$. Estimates of the coefficients were extracted from the analysis of eleven complete 3456-d sets, having input d_n ranging from −0.5 to 0.5 μHz. Top row – results for datasets with fractional duty cycle $D = 1.00$. Bottom row – results after each dataset was modulated by the 3456-d BiSON window function, giving a fractional duty cycle of $D = 0.76$.

To properly test the dependence we assume the differences can indeed be modelled as a linear function of the product $\Delta\nu_n d_n$ (see above), plus the offset d_n. We then have:

$$\nu^{sym}_{n,2} - \nu_{n,2,0} = d_n + c_0 + c_1 \Delta\nu_n d_n, \quad (6)$$

where c_0 and c_1 are the zeroth and first-order coefficients to be extracted. The above can be rearranged to give an equation of the form

$$y = c_0 + c_1 x.$$

so that a linear fit of y against x yields the sought-for coefficients, i.e.,

$$\underbrace{\nu^{sym}_{n,2} - \nu_{n,2,0} - d_n}_{y} = c_0 + c_1 \underbrace{\Delta\nu_n d_n}_{x}. \quad (7)$$

We extracted estimates of the coefficients from four analyses: from fits to the eleven, full 3456-d sets, both with and without the BiSON window; and to the sets subdivided into contiguous, short 108-d lengths (again, both with and without the window). The results from the 3456-d analyses are plotted in the various panels of Fig. 4. The coefficients returned from the 108-d analyses were very similar and are therefore not plotted.

The zero-order coefficient, c_0, indicates the amount by which the fitted frequency differs from $\nu_{n,2,0}$ in the absence of any asymmetry in the multiplet (i.e., for $d_n = 0$). Under these circumstances the symmetric model provides a precise description of the frequency arrangement of the components, and any fit should in principle extract the underlying (2, 0) frequency. The lower four panels of Fig. 4 show that under

conditions of complete data fill (duty cycle $D = 1.00$) the offsets are indeed consistent with zero over the range of n fitted. As such, the correct frequency is given. However, when the window function is introduced this is no longer the case, and a negative non-zero offset is present that is peaked at $\sim$2600 μHz. Interestingly, the characteristics of this are similar in both the short 108-d, and long 3456-d, sets. Its location near the centre of the p-mode spectrum reflects, we believe, the deleterious impact of the complicated background that is introduced between the modes by the window (whose characteristics are discussed at length in Chaplin et al. 2003b).

The first-order coefficient, c_1, takes a reasonably constant value of $\approx -0.1\,(\mu\mathrm{Hz})^{-1}$ over the central part of the p-mode range (irrespective of the length of data fitted, and the presence or absence of the BiSON window). To understand the implications of this we re-arrange Eq. (6) slightly to:

$$\nu^{\mathrm{sym}}_{n,2} = \nu_{n,2,0} + c_0 + d_n[1 + c_1\Delta\nu_n]. \quad (8)$$

If $c_o \approx 0$, it then follows that the product $c_1\Delta\nu_n$ is the fractional amount by which the symmetric-fitted frequency differs from $\nu_{n,2,0} + d_n$. Put another way, it is a fractional measure of the extent to which the (2, 0) component determines (pulls) the fitted $\nu^{\mathrm{sym}}_{n,2}$. If $\mathcal{S}_n$ is the fractional contribution of the (2, 0) component to the symmetric-model frequency, i.e.,

$$\nu^{\mathrm{sym}}_{n,2} = \mathcal{S}_n\nu_{n,2,0} + (1 - \mathcal{S}_n)\nu_{n,2,|2|}, \quad (9)$$

the following is then a quantitative statement of the above:

$$\mathcal{S}_n = -\frac{c_0 + d_n c_1\Delta\nu_n}{d_n} \approx -c_1\Delta\nu_n. \quad (10)$$

Since the linewidth at the centre of the spectrum is $\approx 1\,\mu$Hz, this suggests that $\mathcal{S}_n \approx 10\%$. So, even though the (2, 0) mode has a height that is $\sim$55% of that of its (2, 2) neighbours, its influence on the fitted frequency is somewhat less than is implied by the visibility ratio.

