A&A manuscript no.
(will be inserted by hand later)

Your thesaurus codes are:
06(03.13.2; 03.13.6; 03.13.5; 06.15.1)

ASTRONOMY AND ASTROPHYSICS

The art of fitting p-mode spectra: Part II. Leakage and noise covariance matrices

Thierry Appourchaux[1], **Maria-Cristina Rabello-Soares**[1,2], **Laurent Gizon**[1,3]

[1] Space Science Department of ESA, ESTEC, NL-2200 AG Noordwijk
[2] Teoretisk Astrofysik Center, Århus Universitet, DK-8000 Århus C
[3] W.W.Hansen Experimental Physics Laboratory, Center for Space Science and Astrophysics, Stanford University, Stanford, CA 94305-4085, USA

Received 15 October 1997 / Accepted 24 February 1998

Abstract. In Part I we have developed a theory for fitting p-mode Fourier spectra assuming that these spectra have a multi-normal distribution. We showed, using Monte-Carlo simulations, how one can obtain p-mode parameters using 'Maximum Likelihood Estimators'. In this article, hereafter Part II, we show how to use the theory developed in Part I for fitting real data. We introduce 4 new diagnostics in helioseismology: the (m,ν) echelle diagram, the cross echelle diagram, the inter echelle diagram, and the cross spectrum ratio. These diagnostics are extremely powerful to visualize and understand the covariance matrices of the Fourier spectra, and also to find bugs in the data analysis code. The diagrams are used to verify the computation of the leakage matrices, and also to measure quantitatively these matrices. Cross spectrum ratios are used to obtain quantitative information on the noise covariance matrices. Numerous examples using the LOI/SOHO and GONG data are given.

Key words: Methods: data analysis – statistical – observational – Sun: oscillations

1. Introduction

The physics of the solar interior is known from inversion of solar p-mode frequencies and splittings. These measurements are derived from fitting p-mode Fourier spectra. Schou (1992) was the first one to assume a multi-normal distribution for p-mode Fourier spectra and using a real leakage matrix. Following this pioneering work, Appourchaux et al. (1997) (hereafter Part I), generalized the theoretical background for fitting p-mode Fourier spectra to complex leakage matrix, and included explicitly the correlation of the noise between the Fourier spectra. Using Monte-Carlo simulations, we showed that our fitted parameters were unbiased. We also studied systematic errors due to an imperfect knowledge of the leakage covariance matrix. Unfortunately, a theoretical knowledge of fitting data is not enough as only real data will teach us if our approach is correct. Contrary to fitting p-mode power spectra, the process of fitting the Fourier spectra as described in Part I is rather difficult to understand and visualize. Schou (1992) gave a few diagnostics for understanding how the Fourier spectra are fitted but without showing an easy way to visualize the covariance matrices.

In this paper, we show how one can easily visualize the mode and noise correlation matrices, and then derive the mode leakage matrix. In the first section, we describe 4 new diagrams that have various diagnostics power. In the second section, we describe how we use those diagrams for inferring the leakage and noise covariance matrices for the data of the Luminosity Oscillations Imager (LOI) on board the Solar and Heliospheric Observatory (SOHO) data (Appourchaux et al, 1997), and for the data of the Global Oscillations Network Group (GONG) (Hill et al, 1996). The LOI time series starts on 27 March 1996 and ends on 27 March 1997 with a duty cycle greater than 99%. The GONG time series starts on 27 August 1995 and ends on 22 August 1996 with a 75% duty cycle. In the last section we conclude by emphasizing the usefulness of these diagrams.

2. Diagnostics for helioseismic data analysis

The echelle diagram was first introduced by Grec (1981). It is based on the fact that the low-degree modes are essentially equidistant in frequency for a given l; the typical spacing for $l = 0$ is 135 μHz. The spectrum is cut into pieces of 135 μHz which are stacked on top of each other. Since the modes are not truly equidistant in frequency, the echelle diagram shows up power as distorted ridges; an example is given in Fig. 1 for the LOI/SOHO instrument seeing the Sun as a star.

Another useful diagram was introduced by Brown (1985), the so-called (m,ν) diagram which shows how the frequency of an l,m mode depends upon m. Most often this

Send offprint requests to: thierrya@so.estec.esa.nl

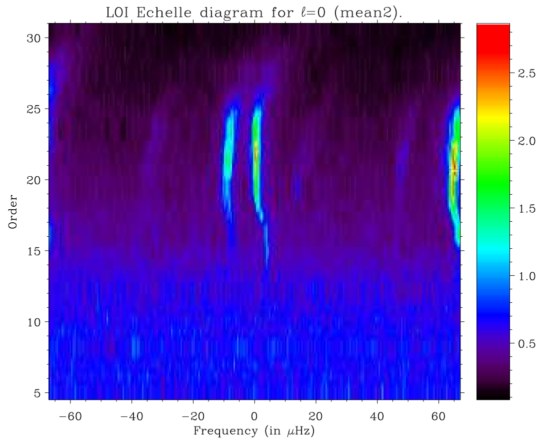


Fig. 1. Amplitude spectra echelle diagram for 1 year of LOI data seeing the Sun as a star. The scale is in part-per-million $\mu\mathrm{Hz}^{-1/2}$ (ppm $\mu\mathrm{Hz}^{-1/2}$). The spacing is tuned for $l=0$ ($\Delta\nu$=134.8 μHz). The $l=0$ modes are at the center, the $l=2$ modes are about 10 μHz on the left hand side, the $l=1$ modes are about +65 μHz on the right hand side, the $l=3$ can be faintly seen at about 12 μHz from the left hand side of the $l=1$. Other modes such as $l=4$ and $l=5$ can also be seen faintly seen at -35 μHz and +15 μHz, respectively. The distortion of the ridges are due to sound speed gradients in the solar core.

diagram is only shown for a single n and for intermediate degrees $l \geq 10$.

The purpose of these diagrams is always to show an estimate of the variance of the spectra. In our case we also want to visualize not only the variance but also the covariance of the Fourier spectra. Here we briefly recall from Part I that the observed Fourier spectra ($\boldsymbol{y}$) can be related to the individual Fourier spectra of the normal modes ($\boldsymbol{x}$) by the leakage matrix $\mathcal{C}^{(l,l')}$ by:

$$\boldsymbol{y} = \mathcal{C}^{(l,l')}\boldsymbol{x} \tag{1}$$

where $\boldsymbol{x}(\nu)$ and $\boldsymbol{y}(\nu)$ are 2 complex vectors made each of $2l+2l'+2$ components: $2l$+1 components for l, $2l'$+1 components for l'. Here only two l values have been used for simplicity, there is no difficulty to extend the expression of the leakage matrix for a number of degrees greater than 2. The covariance matrix $\mathsf{V}^{(l,l')}_{m,m'}$ of $\boldsymbol{z_y}$ ($\boldsymbol{z}^{\mathrm{T}}_{\boldsymbol{y}} = (\mathrm{Re}(\boldsymbol{y}^{\mathrm{T}}), \mathrm{Im}(\boldsymbol{y}^{\mathrm{T}}))$), can be derived from the sub-matrix $\mathcal{V}^{(l,l')}$ whose elements can be expressed as:

$$2\mathcal{V}^{(l,l')}_{m,m'} = E[y_{l,m}(\nu)y^{*}_{l',m'}(\nu)] =$$

$$2\sum_{l"=l,l'}\sum_{m"=-l"}^{m"=l"} \mathcal{C}^{(l',l")}_{m',m"}\mathcal{C}^{(l,l")*}_{m,m"} f^{l''}_{m"}(\nu) + 2\mathcal{B}^{(l,l')}_{m,m'} \tag{2}$$

where E is the expectation, $f^{l''}_{m''}(\nu)$ is the model of the line shape of the power spectrum of the (l'',m'') mode, $\mathcal{B}^{(l,l')}$ is the covariance matrix of the noise, and with the $y_{l,m}(\nu)$ having a mean of zero. The factor 2 comes from the fact that the real part of $\mathcal{V}^{(l,l')}$ represents both the covariance of the real or imaginary parts of the Fourier spectra (See Part I, section 3.3.2); the same property applied to the imaginary part of $\mathcal{V}^{(l,l')}$ which represents the covariance between the real part and the imaginary part of the Fourier spectra. Equation (2) contains all the information that we need for visualizing an estimate of the real and imaginary parts of $\mathcal{V}^{(l,l')}$. Drawing from the usefulness of the diagrams of Grec and Brown, we created four new diagrams for visualizing an estimate of $\mathcal{V}^{(l,l')}$, all having various diagnostics power:

1. (m,ν) echelle diagram: estimate of the diagonal elements of $\mathcal{V}^{(l,l)}$ $(l=l')$
2. cross echelle diagram: estimate of the off-diagonal elements of $\mathcal{V}^{(l,l)}$ $(l=l')$
3. inter echelle diagram: estimate of the off-diagonal elements of $\mathcal{V}^{(l,l')}$ (l/nel')
4. cross spectrum ratio: estimate of the ratio of the elements of $\mathcal{B}^{(l,l')}$

Each diagnostic is described hereafter in more detail.

