A MODE FITTING METHODOLOGY OPTIMIZED FOR VERY LONG TIME SERIES

S. G. Korzennik

Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, U.S.A.

ABSTRACT

I describe and present the results of a newly developed fitting methodology optimized for very long time series. The development of this new methodology was motivated by the fact that we now have more than half a decade of nearly uninterrupted observations by GONG and MDI, with fill factors as high as 89.8% and 82.2% respectively. The fitting procedure uses an optimal sine-multi-taper spectral estimator – where the number of tapers is based on the mode linewidth – the complete leakage matrix (*i.e.*, horizontal as well as vertical components), and an asymmetric mode profile to fit simultaneously all the azimuthal orders with individually parameterized profiles. This method was applied to 2088-day-long time series of MDI and GONG observations, as well as 728-day-long subsets, and for spherical harmonic degrees between 1 and 25. The values[1] resulting from these fits are inter-compared (MDI versus GONG) and compared to equivalent estimates from the MDI team and the GONG project.

1. MOTIVATION

I have developed a new methodology to fit modes using very long time series, since we now have access to more than half a decade of nearly uninterrupted times series of solar observations. This methodology includes an "optimal" sine-multi-taper spectral estimator, the complete leakage matrix (*i.e.*, horizontal as well as vertical components), an asymmetric profile and the simultaneous fitting of individual profiles at all the azimuthal orders (m) for a given (n, ℓ) mode.

The primary goal for this work was to extend the mode fitting to low-order and low-degree modes. The low degree modes carry information on the structure and dynamics of the deep interior while the low-order low-degree modes are long lived (*i.e.*, show very narrow peaks) and can thus be measured with very high precision.

---

[1]The tables of mode frequencies are available at: `ftp://cfa-ftp.cfa.harvard.edu/pub/sylvain/tables/`

2. DATA SETS USED

I used times series of spherical harmonic coefficients computed from full-disk observations by the MDI and GONG instruments and limited to $\ell \leq 25$. The MDI 2088-day-long time series was used to develop the methodology. It was then applied to the co-eval GONG data set. Both data sets were also subdivided in five 728-day-long overlapping segments each offset by some 364 days from the previous one. The fill factors of the 2088-day-long time series, before detrending, are 89.8% and 82.2% for MDI and GONG observations respectively.

3. METHODOLOGY

The key original aspects of this new methodology are the use of an "optimal" sine-multi-tapered spectral estimator, the simultaneous fitting of all m spectra and the use of an asymmetric profile. The fitted model includes both radial and horizontal leakage components and the leakage of nearby modes inside the fitting range.

I used the Nth order sine-multi-taper as power spectrum estimator, defined as

$$P_{\ell,m}^{(N)}(\nu) = \sum_{k=1}^{N} \left| \text{FFT}\left[\sin(\frac{\pi\, k\, i}{M+1}) c_{\ell,m}(t_i) \right] \right|^2 \quad (1)$$

where $c_{\ell,m}(t_i)$ represents the spherical harmonic coefficient for ℓ and m at the time t_i, and M is the length of the time series. These sine-multi-tapered power spectra were computed with an oversampling factor of 2, and for a pre-selected list of number of tapers. For the 2088-day-long time series, I used 5, 9, 21, 45 and 91 tapers, while for the 728-day-long time series I used 3, 7, 15, 31 and 63 tapers. The choice of the optimal number of tapers amongst that list is explained below.

I used a downhill simplex minimization to fit simultaneously, and in the least-squares senses, all the multiplets for a given mode, *i.e.* all m for a given n, ℓ. The fitting is done iteratively, over a frequency range limited to only encompass the closest spatial leaks ($\delta m = \pm 2, \delta \ell = 0$),

using the "optimal" sine-multi-taper power spectrum and fitting only for the modes whose amplitudes are above some prescribed threshold.

The fitted profile is an asymmetric Lorentzian:

$$P_{n,\ell,m}(\nu) = \frac{1 + \alpha_{n,\ell}(x_{n,\ell,m} - \alpha_{n,\ell}/2)}{x_{n,\ell,m}^2 + 1} \quad (2)$$

where

$$x_{n,\ell,m} = \frac{\nu - \nu_{n,\ell,m}}{\Gamma_{n,\ell}/2} \quad (3)$$

The power spectrum is thus modeled as the superposition of the mode profile and the spatial leaks present in the fitting range:

$$
\begin{aligned}
P_{\ell,m}(\nu) &= A_{n,\ell,m}P_{n,\ell,m}(\nu) + B_{n,\ell,m} \quad (4) \\
&+ \sum_{m'} A_{n,\ell,m'}C(n,\ell,m';n,\ell,m)P_{n,\ell,m'}(\nu) \\
&+ \sum_{n',\ell',m'} A_{n',\ell',m'}C(n',\ell',m';n,\ell,m)P_{n',\ell',m'}(\nu)
\end{aligned}
$$

The sum on m' is actually limited to $m' = m - 2$ and $m' = m+2$ by the choice of the fitting range. The sum on n', ℓ', m' is included only if a nearby mode ($|\ell - \ell'| \leq 3$ & $|m - m'| \leq 3$) fall within the widened fitting window.