At frequencies above $\sim$3200 μHz the widths $\Delta\nu_n$ rise sharply with the result that adjacent components suffer from substantial blending. It then becomes increasingly difficult to determine accurately the presence of a non-zero d_n. This explains why the high-frequency values of c_1 tend toward the zero level.

At very low n the differences in frequency are so small that it is difficult to extract a reliable measure of the c_1. Furthermore, beating with noise under conditions of low signal to noise, and the influence of realization noise when the components are confined to very few frequency bins, can influence dramatically the estimates. As such, the values plotted below $\sim$2000 μHz in Fig. 4 should be regarded with caution. Regardless, since these peaks are so narrow fits of different models to the modes can be judged more easily on their individual merit, i.e., if all three components are clearly visible it will be obvious if the fitted-model components do not lie on top of them. This is a luxury not available at higher frequencies.

5.2. Asymmetric fits

In Fig. 5 we show results from fitting the asymmetric models to the same cohort of artificial data (this time on just the 3456-d length scale).

The top two rows show the difference between the asymmetric-model-fitted frequencies and the input (2, 0) frequency. The panels in the top row show results from fitting for the mean frequency of the $|m| = 2$ components, $\nu^{\mathrm{ass}}_{n,2,|2|}$. Since the fitting model now describes the asymmetric arrangement of the components, $\nu^{\mathrm{ass}}_{n,2,|2|}$ should correspond to $\nu_{n,2,0} + d_n$ (prediction shown as the dotted line in each figure) and not be offset from this, as was the case for the symmetric-model fits. The results in the second row are from fits for the (2, 0) component, $\nu^{\mathrm{ass}}_{n,2,0}$.

The results for all input values of d_n suggest that over much of the frequency range unbiased estimates of both the (2, 0) and mean (2, |2|) frequencies are extracted (at the level of precision of the data). At frequencies above $\approx$3500 μHz the blending of adjacent components means the fitted estimates show larger scatter and are poorly constrained. The implication here is that an asymmetric model can be applied successfully to data of this length and quality (except at very high n) with the result that more accurate estimates of the frequencies are obtained (i.e., in comparison to those given by the symmetric-model approach). However, this comes at the cost of reduced precision in the estimates.

The bottom row of Fig. 5 plots ratios of the fitted frequency errors from the asymmetric and symmetric fits. Stars show the ratio of the uncertainties in the fitted (2, |2|) frequency to those in the symmetric-fitted $\nu^{\mathrm{sym}}_{n,2}$; and diamonds the ratio of uncertainties in the fitted (2, 0) to those in $\nu^{\mathrm{sym}}_{n,2}$. Over most of the frequency range the (2, |2|) estimates from the asymmetric model have similar (slightly inferior) precision to their symmetric-fitted counterparts. Only at high frequencies – where the asymmetric fits show poor convergence – does the ratio increase noticeably. In contrast, the precision in the fitted (2, 0) is significantly inferior, showing a steady deterioration that follows approximately the increase in linewidth.

6. Results from time-varying d_n simulations

Figure 6 shows results from fitting a 3456-d set into which a time-dependent variation of the $\ell = 2$ component frequencies was introduced (see Sect. 3). This was done in such a way as to match approximately the predictions of Dziembowski et al. (2000), and to therefore provide a proxy for the content of real data.

In the top row we plot the difference between the symmetric-model fitted frequency and that of the actual mean of the (2, 0) component over the 3456-d simulated period. As for the plots in Fig. 3 we show also the difference that would result were the (2, 0) component to have had no influence on the extracted value; the fitted frequency would then be that of the mean of the $|m| = 2$ modes, meaning the offset would correspond to the mean d_n for the simulated epoch (rendered as the solid line). Also shown (dashed line) is the prediction from Eq. (8) of the "pulled" frequency that actually results. To make the latter prediction – which the fitted values follow closely, both with and without the window function – we assumed a fixed c_1 of $-0.1\,(\mu\mathrm{Hz})^{-1}$ at all n.