2.1. (m,ν) echelle diagram

The (m,ν) *echelle diagram* is made of $2l+1$ echelle diagrams of each l,m power spectra or $|y_{l,m}(\nu)|^2$. The $2l+1$ echelle diagrams are stacked on top of each other to show the dependence of the mode frequency upon m. These diagrams give an estimate of the diagonal of the covariance matrix of the observations as:

$$2\widetilde{\mathcal{V}}^{(l,l)}_{m,m}(\nu) = |y_{l,m}(\nu)|^2 \tag{3}$$

where $\widetilde{\mathcal{V}}^{(l,l)}$ symbolizes the estimate of $\mathcal{V}^{(l,l)}$. It is important when one makes these diagrams to tune the spacing for the degree to study. The spacing for a given l can be computed from available p-mode frequencies. Since the spacing varies with the degree, other modes with a significant different spacing can be seen more like diagonal ridges crossing the (m,ν) diagrams; this is a powerful tool to identify other degrees.

Nevertheless, the diagnostics power of the m,ν echelle diagram is rather limited for deriving the leakage matrix: it can be shown using Eq. (2) that the diagonal elements of $\mathcal{V}^{(l,l)}$ can be expressed as:

$$\mathcal{V}^{(l,l)}_{m,m} = \sum_{m"=-l}^{m"=l} |\mathcal{C}^{(l,l)}_{m,m"}|^2 f^{l}_{m"}(\nu) + \mathcal{B}^{(l,l)}_{m,m} \tag{4}$$

As we can see with Eq. (4), the sign information of the elements of $\mathcal{C}^{(l,l)}$ is lost; second, their magnitude being typically less than 0.5, the leakage elements cannot be easily seen in the power spectra. Another kind of diagram that preserves the sign of the leakage elements had to be devised.

2.2. Cross echelle diagram

The *cross echelle diagram* of an l,m mode is made of $2l+1$ echelle diagrams of the cross spectrum of m and m' or

$y_{l,m}(\nu)y^*_{l,m'}(\nu)$. The $2l+1$ real (or imaginary) parts of the cross spectra are stacked on top of each other to show the dependence upon m of the mode frequency. These diagrams give an estimate of the rows (or columns) of the covariance matrix of the observations as:

$$2\widetilde{\mathcal{V}}^{(l,l)}_{m,m'}(\nu) = y_{l,m}(\nu)y^*_{l,m'}(\nu) \quad (5)$$

Of course when $m = m'$ we get the echelle diagrams of the previous section. Only $l+1$ cross echelle diagrams are shown as the matrix $\mathcal{V}^{(l,l')}$ is hermitian by definition.

The imaginary part of the cross spectra has some diagnostic power: it represents the correlation between the real and imaginary parts of the Fourier spectra. When the leakage matrix is real, which is generally the case, there is no correlation between the real and imaginary parts. Nevertheless the imaginary part could be helpful to find errors in the filters applied to the images (See Part I, section 3.3.1).

It can be shown that the elements of $\mathcal{V}^{(l,l)}$ can be expressed as:

$$\mathcal{V}^{(l,l)}_{m,m'} = \mathcal{C}^{(l,l)}_{m,m'} f^l_{m'}(\nu) + \mathcal{C}^{(l,l)*}_{m',m} f^l_m(\nu) + \sum_{m"\neq m',m} \mathcal{C}^{(l,l)}_{m',m"} \mathcal{C}^{(l,l)*}_{m,m"} f^{l''}_{m''}(\nu) + \mathcal{B}^{(l,l)}_{m,m'} \quad (6)$$

As we can see with Eq. (6), these diagrams preserve the sign of the leakage matrix elements. In general, the cross spectra for m, m', representing $\mathcal{V}^{(l,l)}_{m,m'}$ carries information over the sign of the leakage elements $\mathcal{C}^{(l,l)}_{m,m'}$ and $\mathcal{C}^{(l,l)*}_{m',m}$. The other additional terms expressed as product of leakage elements are sometimes more difficult to interpret.

But the power of these diagrams is not only restricted to checking the sign of the elements of the leakage matrix. They are also real tools to get a first order estimate of the leakage matrix. We have shown in Part I, that there is no difference between fitting data for which the leakage matrix *is* the identity, and data for which the leakage matrix is *not*. We showed in part I, that the covariance matrix of $\tilde{\boldsymbol{x}}$ ($\tilde{\boldsymbol{x}} = \mathcal{C}^{-1}\boldsymbol{y}$) can be written, similarly to that of Eq. (2)), by the sum of 2 matrices: the first one represents the mode covariance matrix and is diagonal, and the second term represents the covariance matrix of the noise. Therefore, applying the inverse of the leakage matrix to the data should, in principle, remove all artificial mode correlations between the Fourier spectra of $\tilde{\boldsymbol{x}}$: this can be verified using the cross echelle diagrams. This is the most powerful test for deriving the leakage matrices.

The cross echelle diagrams are useful to verify the correlation within a given degree, but other degrees are known to leak into the target degree, such as $l = 6$, 7 and 8 into $l = 1$, 4 and 8, respectively. The purpose of the next diagram is to assess the magnitude of these leakages.

2.3. Inter echelle diagram

The *inter echelle diagram* of an l, m mode for the degree l' is made of $2l'+1$ echelle diagrams of the cross spectrum of l, m and l', m' or $y_{l,m}(\nu)y^*_{l',m'}(\nu)$. The $2l'+1$ real part of the cross spectra are stacked on top of each other to show the dependence upon m'. These diagrams give an estimate of the rows (or columns) of the covariance matrix of the observations as:

$$2\widetilde{\mathcal{V}}_{m,m'}(\nu) = y_{l,m}(\nu)y^*_{l',m'}(\nu) \quad (7)$$

Similarly as for the cross echelle diagram, it will help to visualize the covariance matrix between different degrees, and to derive leakage elements of the full leakage matrix $\mathcal{C}^{(l,l')}$. One can derive an equation similar to Eq. (6) for different degrees showing that the inter echelle diagram carries information over the sign of the leakage elements $\mathcal{C}^{(l,l')}_{m,m'}$ and $\mathcal{C}^{(l',l)*}_{m',m}$.

As mentioned above applying the inverse of the leakage matrix will help to verify to the first order that there is no artificial correlation due to the p modes. When different degrees are involved the full leakage matrix $\mathcal{C}^{(l,l')}$ has to be used, producing diagrams that should have no artificial correlation due to the p modes.

2.4. Cross spectrum ratio

All the previous diagrams are helpful to understand and visualize the mode covariance matrices. Unfortunately, due to the high signal-to-noise ratio, these diagrams cannot be used to evaluate the correlation of the noise in the Fourier spectra. In between the p modes, these correlations can be more easily visualized as we have:

$$\mathcal{V}^{(l,l)}_{m,m'} \approx \mathcal{B}^{(l,l)}_{m,m'} \quad (8)$$

But instead of visualizing $\mathcal{B}^{(l,l')}$, we prefer to look directly at the correlation by computing the ratio of the cross spectra as:

$$\widetilde{\frac{\mathcal{B}^{(l,l')}_{m,m'}}{\mathcal{B}^{(l,l)}_{m,m}}} = \frac{(y_{l,m}(\nu)y^*_{l,m'}(\nu)}{|y_{l,m}(\nu)|^2} \quad (9)$$

This ratio is called the *cross spectrum ratio*. The cross spectrum ratio gives an estimate of the ratio matrix $\mathcal{R}^{(l,l')}$ (See Part I) which gives a better understanding of how much the noise background is correlated between the Fourier spectra. Nevertheless, in the p-mode frequency range, the cross spectrum ratio is more difficult to interpret as the noise correlation is affected by the presence of the modes. By looking away from the modes (at high or low frequency) or by looking between the modes, one could obtain a reasonable good estimate of the noise correlation.