The fitting range is set to be $\tilde{\nu}_{n,\ell,m} \pm \delta\nu$, where $\delta\nu$ is given by

$$\delta\nu = 4\,\Gamma_{n,\ell}^{(\text{eff})} + \Delta\nu \quad (5)$$

where $\Gamma^{(\text{eff})}$ is the effective mode linewidth and $\Delta\nu = 800$ nHz.

The effective mode linewidth, $\Gamma_{n,\ell}^{(\text{eff})}$, is estimated by

$$(\Gamma_{n,\ell}^{(\text{eff})})^2 = \tilde{\Gamma}_{n,\ell}^2 + \Gamma_{r,N}^2 \quad (6)$$

where $\tilde{\Gamma}_{n,\ell}$ is some estimate of the mode linewidth and $\Gamma_{r,N}$ is the resolution of the Nth order multi-taper power spectrum, given by

$$\Gamma_{r,N} = N\,\Gamma_r = \frac{N}{T} \quad (7)$$

The "optimal" N^{th} order sine-multi-taper power spectrum is defined as the highest order sine-multi-taper spectrum from a pre-selected list having a resolution at least five times better than the effective mode linewidth, whenever possible.

The value of N is thus selected to satisfy:

$$\frac{\tilde{\Gamma}_{n,\ell}}{5} \geq \Gamma_{r,N} = N\,\Gamma_r \quad (8)$$

At low frequencies, the mode amplitude becomes comparable if not smaller than the background noise all the
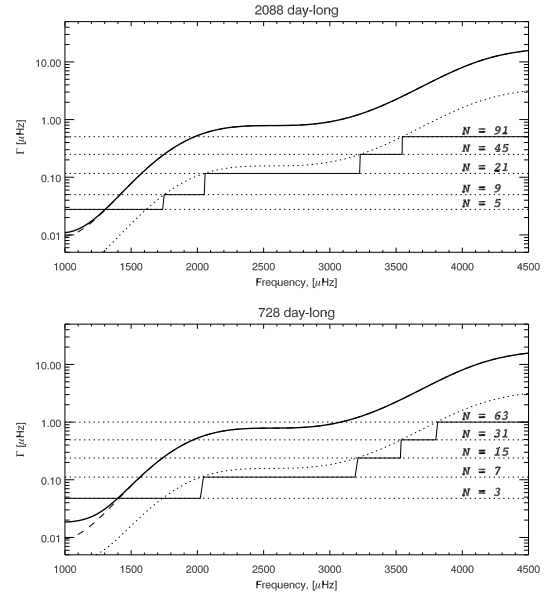

Figure 1. Mode linewidth (long dash) and effective linewidth (solid bold curve), compared to multi-tapers spectral resolution (dotted lines). The optimal number of multi-tapers (see description in text) is indicated by the stepwise solid line, and is such that the spectral resolution remains 5 times smaller than the effective linewidth, whenever possible. Top and bottom panels correspond to 2088-day-long and 768-day-long time series, respectively.

while the mode linewidth becomes smaller than the spectral resolution, forcing the fitting procedure to hunt for small and narrow peaks. Only modes with a power amplitude 3 times greater than the RMS of the residuals to the fit were kept. As the fitting proceeds a sanity check rejects any mode whose amplitude has dropped below that threshold.

I used leakage matrices computed by J. Schou, for both MDI and GONG observations. Both, horizontal and vertical components were computed and the horizontal to vertical displacement ratio, β, taken to be the theoretical prediction.

Error bars were estimated from the covariance matrix of the problem, itself based on the Hessian matrix computed using numerical estimates of the second derivative of the merit function.

4. RESULTS

Examples of fitting of the MDI 2088-day-long time series for $\ell = 9$ and a selection of values of n are presented in the various panels of Fig. 2.