The second row shows the result of fitting the asymmetric model to extract estimates of the mean (2, |2|) frequencies, with the solid line following the expected offset (that of the

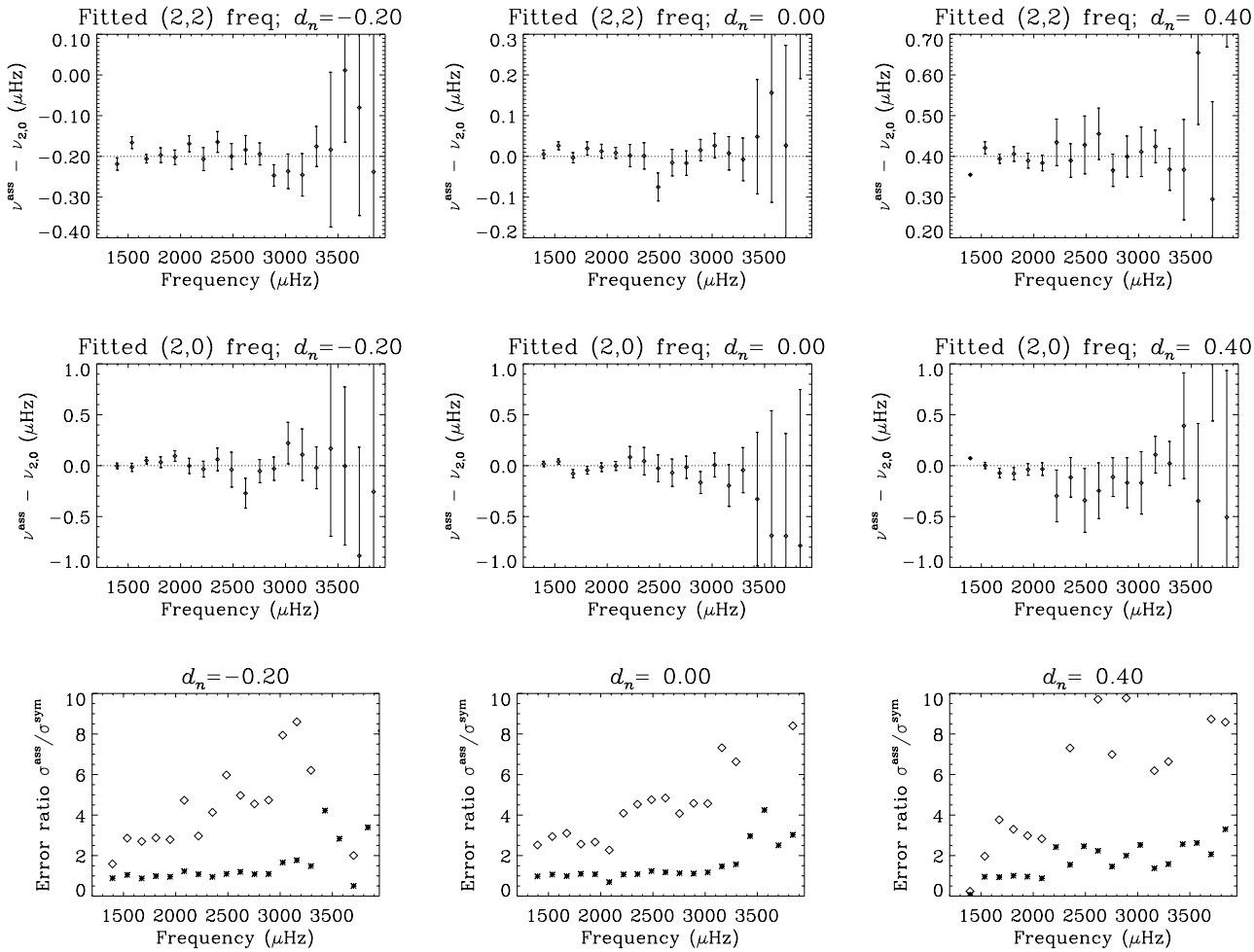

Fig. 5. Results from fitting asymmetric-component models to the first cohort of artificial 3456-d sets in which all $\ell = 2$ modes were given the same frequency asymmetry, d_n. Top row – difference between the fitted mean frequency of the $|m| = 2$ components, $\nu^{ass}_{n,2,|2|}$, and the input (2, 0) values. Middle row – difference between the fitted ($\nu^{ass}_{n,2,0}$) and input (2, 0) frequencies. Bottom row – ratios of the fitted frequency errors from the asymmetric and symmetric fits. Stars show the ratio of the uncertainties in the fitted $\nu^{ass}_{n,2,|2|}$ to those in the symmetric-fitted $\nu^{sym}_{n,2}$; and diamonds the ratio of the uncertainties in the asymmetrically fitted $\nu^{ass}_{n,2,0}$ to those in $\nu^{sym}_{n,2}$.

mean d_n). The third row shows the residuals from fitting directly for the (2, 0) frequency.

Again, we find that both asymmetric strategies extract reliable estimates of the sought-for parameters below ∼3500 μHz. The impact of the window function is manifested only by a modest increase in the size of the uncertainties (compared to fits made to continuous data). This observed increase is consistent with predictions based upon the reduction in signal to noise caused by the window (see Chaplin et al. 2002). The duplication of patterns in the fitting with and without the window (compare left and right-hand panels in Fig. 6) suggests they are the result of realization noise.

Because the average, underlying asymmetry over the 10-yr period is very small for most of the modes (solid line in upper four panels), the symmetric-model-fitted and asymmetric $(2, |2|)$ fitted frequencies returned are very similar. Only above ∼3200 μHz do the two depart from one another, and even then the observed differences are not significant given the level of precision in the fits. At frequencies above 3500 μHz the asymmetric fits are poorly constrained.

The difference between the symmetric-model fitted and underlying (2, 0) frequency is, however, significant over the central part of the spectrum. To illustrate this we plot in the bottom panel of Fig. 6 the difference normalized by the uncertainty in the symmetric fitted frequency (triangles joined by dashed line). This is greater than 4σ between ∼2900 and 3400 μHz.

Also shown is the difference between the asymmetric-model fitted (2, 0) frequency and the actual underlying value, again normalized by the fitted error (diamonds joined by dotted line). Over the fitted range it is apparent that the (2, 0) estimates are accurate, at the level of precision of the fits; however, the fitted uncertainties are much larger than than those in the symmetric-model frequencies (in line with the results in Sect. 5).

7. Conclusions and implications for real data

We have studied in detail the extraction of estimates of $\ell = 2$ p-mode frequencies from unresolved observations of the visible disc of the Sun. This is complicated in practice by the