3. Application to data

3.1. (m,ν) echelle diagram for the LOI/SOHO data

Example of these diagrams can be seen in Figs. 2, 3 and 4 for 1 year of LOI data for $l = 1, 2$ and 5, respectively.

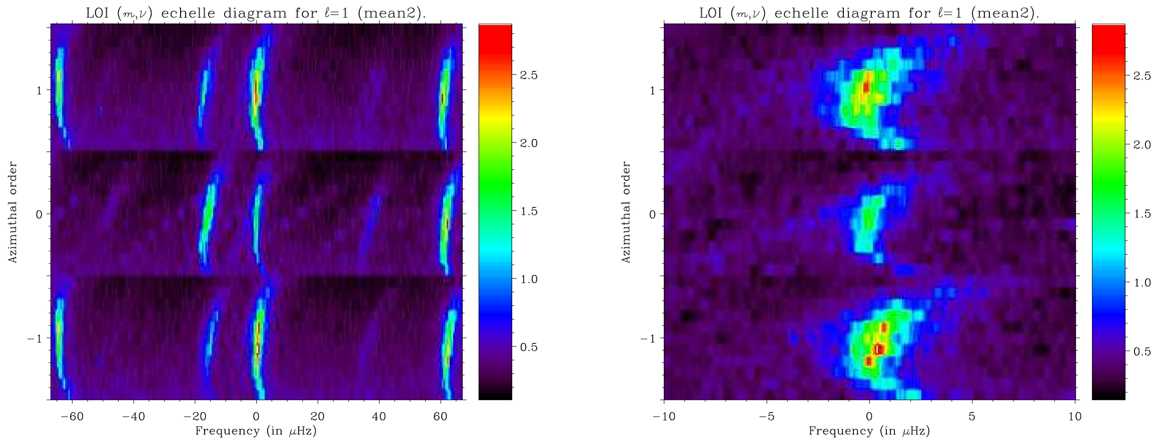


Fig. 2. (m,ν) echelle diagram for 1 year of LOI data for $l=1$. The scale is in ppm μHz$^{-1/2}$. The spacing is tuned for $l=1$ ($\Delta\nu$=134.8 μHz). (Left) The full diagram centered on $l=1$. The $l=3$ modes are located about 15 μHz at the left hand side of the $l=1$. The $l=2$ modes are on the right edge, while the $l=0$ are on the opposite side. The $l=4$ modes can be seen around +40 μHz. (Right) The same diagram but enlarged around $l=1$. The frequency shift or splitting of the modes due to the solar rotation can be seen: the 2 patches of power for $m=-1$ and $m=+1$ are slightly displaced from each other. The artificial correlation between $m=-1$ and $m=+1$ is not as clear.

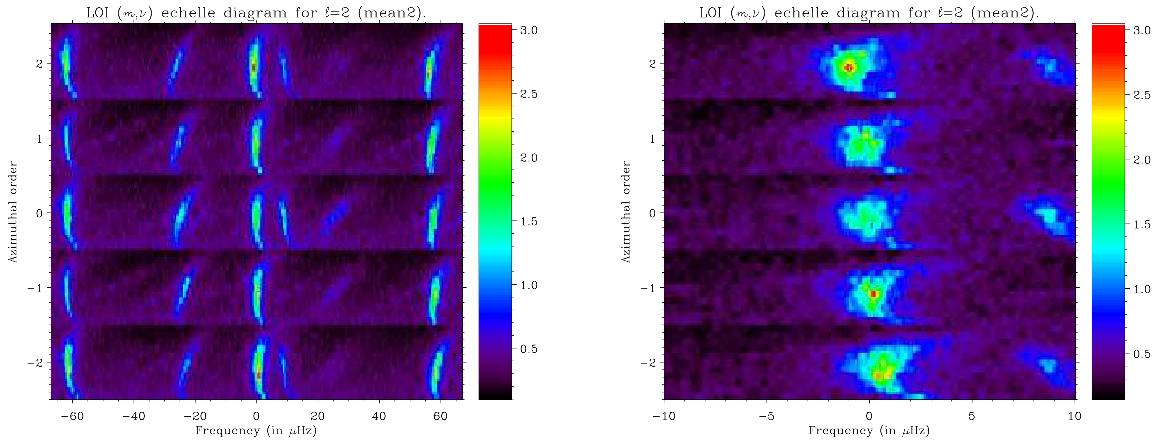


Fig. 3. (m,ν) echelle diagram for 1 year of LOI data for $l=2$. The scale is in ppm μHz$^{-1/2}$. The spacing is tuned for $l=2$ ($\Delta\nu$=135.1 μHz). (Left) The full diagram centered on $l=2$ with a spacing of 135. μHz. The $l=0$ modes are located about 10 μHz at the right hand side of the $l=2$ modes. The $l=3$ modes are on the right edge, while the $l=1$ modes are on the opposite side. The $l=4$ modes are easily seen at about -25 μHz; the $l=5$ starts to appear at +25 μHz. (Right) The same diagram but enlarged around $l=2$ with the $l=0$ modes at the right hand side. The splitting of the modes is clear. Note the absence of the $l=0$ modes for $m=\pm1$

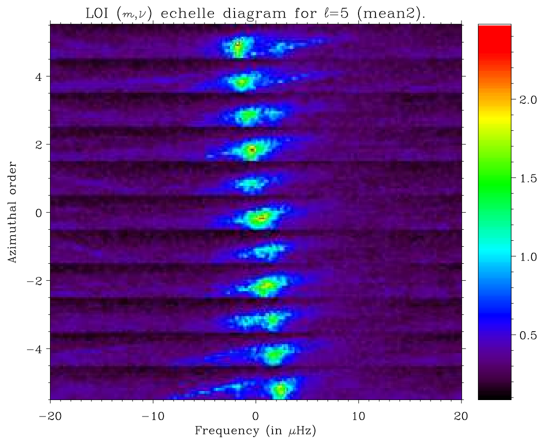


Fig. 4. (m,ν) echelle diagram for 1 year of LOI data for $l=5$. The scale is in ppm μHz$^{-1/2}$. The spacing is tuned for $l=5$ ($\Delta\nu$=136.2 μHz). Here we can clearly see the splitting of the $l=5$ modes. Unfortunately, the $l=5$ data of the LOI are heavily polluted by the presence of the $l=8$ modes. But since the spacing for the $l=8$ modes is different compared with that of the $l=5$ modes, the $l=8$ modes are seen as ridges going from lower left to upper right in each m panel; while the $l=5$ modes are seen more like straight ridges in each m panel. It can also be seen that the $|m|=5$ Fourier spectra contain information both about $m=\pm5$. This is the result of the large LOI pixel size creating oversampling. The $l=8$ modes are also distinguished because the 'splitting' of this alias is in the opposite direction compared with that of the $l=5$ modes.

It is important when one makes such diagrams to tune the spacing for the degree to study. For example, one can see in Fig. 2 that the ridges of power of $l = 0, 1, 2$ and 3 have different shapes than in Fig. 3. Other modes with a significant different spacing can be seen more like diagonal ridges crossing the (m, ν) diagrams; this is a powerful tool to identify other degrees. In Fig. 4, the $l = 5$ (m, ν) echelle diagram is clearly contaminated by an other degree, i.e. $l = 8$, which appears at different frequencies depending on m. For $l = 5, m = -5$, the $l = 8, m = +8$ is quite strong; while for $l = 5, m = +5$, this is $l = 8, m = -8$ which shows up. This kind of 'anti'-splitting behaviour is typical of any aliasing degrees. It is more prominent in the LOI data because of the undersampling effect due to the large size of the LOI pixels.