Fig. 3 compares the fitting results from the two 2088-day-long co-eval MDI and GONG time series. Comparisons
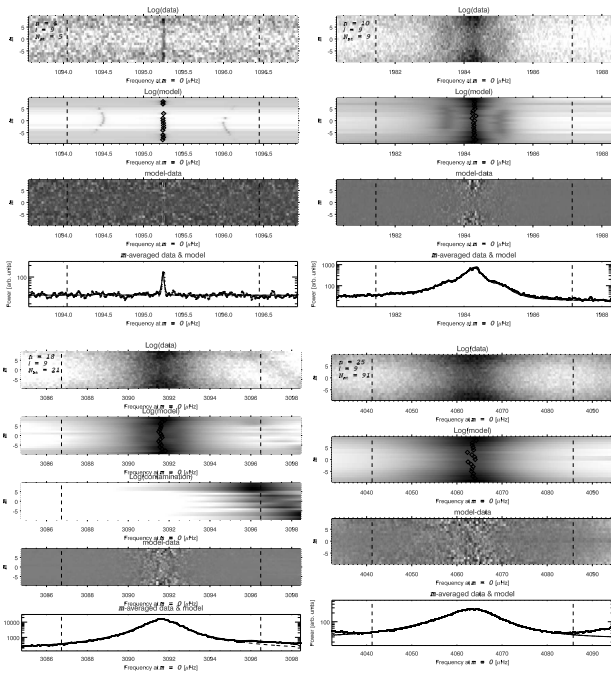
Figure 2. Examples of fitting, for MDI 2088-day-long time series, for $\ell = 9$ and for various n & N. Notice how for $n = 4$ & $N = 5$ some of the modes amplitudes were not large enough to be fitted; how for $n = 10$ & $N = 9$ the closest spatial leaks ($\delta m = \pm 2$, $\delta \ell = 0$) are barely resolved and blend with the main peak in the m-averaged spectrum; for $n = 18$ & $N = 21$ the contamination by $n' = n \pm 1$, $\ell' = \ell \mp 3$; and for $n = 25$ & $N = 91$ that mode linewidth is so large that a very large number of tapers has been used.
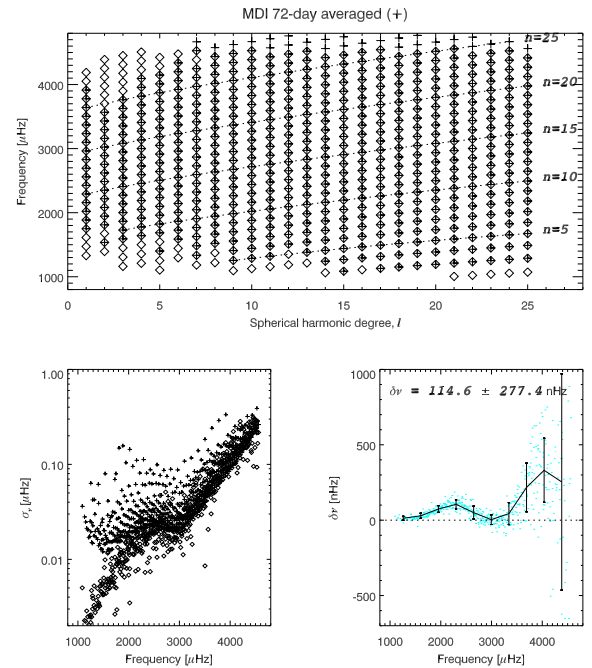

Figure 4. Comparison between singlets resulting from fitting the 2088-day-long MDI time series (diamonds) and the corresponding MDI average values computed from 27 tables resulting from fitting 72-day-long times series. Top panel shows the respective coverage in an $\ell - \nu$ diagram. The lower left panel compares the frequency uncertainties, while the lower right shows the frequency differences (dots) and these differences binned over 10 equispaced frequency bins – the error bars represent the standard deviation inside each bin.

Figure 3. Comparison of fitting results (singlets) from the two 2088-day-long co-eval MDI and GONG time series.

for the five 728-day-long segments show similar results. A more extensive presentation of these results will be soon available in Korzennik (2004).

4.1. Comparisons with Previous Estimates

Fig. 4 compares singlets resulting from fitting the 2088-day-long MDI time series (diamonds) and the corresponding MDI average values computed from 27 tables resulting from fitting 72-day-long times series (Schou, 1999). It shows a systematic difference between frequencies, with a specific frequency dependence. One obvious reason for this difference is the fact that Schou's fitting uses a symmetric profile while I am fitting an asymmetric one.