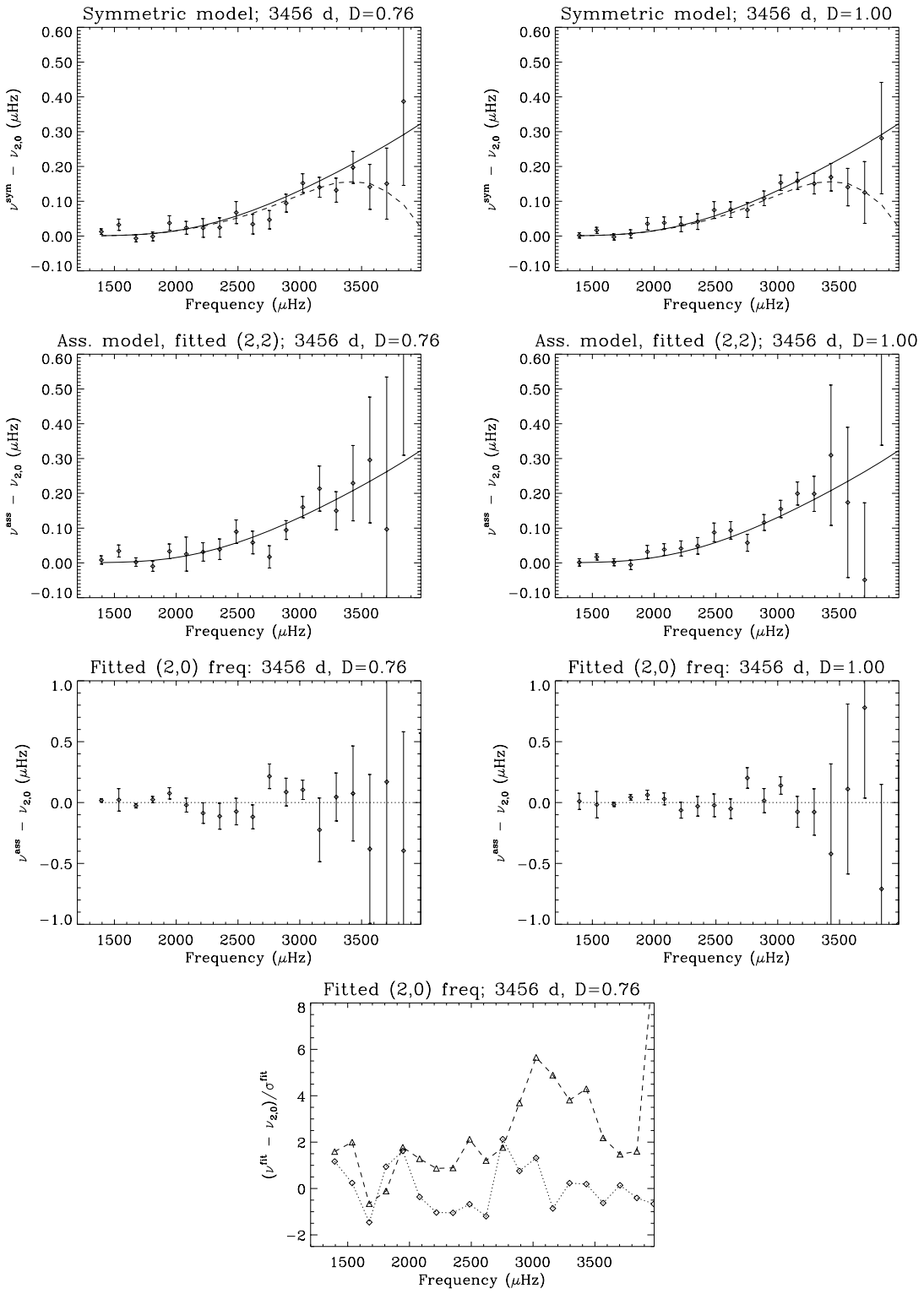

Fig. 6. Results from fitting a 3456-d set – both with and without the BiSON window function – into which a time-dependent variation of the $\ell = 2$ component frequencies was introduced (see Sect. 3). Top row – difference between the symmetric-model fitted frequency and that of the underlying central (2, 0) component. Also shown is the residual that would result were the $m = 0$ component to have *no* influence on the fitted frequency (solid line); this corresponds to the input d_n. The dashed line is a prediction of the difference that is proportional to the product of the input d_n and the linewidth, $\Delta\nu_n$. Second row – difference between the asymmetric-model fitted mean frequency of the $|m| = 2$ components, $\nu^{\rm ass}_{n,2,|2|}$, and the input (2, 0) values. Third row – difference between the asymmetric-model fitted $\nu^{\rm ass}_{n,2,0}$ and input (2, 0) frequencies. Lower panel – difference between the symmetric-fitted frequency and the actual (2, 0) frequency, normalized by the fitted uncertainty (triangles); same, but for the asymmetrically fitted (2, 0) and actual (2, 0) frequency (diamonds).

asymmetric arrangement in frequency of the three components ($m = -2$, 0 and 2) that are prominent in such data.