3.2. A useful detail

Before using the other diagrams on real data, we need to point out a very important property coming from the way the m signals are built. If the weights $W_{l,m}$ applied to the images, to extract an l, m mode, have the same properties as of the spherical harmonics then we have:

$$W_{l,m} = W^*_{l,-m} \tag{10}$$

which means that both the Fourier spectra of $+m$ and $-m$ can be obtained from using a single filter $W_{l,+m}$. The Fourier spectrum of $+m$ will be in the positive part of the frequency range, while that of $-m$ will be in the negative part. This approach was first used by Rhodes et al. (1979) for measuring solar rotation, and mentioned theoretically by Appourchaux and Andersen (1990) for the case of the LOI. Due to the property of the Fourier transform, we recover, in the negative part of the frequency range, *not* the Fourier spectrum of $-m$ but its *conjugate*. This fact is very important, if one does not take care of the sign of the imaginary part of $-m$, it will lead to very serious problem. Needless to say that fitting data without taking into account this detail will have devastating effects. Obviously this important detail cannot be detected in the power spectra.

3.3. Leakage matrix measurement for a single degree

3.3.1. LOI/SOHO data

According to theoretical computation of the p-mode sensitivities of the LOI pixels (Appourchaux and Andersen, 1990) and using the real shape of the LOI pixels (Appourchaux and Telljohann, 1996), the leakage matrices of $l = 1$ and 2 are given by:

$$\mathcal{C}^{(1,1)} = \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & 0 \\ \alpha & 0 & 1 \end{pmatrix}$$

$$\mathcal{C}^{(2,2)} = \begin{pmatrix} 1 & 0 & \alpha_1 & 0 & \alpha_4 \\ 0 & 1 & 0 & \alpha_2 & 0 \\ \alpha_3 & 0 & 1 & 0 & \alpha_3 \\ 0 & \alpha_2 & 0 & 1 & 0 \\ \alpha_4 & 0 & \alpha_1 & 0 & 1 \end{pmatrix} \tag{11}$$

with: $\alpha = 0.474$, $\alpha_1 = \alpha_3 = -0.308$, $\alpha_2 = 0.576$, $\alpha_4 = -0.216$. These leakage matrices are mean value over one year. The leakages vary throughout the year because the distance between SOHO and the Sun varies. There is no B angle effect as the mean B is zero over 1 year. All the leakage elements are real as the LOI filters have the same symmetry as the spherical harmonics.

Figure 5 and 6 display the cross echelle diagram for $l = 1$ and 2, respectively. From Figs. 5 and 6, we can directly verify using Eq. (6) the sign, and sometimes even the magnitude of the leakage elements. The best cross check that our theoretical computations are correct is to apply the inverse of the leakage matrices to the original data (See Section 2.2). Figures 7 and 8 show how the artificial correlations (or m leaks) can be removed from the LOI data; for the latter we also cleaned the original data from the presence of the $l = 0$ modes. In this latter case, $\boldsymbol{y}$ and $\boldsymbol{x}$ (cf Eq. (1)) are both 6-element vectors where the first element is the observed Fourier spectrum for $l = 0$, and the other 5 are the Fourier spectra for $l = 2$ and the associated m's. Both vectors are related to each other by the leakage matrix $\mathcal{C}^{(0,2)}$ of dimension 6×6. We shall see later on with the GONG data that cleaning, similar to that of the LOI, can be achieved not only for 2 degrees but also for 3 degrees ($l = 1, 6$ and 9).

Unfortunately, for the LOI data the leakage matrices of $l = 4, l' = 7$ and $l = 5, l' = 8$ cannot be inverted. It means that the $2l + 2l' + 2$ Fourier spectra of $l \geq 4$ are linearly dependent. This is the result of the pixel undersampling and has two dramatic consequences: the leakage matrix cannot be verified as for the other degrees, and fitting the data as described in Part I is not valid anymore as the leakage matrix needs to be invertible. This problem is specific to the LOI data. One way around the problem is to restrict the fitting to a subset of the spectra having no linear dependence.

3.3.2. GONG data

The leakage matrices of the GONG data have been computed by R.Howe (1996, private communication). We also computed similar leakage matrices using the equations developed in Part I. The integration was made in the θ, ϕ plane with $\theta_{max} = 65.2°, \phi_{max} = 53.5°$ without apodization. We also took into account the effect of subtracting the velocity of the Sun seen as a star which affects GONG instrument's sensitivity to modes only detected by integrated sunlight instrument (mainly the modes for which $l + m$ is even). The theoretical leakage matrices of $l = 1$

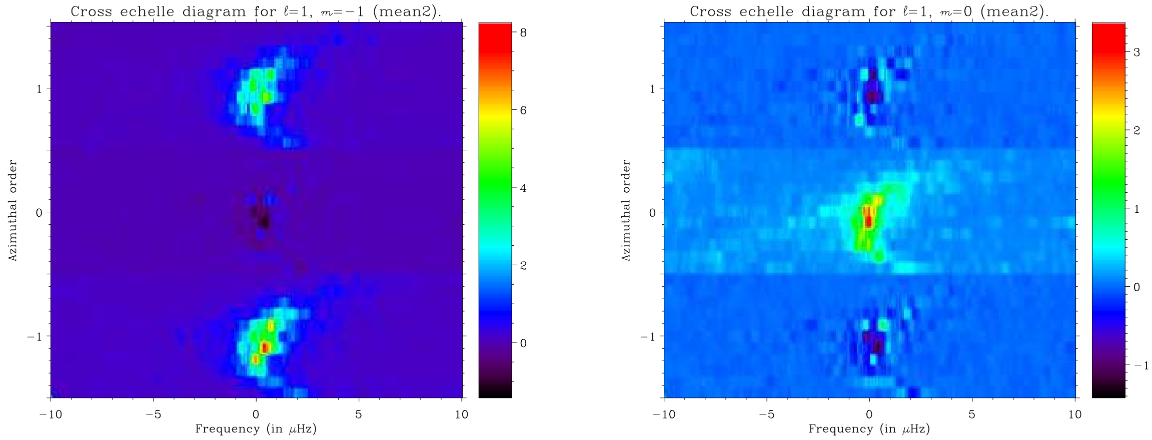
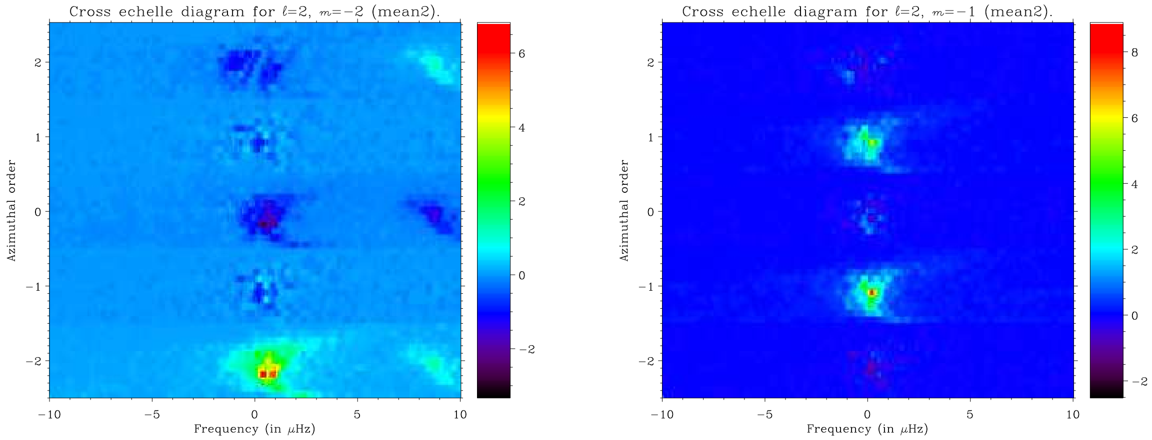


Fig. 5. Real part of the cross echelle diagrams for 1 year of LOI data for $l = 1$. The scale is in ppm$^2\,\mu$Hz^{-1}. The spacing is tuned for $l = 1$ ($\Delta\nu$=134.8 μHz). (Left) For $l = 1, m = -1$. The lower panel represents the echelle diagram of $m = -1$. The 2 other panels are the cross spectra with $m = 0$ and $m = 1$. As predicted, there is no visible correlation between $m = -1$ and $m = 0$ while there is between $m = -1$ and $m = +1$. Here we visualize the first line of the real part of the matrix $\mathcal{V}^{(1,1)}$. (Right) For $l = 1, m = 0$. The bottom panel is the same as the middle panel of the left diagram but with a different color scale. The middle panel is the power spectra of $m = 0$ already shown in Fig. 2; the $l = 6$ modes are visible as a diagonal ridge crossing the ridge of the $l = 1$ modes.