Fig. 5 compares results from my fit to the GONG 2088-day-long time series to average values based on 58 tables resulting from fitting 108-day-long times series that covers the same time span and are routinely computed by the GONG project (Hill et al., 1996). The lower SNR of the GONG observations at low frequency does not allow to push mode fitting down to orders as low as for the MDI observations. But using a long time series still allowed
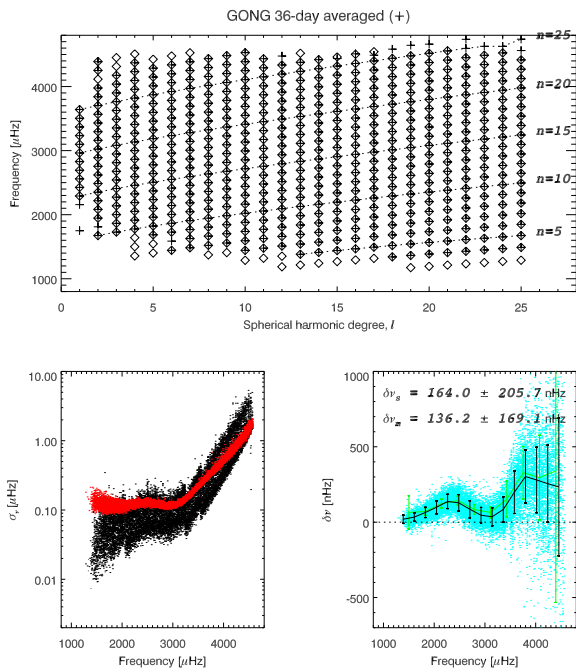
Figure 5. Comparison between results from fitting the 2088-day-long GONG time series (diamonds or black dots) and the corresponding GONG average values computed from 58 tables resulting from fitting 108-day-long times series (crosses or red dots).
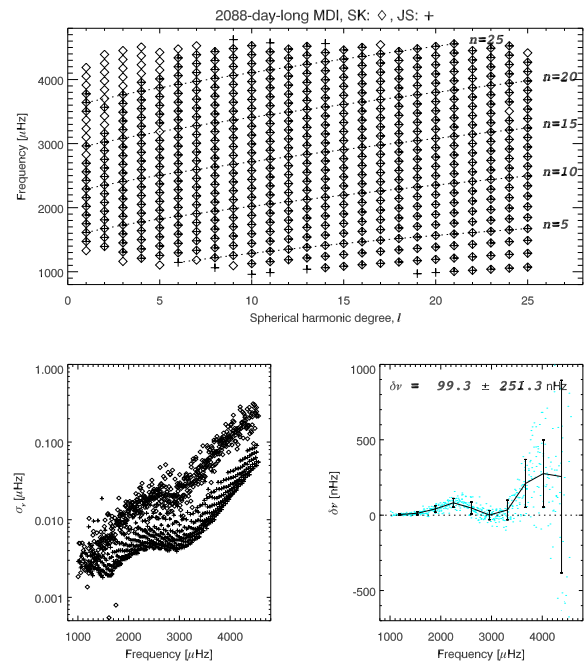

Figure 6. Comparison of MDI singlets resulting from this work (crosses) and from Schou's fitting to the same 2088-day-long time series (diamonds). Coverage in the $\ell - \nu$ diagram is very similar. My estimate of frequency uncertainties appears too conservative while the frequency differences show a systematic pattern with frequency.

me to fit low order modes that are rarely (less than 10 out of 58 times) fitted by the GONG project.

Back in 1991, Schou (private communication) fitted one 72-day-long MDI time series using an asymmetric profile as well as a symmetric one. Comparison of his asymmetric versus symmetric fits, when limited to $\ell \leq 25$ shows the very same residual systematic difference. More recently, Schou (private communication, 2004) has carried out a fitting of the 2088-day-long time series as well – using the same methodology he uses for fitting the 72-day-long time series. Comparison with the results of that fit is show in Fig. 6.

To estimate the effect of fitting a symmetric profile to an asymmetric peak I have computed a grid of isolated asymmetric profiles and fitted them with symmetric ones. The resulting offset in frequency varies nearly linearly with the asymmetry coefficient, as defined in my parameterization, for a given FWHM. The systematic error introduced by fitting an asymmetric peak with a symmetric profile is thus given by

$$\nu_{\mathrm{asymmetric}} - \nu_{\mathrm{symmetric}} = -\alpha\,\Gamma/2 \qquad (9)$$

Roughly half of the frequency differences seen in Figs. 4, 5 and 6 can be explained by this model. This simple model is unable to reproduce the systematic difference seen in Schou's symmetric versus asymmetric fits.

Nevertheless, residual differences after correcting for the

asymmetry using Eq. 9, are marginally significant, *i.e.* at the 4σ level, but still show a systematic trend with frequency.

ACKNOWLEDGMENTS

REFERENCES

Hill, F., et al. 1996, Science, 272, 1292

Korzennik, S.G., 2004, *submitted to ApJ*

Schou, J. 1999, ApJL, 523, L181