To effect our study we constructed, and then analyzed, a series of 10-yr artificial datasets into which varying degrees of asymmetry, d_n, were introduced: first, at levels that were held constant over the duration of the simulated observations; and then as a time-varying parameter designed to mimic the observed effect of the solar activity cycle. Our principal conclusions are as follows:

1. It has for a long time been standard practice to fit the $\ell = 2$ modes to a model that assumes the components are arranged symmetrically. When asymmetric data are fitted to a model of this type, the fitted frequency for multiplets over the central part of the spectrum is given by:

   $\nu_{n,2}^{\rm sym} \sim 0.1\nu_{n,2,0} + 0.9\nu_{n,2,|2|}$,

   where $\nu_{n,2,0}$ is the frequency of the inner, (2, 0) component and $\nu_{n,2,|2|}$ is the mean of that of the outer, $|m| = 2$ components. The fitted value – although determined largely by the more-prominent (2, 2) and (2, −2) modes – is therefore "pulled" toward the (2, 0) component.
   - Dataset segments of length of 108 d, or similar, are often used in studies of the solar activity cycle. Because of their short lengths sets made from observations coincident with high levels of activity may have large $\ell = 2$ frequency asymmetries present within them. However, they offer insufficient resolution to extract reliable estimates of either the asymmetry itself or the frequency of the (2, 0) component. The impact of the asymmetry will therefore be most conspicuous at times of high activity. The fitted frequency of modes from ∼2500 to ∼3500 μHz may differ from: that of the underlying (2, 0) component by as much as ∼0.3 μHz; and from that of the mean of the $m = 2$ components by ∼0.03 μHz. This is to be compared with typical 108-d frequency uncertainties for the most prominent quadrupole modes of ≈0.20 μHz.
   - When the symmetric model is applied to dataset lengths of ∼10 yr the difference between the fitted and $|m| = 2$ frequency is unlikely to be significant in spite of the vastly improved precision (frequency uncertainties now ≈0.03 μHz). This is because the underlying asymmetry present is an *average* over the minimum-to-maximum range of activity, and therefore smaller in magnitude than for a 108-d period at or around cycle maximum. Our analyses suggest also that the difference between the symmetric-fitted and underlying (2, 0) frequencies is significant for modes at the centre of the spectrum. Although the difference increases at higher frequencies, it is less significant there owing to the increase in the frequency uncertainties.
2. We also tested asymmetric models formulated to extract the locations of all three modes present. Our results demonstrate that for time-series lengths comparable to the 11-yr cycle, and data quality similar to BiSON or GOLF, robust, unbiased estimates of the frequencies of the (2, 0) and (2, |2|) components can be extracted both with and without the BiSON window. At frequencies above ∼3500 μHz the increased linewidth of the modes gives rise to substantial blending between the components with the result that the asymmetric procedure shows poor convergence, yielding weakly constrained estimates of the parameters. While direct estimates of the (2, 0) frequency are encouragingly accurate across a wide fitting range – irrespective of whether or not the BiSON window is applied – the associated uncertainties are approximately twice as large as those from the symmetric-model for modes at ∼1500 μHz, some five times larger at ∼3000 μHz, and roughly an order of magnitude larger at ∼4000 μHz.

Acknowledgements. This work utilizes helioseismic data collected by the ground-based Birmingham Solar-Oscillations Network (BiSON), and the SOI/MDI instrument on board the SOHO satellite. BiSON is funded by the UK Particle Physics and Astronomy Research Council. SOHO is a project of international cooperation between ESA and NASA.

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 and P Whitelock 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. We are also grateful to T. Bedding and colleagues for providing a pre-print of their paper.

Appendix A: Comparison with resolved observations

The full set of m frequencies extracted from resolved-Sun observations are often described using a polynomial expansion:

$$\nu_{n\ell,m} = \nu_{n\ell}^{\rm cen} + \sum_{j=1}^{j_{\rm max}} a_j(n,\ell)\,\ell\,\mathcal{P}_\ell^j(m). \quad \text{(A.1)}$$

Where the basis functions are polynomials that are related to Clebsch-Gordan coefficients (Ritzwoller & Laveley 1991).