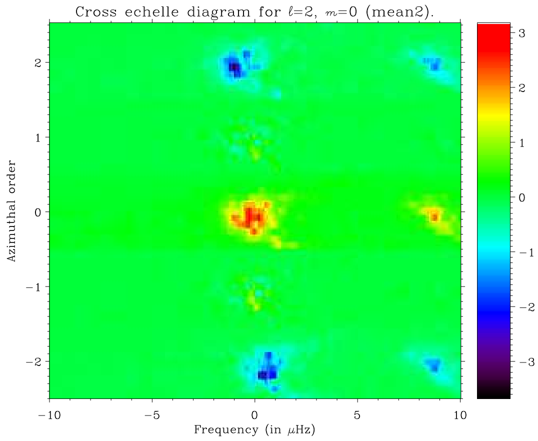


Fig. 6. Real part of the cross echelle diagrams for 1 year of LOI data for $l = 2$. The scale is in ppm$^2\,\mu$Hz^{-1}. The spacing is tuned for $l = 2$ ($\Delta\nu$=135.1 μHz). (Top, left) For $m = -2$. The middle panel (for $m' = 0$) has negative 'power' close to the frequencies of the $m = -2$, indicating that α_3 is negative. There is some indication that α_1 may be negative. The upper panel (for $m' = 2$) has 2 patches of negative 'power' close to the frequencies of $m = -2$ and $m = +2$, showing that α_4 is also negative. These 2 patches are clearly separated because of the mode splitting. (Top, right) For $m = -1$. Here it is obvious that α_2 is positive. (Opposite) For $m = 0$. There is now 2 patches of negative power in 2 opposite panels ($m' = -2$ and $m' = 2$) due to the negativity of α_3. It is likely that α_1 is also negative.

and 2 for GONG are also given by Eq. (11) but with: $\alpha = -0.55$, $\alpha_1 = -0.268$, $\alpha_2 = 0.451$, $\alpha_3 = -0.122$, and $\alpha_4 = -0.290$. For $l = 1$, the leakage between $m = -1$ and $m = +1$ is negative due to the subtraction of the mean velocity which affects these modes. For $l = 2$, the leakage between $m = -1$ and $m = +1$ has about the same value as for the LOI; these modes are not affected by the subtraction.

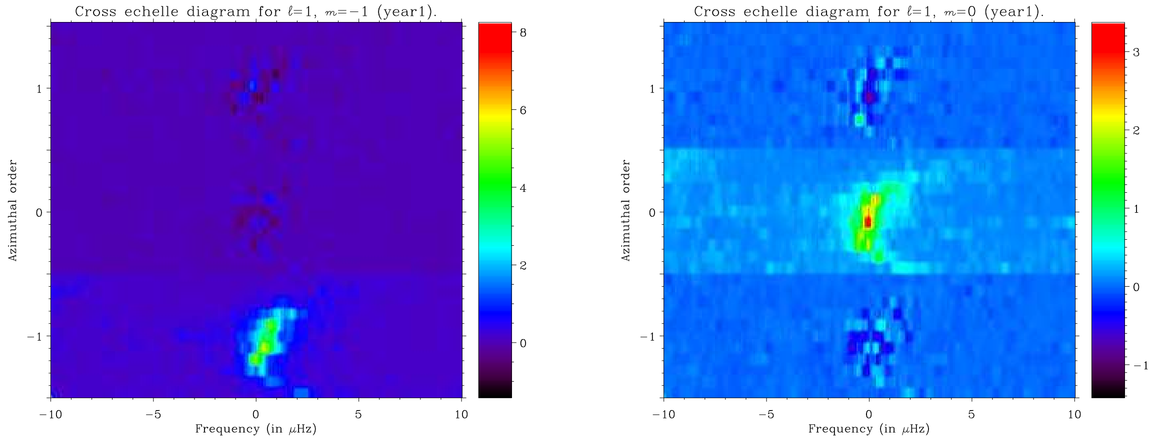
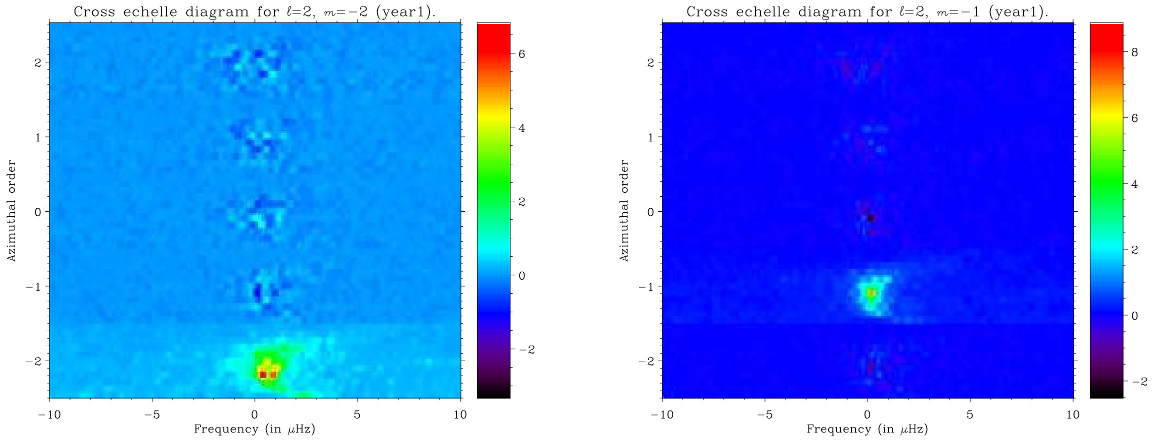


Fig. 7. Real part of the cross echelle diagrams of 1 year of LOI data for $l = 1$. The scale is in ppm$^2\mu$Hz^{-1}. The spacing is tuned for $l = 1$ ($\Delta\nu$=134.8 μHz). The inverse of the theoretical leakage matrix has been applied to the original data with α=0.474. (Left) For $l = 1, m = -1$. The artificial correlation between $m = -1$ and $m = +1$ has been entirely removed (See Fig. 5 for comparison). (Right) For $l = 1, m = 0$. There is no improvement as there was no correlation before.



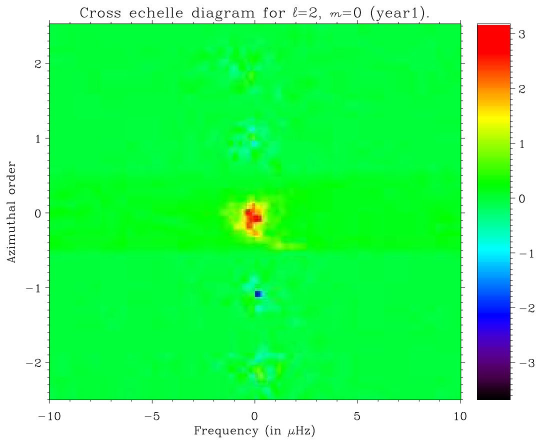


Fig. 8. Real part of the cross echelle diagrams of 1 year of LOI data for $l = 2$ after having applied the inverse of the leakage matrix. The scale is in ppm^2 μHz^{-1}. The spacing is tuned for $l = 2$ ($\Delta\nu$=135.1 μHz). (Top, left) For $m = -2$. (Top, right) For $m = -1$. (Opposite) For $m = 0$. Using the inverse of the leakage matrix we have removed all artificial correlations (See Fig. 6 for comparison). The spectra have been cleaned both from the m leaks and from $l = 0$ modes because we used the leakage matrix of $l = 2$ and $l = 0$ only; the leakages from all the other degrees were ignored.

The cross echelle diagrams of the $l = 2$ GONG data are very close to those of the LOI (See Fig. 6). Figure 9 shows an example of a GONG cross echelle diagram after having applied the inverse of the leakage matrix. It is clear that the p-mode correlations are removed. For $l = 1$ we have used the cross echelle diagram of $m = -1$ and $m = 1$ for inferring quantitatively the off-diagonal leakage element α. We first applied the inverse of a leakage matrix to the $l = 1$ data, and then constructed the cross echelle diagram of $m = -1$ and $m' = 1$ for these data. We then collapsed this diagram by adding up all the modes with $n = 10 - 26$, and finally we corrected the collapsed diagram from the solar noise background. The collapsed diagram is shown in Fig. 10 (Left) for no correction ($\alpha = 0$) and for an α

of -0.53; the corrected surface of the collapsed diagram as a function of α is shown on Fig. 10 (Right). When the corrected surface is close to 0, there is no artificial correlation remaining.

Cleaning the data from artificial correlations has also been done in a different way by Toutain et al. (1998). Using Singular Value Decomposition, they recomputed pixel filters for the MDI/LOI proxy so as to remove the m leaks and the other aliasing degrees. Here we have shown, that data cleaning is also possible *without* having the pixel time series, but using the Fourier spectra. This latter cleaning technique is more useful as one has to reduce a smaller amount of data.

3.4. Leakage matrix measurement for many degrees

3.4.1. LOI/SOHO data

Figure 11 shows the inter echelle diagram for $l = 1$, $l' = 6$. This diagram is more difficult to interpret, but using specific spacing for l or l', one can unambiguously identify which degree contributes to the correlation. For the LOI data, the $l = 5$ and $l = 8$ (similarly $l = 4$ and $l = 7$) are strongly contaminated by each other. This is again due to the spatial undersampling which prevents the LOI data to be cleaned from aliasing degrees making the leakage matrix not invertible. Fortunately, the $l = 1$ LOI data are far less affected by the presence of $l = 6$ compared with that of GONG (See in the next section); this is the result of an effective apodization function (limb darkening) which creates a narrower spatial response than that of GONG (line-of-sight projection).

Figure 8 has already shown that we could clean the $l = 2$ LOI data from the $l = 0$ mode. This was done by taking into account the leakage matrix $\mathcal{C}^{(0,2)}$. Unfortunately as mentioned above for the $l \geq 4$, the LOI data cannot be cleaned from aliasing degrees because the leakage matrice $\mathcal{C}^{(1,6)}$, $\mathcal{C}^{(4,7)}$ and $\mathcal{C}^{(5,8)}$ are singular.

3.4.2. GONG data

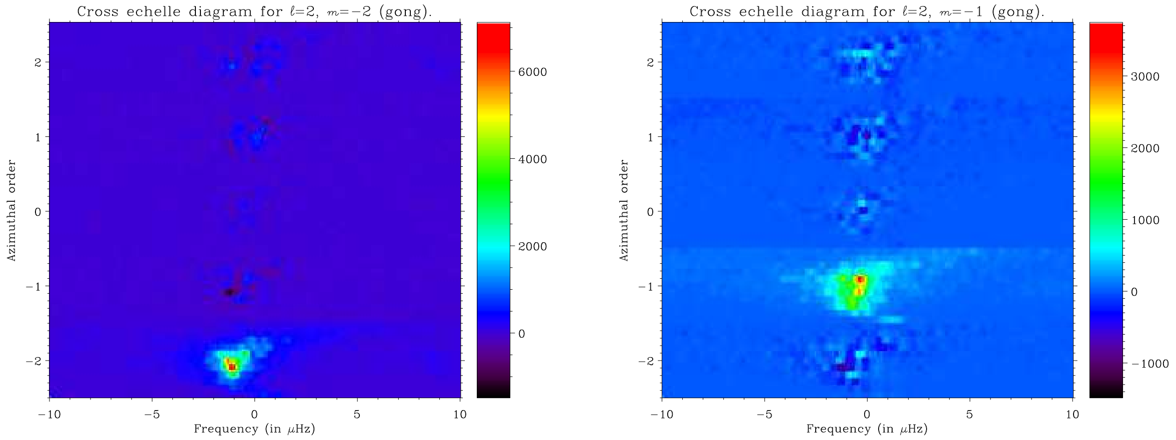
Figure 12 shows the inter echelle diagram for $l = 1, l' = 6$. The correlations are clearly visible. If they are not taken into account they are likely to bias the frequency and splittings of the $l = 1$ (See Rabello-Soares and Appourchaux, 1998, in preparation). By using the inverse of the leakage matrix of $l = 1, 6$ and 9 (i.e. $\mathcal{C}^{(1,6,9)}$), one can clean the data from these spurious correlations arising both from the two aliasing degrees ($l = 6$ and 9) and from the m leaks. Figure 13 shows the 'cleaned' inter echelle diagram which should be compared with Figs. 12. The correlations have been almost entirely removed. Some correlations between the $l = 1, m = -1$ modes and the $l = 6$ modes are still present. This is due to the fact that the computation of the leakage for $l = 1, m = \pm 1$ are very sensitive to the subtraction of the velocity of the Sun seen as a star. This is not the case for the $l = 1, m = 0$ modes as they are insensitive to this effect. Nevertheless, our imperfect knowledge of the leakage matrix can be adjusted in order to remove the residual correlations. It helped to reduce the systematic errors of the $l = 1$ splitting at high frequencies, mainly where the $l = 6$ and 9 modes start to cross over the $l = 1$ modes (Rabello-Soares and Appourchaux 1998, in preparation).

It is possible to apply this cleaning technique to higher degree modes ($l \approx 50 - 100$. In principle, this is feasible, although the plethora of data involved may be fairly substantial. For higher degrees, the ridges of the modes in the (m, ν) echelle diagrams are almost parallel to each other. The idea would be not to use the full Fourier spectra, as we did for the low degree GONG data, but to use only a small frequency range of about $\pm$ 30 μHz around the target degree. By doing so, we will not only clean the data from the aliasing degrees but also from the m leaks. In this case, the mode covariance matrix is diagonal. If we assume, wrongly, that the noise covariance matrix is diagonal, the statistics of each cleaned spectrum, for high degree modes, could be approximated by a χ^2 with 2 degree of freedom. Neglecting the off-diagonal elements of the noise covariance matrix will lead to larger mode linewidth, thereby producing underestimated splittings; this systematic bias will decrease as the degree. This approximation will produce much less systematic errors than the approximation used by Hill et al. (1996) for the original data. The use of the suggested approximation for the GONG data may help, at the same time, to reduce systematic errors, and to increase computing speed for higher degree modes. The degree where one needs to switch between this approximation and the correct analysis needs to be determined.

3.5. Noise covariance matrix measurement

3.5.1. LOI/SOHO data

The leakage and ratio matrices have similar properties (See Part I). For example, for a given l there is no correlation between the m for which $m + m'$ is odd for either matrix. As outlined in Part I, for a given degree, the ratio matrix is also close to the leakage matrix. For example, Fig. 14 shows that for the LOI data the measured ratio matrix element between $m = -1$ and $m = +1$ is about $+0.55$ which is rather close to the $\alpha = 0.474$ given in Eq. 11. The ratio matrix is also useful if one wants to reduce the number of noise parameters to be fitted. This is useful when the noise background, mainly of solar origin, varies slowly over the p-mode range. In this case the ratio can



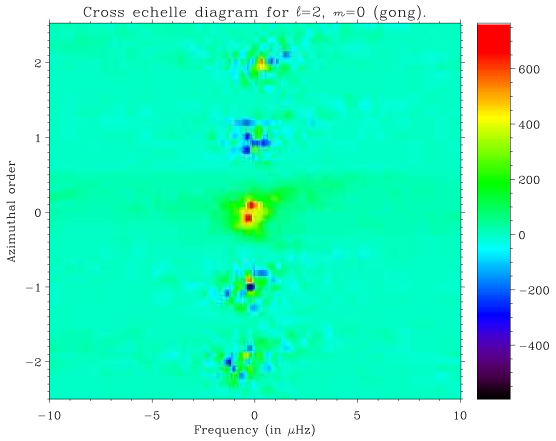


Fig. 9. Real part of the cross echelle diagrams for 360 days of GONG data for $l = 2$ after having applied the inverse of the leakage matrix. The scale is in $\mathrm{m^2\,s^{-2}\,Hz^{-1}}$. The spacing is tuned for $l = 2$ ($\Delta\nu$=135.1 μHz). (Top, left) For $m = -2$. (Top, right) For $m = -1$. (Opposite) For $m = 0$. As for the LOI data the artificial correlations are removed. The remaining correlations are artifact due to the statistical fluctuations that are enhanced when the signal-to-noise ratio is low as for $m = 0$.