The offset term in the expansion, $\nu_{n\ell}^{\rm cen}$, is the so-called central frequency of the multiplet (i.e., its centroid). The even a_j coefficients reflect perturbations that are non-spherically symmetric in nature. The d_n for unresolved data can therefore be made from an appropriate combination of the a_j. Those terms needed to describe an $\ell = 2$ multiplet [with $L^2 = \ell(\ell+1)$] are:

$$\mathcal{P}_\ell^1(m) = m/\ell,$$

$$\mathcal{P}_\ell^2(m) = \frac{6m^2 - 2L^2}{6\ell^2 - 2L^2},$$

$$\mathcal{P}_\ell^3(m) = \frac{20m^3 - 4m(3L^2-1)}{20\ell^3 - 4\ell\,(3L^2-1)},$$

and

$$\mathcal{P}_\ell^4(m) = \frac{70m^4 - 10m^2\left(6L^2-5\right) + 6L^2\left(L^2-2\right)}{70\ell^4 - 10\ell^2\,(6L^2-5) + 6L^2\,(L^2-2)}.$$

From these it is straight forward to show that the frequencies of the three components visible in unresolved observations are given by:

$$\nu_{n,2,2} = \nu_{n\ell}^{\rm cen} + 2a_1(n,2) + 2a_2(n,2) + 2a_3(n,2) + 2a_4(n,2), \quad \text{(A.2)}$$

$$\nu_{n,2,-2} = \nu_{n\ell}^{\rm cen} - 2a_1(n,2) + 2a_2(n,2) - 2a_3(n,2) + 2a_4(n,2), \quad \text{(A.3)}$$

and

$$\nu_{n,2,0} = \nu_{n\ell}^{\rm cen} - 2a_2(n,2) + 12a_4(n,2). \quad \text{(A.4)}$$

It then follows that the unresolved-data frequency asymmetry is given by:

$$d_n = 4a_2(n,2) - 10a_4(n,2). \quad \text{(A.5)}$$

References

Appourchaux, T., Anderson, B., & Sekii, T. 2002, in SOHO 11 Symposium From Solar Min to Max: Half a Solar Cycle with SOHO, ed. C. Fröhlich, & A. Wilson, ESA SP-508, 47

Bertello, L., Henney, C. J., Ulrich, R. K., et al. 2000, ApJ, 535, 1066

Butler, R. P., Bedding, T. R., Kjeldson, H., et al. 2003, ApJ, submitted

Chaplin, W. J., Elsworth, Y., Howe, R., et al. 1997, MNRAS, 287, 51

Chaplin, W. J., & Appourchaux, T. 1999, MNRAS, 309, 761

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

Chaplin, W. J., Elsworth, Y., Isaak, G. R., et al. 2001a, MNRAS, 327, 1127

Chaplin, W. J., Appourchaux, T., Elsworth, Y., Isaak, G. R., & New, R., 2001b, MNRAS, 324, 911

Chaplin, W. J., Elsworth, Y., Isaak, G. R., et al. 2002, MNRAS, 330, 731

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

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

Chaplin, W. J., Elsworth, Y., Isaak, G. R., Miller, B. A., & New, R. 2003c, MNRAS, submitted

Christensen-Dalsgaard, J. 1989, MNRAS, 239, 977

De Toma, G., White, O. R., & Harvey, K. L. 2000, ApJ, 529, 1101

Dziembowski, W. A., & Goode, P. R. 1997, A&A, 317, 919

Dziembowski, W. A., Goode, P. R., Kosovichev, A. G., & Schou, J. 2000, ApJ, 537, 1026

Moreno-Insertis, F., & Solanki, S. K. 2000, MNRAS, 313, 411

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

Ritzwoller, M. H., & Laveley, E. M. 1991, ApJ, 369, 557

Tapping, K. F., & DeTracey, B. 1990, Sol. Phys., 127, 321

Thiery, S., Boumier, P., Gabriel, A. H., et al. 2000, A&A, 355, 743

Toutain, T., Appourchaux, T., Fröhlich, C., et al. 1998, ApJ, 506, L147

Toutain, T., & Kosovichev, A. G. 2000, ApJ, 534, L211