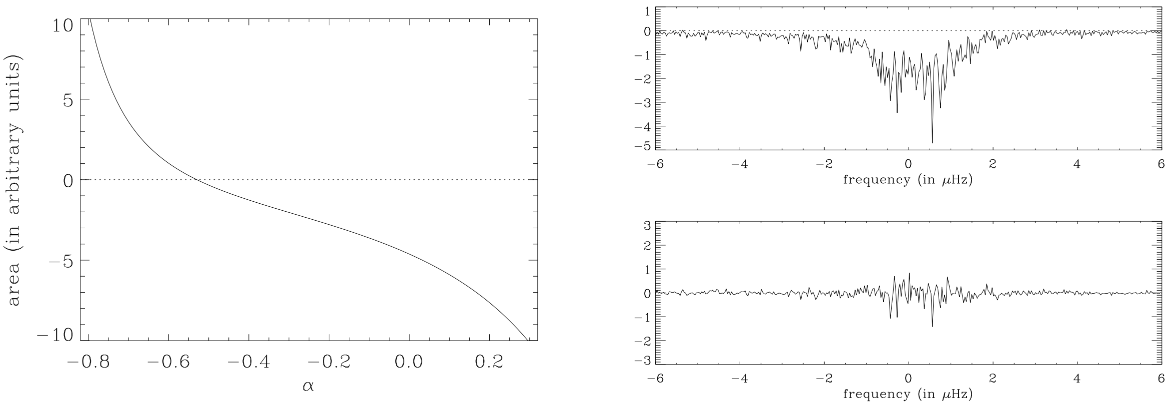


Fig. 10. (Left) Collapsed diagram of the real part of the cross spectrum of $m = -1$ and $m = +1$ for the GONG data after having applied the inverse of the leakage matrix, (top) no correction, (bottom) α=-0.53. (Right) Surface of the collapsed diagram as a function of α, the surface is 0 for α=-0.53.

be assumed to be constant over the p-mode range. For example for $l = 1$ we can fit 2 noise parameters instead of 3. When the noise varies in the p-mode range, it is more advisable to fit as many noise parameters as required.

Figure 15 shows the cross spectrum ratios of $l = 2$ for the LOI data. As this is commonly the case for the LOI, the ratios are independent of frequency. In addition, as outlined in Part I, the ratio matrix is very close to the leakage matrix. Using the values of α given for the LOI in Eq. (11) and Fig. 15, one can see that this is the case for the LOI. The ratio matrix has even the same symmetry property as the leakage matrix (not shown here). In the case of the LOI, we sometimes use the independence of the ratio with frequency to reduce the number of free parameters. It is less straightforward to measure the noise correlation for the GONG data than for the LOI data. We recommend to measure it in between the p modes because that is what the fitting routines will determine.

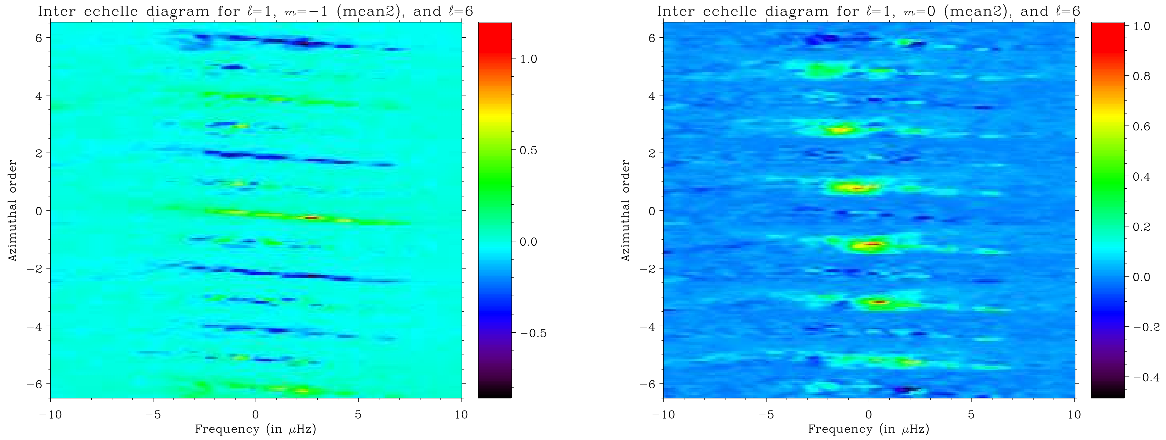


Fig. 11. Real part of the inter echelle diagrams for 1 year of LOI data for $l = 1$ and $l' = 6$. The scale is in ppm$^2\,\mu$Hz^{-1}. The spacing is tuned for $l' = 6$ ($\Delta\nu$=136.7 μHz). (Left) For $l = 1, m = -1$. Given the structure of the ridges, it means that the $l = 1$ modes leak into the even m modes of $l' = 6$, while the $l' = 6$ modes do not leak into the $l = 1$. (Right) For $l = 1, m = 0$. Again using the structure of the ridges, the odd m' of $l' = 6$ leaks weakly into $l = 1, m = 0$.

3.5.2. GONG data

Figures 14 and 15 show the cross spectrum ratios of $l = 2$ for the GONG data. The ratio matrices (as the leakage matrices) are not symmetrical. Figure 14 shows also the effect of the subtraction of the full disk integrated velocity in the GONG data. In this case, the correlation is extremely high (about -0.8) and very different from the $\alpha = -0.55$ computed for GONG. Apart from the $l = 1$, these matrices tend to be very close to those of the leakage matrices (see previous section). It is also clear that the ratios depend upon frequency, probably due to the effect of mesogranulation affecting the spatial properties of the noise with frequency. In this case, the dependence of the noise correlation upon frequency has to be taken into account in the fit either by adding free parameters or by measuring these correlations in the cross spectrum ratios (see Rabello-Soares and Appourchaux 1998, in preparation).

4. Summary and Conclusion

For helping the understanding of fitting p-mode Fourier spectra, we have devised 4 new helioseismic diagnostics:

1. the (m, ν) echelle diagram helps visualizing:
   - which degree can be detected,
   - the mode splitting and
   - spurious degrees.
2. the cross echelle diagram helps verifying:
   - the sign of the imaginary parts of the Fourier spectra for $m < 0$,
   - the sign of the elements of the leakage matrix of the degree l,
   - to the first order the theoretical leakage matrix of the degree l by applying the inverse of this matrix to the data (i.e. cleaning the data from the m leaks),
   - the measurement of the real leakage matrix and
   - the cleaning of the data from m leaks.
3. the inter echelle diagram helps verifying:
   - the sign of the elements of the full leakage matrix for degrees l, l',
   - to the first order the theoretical leakage matrix of the degree l, l' by applying the inverse this matrix to the data,
   - the measurement of the real leakage matrix and
   - the cleaning of the data from aliasing degrees.
4. the cross spectrum ratio helps deriving :
   - the amplitude of the noise correlation between the Fourier spectra and
   - the frequency dependence of the correlations.

These steps will help to evaluate the matrices of the leakage, mode covariance, ratio and noise covariance directly from observations. The fitting of the p modes will be considerably eased by verifying that the theoretical knowledge of these matrices is correct.

These steps have been used both on data for which we design the spatial filtering (LOI instrument), and on data for which we did not design the filtering (GONG). As a matter of fact these diagnostics are so powerful that a theoretical knowledge of the various matrices is not always necessary to understand the data. Very often these diagnostics can also be used to find bugs in the data analysis routines. Nevertheless, it is advisable to know to the first order the leakage and ratio matrices for speeding up the analysis process.

Last but not least, the knowledge of the leakage matrices can be used for cleaning the data from m leaks and from undesired aliasing degrees. This cleaning can be performed either when the pixel time series are available (MDI data, Toutain et al. , 1998) or more simply when only the Fourier spectra are available (LOI, GONG). The cleaning has very useful application for the GONG velocity data for removing aliasing degrees from $l = 1$, 6 and 9 for instance, and for inferring better $l = 1$ splittings (Rabello-Soares and Appourchaux 1998, in preparation). The cleaning is somewhat easier with Fourier spectra as one does not need to reduce large amount of image time series. It is advisable that, in the near future, this cleaning technique be used for higher degree modes. This will hopefully provide helioseismology with frequency and splitting data having much less systematic errors than before.

Acknowledgements. SOHO is a mission of international collaboration between ESA and NASA. This work utilizes data obtained by the Global Oscillation Network Group (GONG) project managed by the National Solar Observatory, a Division of the National Optical Astronomy Observatories, which is operated by AURA, Inc. under a cooperative agreement with the National Science Foundation. The GONG data were acquired by instruments operated by the Big Bear Solar Observatory, High Altitude Observatory, Learmonth Solar Observatory, Udaipur Solar Observatory, Instituto de Astrofísico de Canarias and Cerro Tololo Interamerican Observatory. Many thanks to Takashi Sekii for constructive comments on the manuscript, and for extensive cyberspace chatting. I am grateful Mihir Desai for proofreading the English. Last but not least, many thanks to my wife for her patience during the painful writing of these 2 articles.

References

Appourchaux, T. and Andersen, B.N., 1990, Sol.Phys. 128, 91

Appourchaux, T. and Telljohann, 1996, LOI GSE and VDC specification Doc., Version 1.7, available on the VIRGO home page (http://www.so.esa.nl/~thierrya/VIRGO/virgo.html)

Appourchaux, T., Andersen, B.N., Fröhlich, Jiménez, A., Telljohann, U. and Wehrli, C., 1997, Sol. Phys., 170, 27

Brown, T., 1985, Nature, 317, 591

Grec, G., PhD thesis, Université de Nice, Nice

Hill, F., Stark, R.T., Stebbins, R.T. et al. , 1996, Science, 272, 1292

Rhodes, E.J., Deubner, F-L. and Ulrich, R.K., 1979, ApJ, 227, 629

Schou, J., 1992, PhD thesis, *On the analysis of helioseismic data*, Århus Universitet, Århus

Toutain, T., Kosovichev, A. and Appourchaux, T., 1998, IAU General assembly, Kyoto

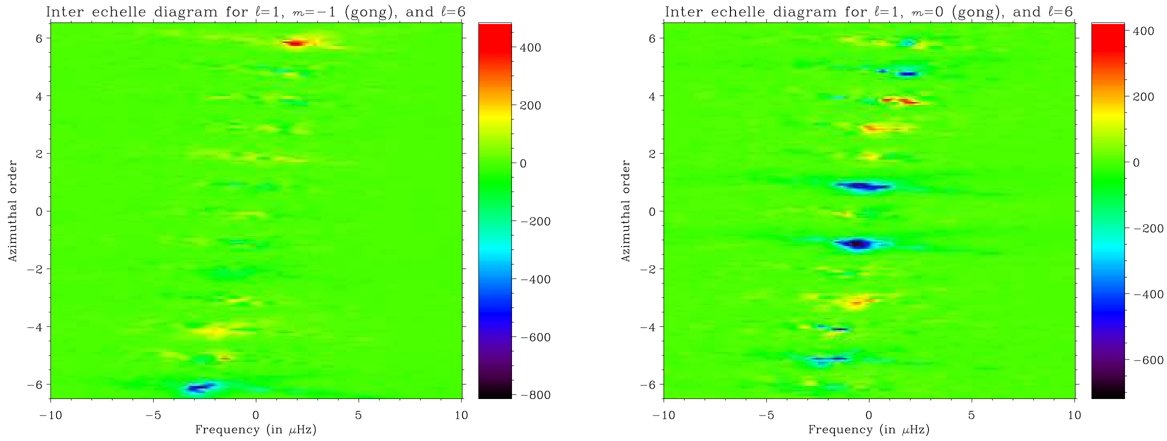


Fig. 12. Real part of the inter echelle diagrams of 1 year of GONG data for $l = 1$ and $l' = 6$. The scale is in $\mathrm{m}^2\,\mathrm{s}^{-2}\,\mathrm{Hz}^{-1}$. The spacing is tuned for $l' = 6$ ($\Delta\nu$=136.7 μHz). (Left) For $l = 1, m = -1$. The strongest correlations are for $m' = \pm 6$ due to $l' = 6$. (Right) For $l = 1, m = 0$. The $l = 6$ modes are strong in all the panels. For $m' = \pm 1$ other modes can be seen as an 'asterisk' ridge: the ridge going from bottom right to upper left represents the $l = 1$ modes; the ridge going from bottom left to upper right represents the $l = 9$ modes; the vertical ridge represents the $l = 6$ modes. At the far left, the ridge parallel to that of the $l = 1$ represents the $l = 3$ modes.

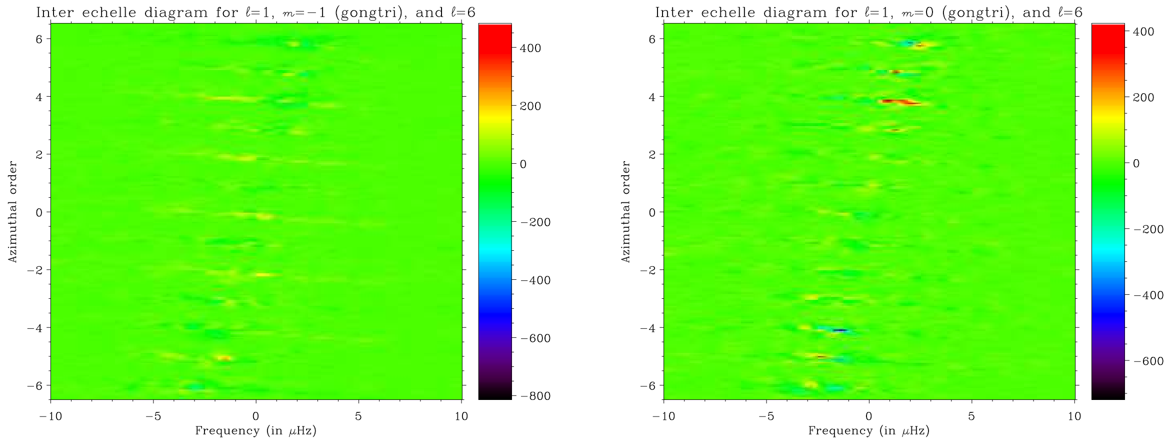


Fig. 13. Real part of the inter echelle diagrams of 1 year of GONG data for $l = 1$ and $l' = 6$ after having applied the inverse of the leakage matrix $\mathcal{C}^{(1,6,9)}$. The scale is in $\mathrm{m}^2\,\mathrm{s}^{-2}\,\mathrm{Hz}^{-1}$. The spacing is tuned for $l' = 6$ ($\Delta\nu$=136.7 μHz). (Left) For $l = 1, m = -1$. The faint yellow oblique ridges are due to the $l = 1$ modes. (Right) For $l = 1, m = 0$. The correlations due to the $l = 1$ and 6 modes are almost entirely removed.

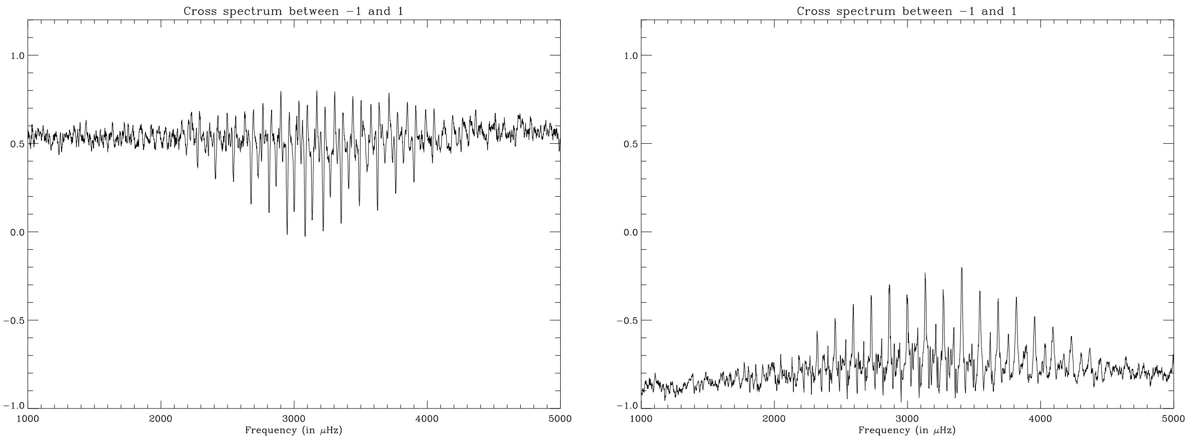


Fig. 14. Real part of the cross spectrum ratios for $l = 1$ smoothed to 10 μHz. (Left) LOI data: for $m = -1$ and m'=1. The noise correlation is about 0.55. (Right) GONG data: for $m = -1$ and $m = 1$. The noise correlation is about -0.77. The large anti-correlation is due to the subtraction of the full-disk velocity which tends to reduce the amplitude of the $m = \pm 1$ modes.

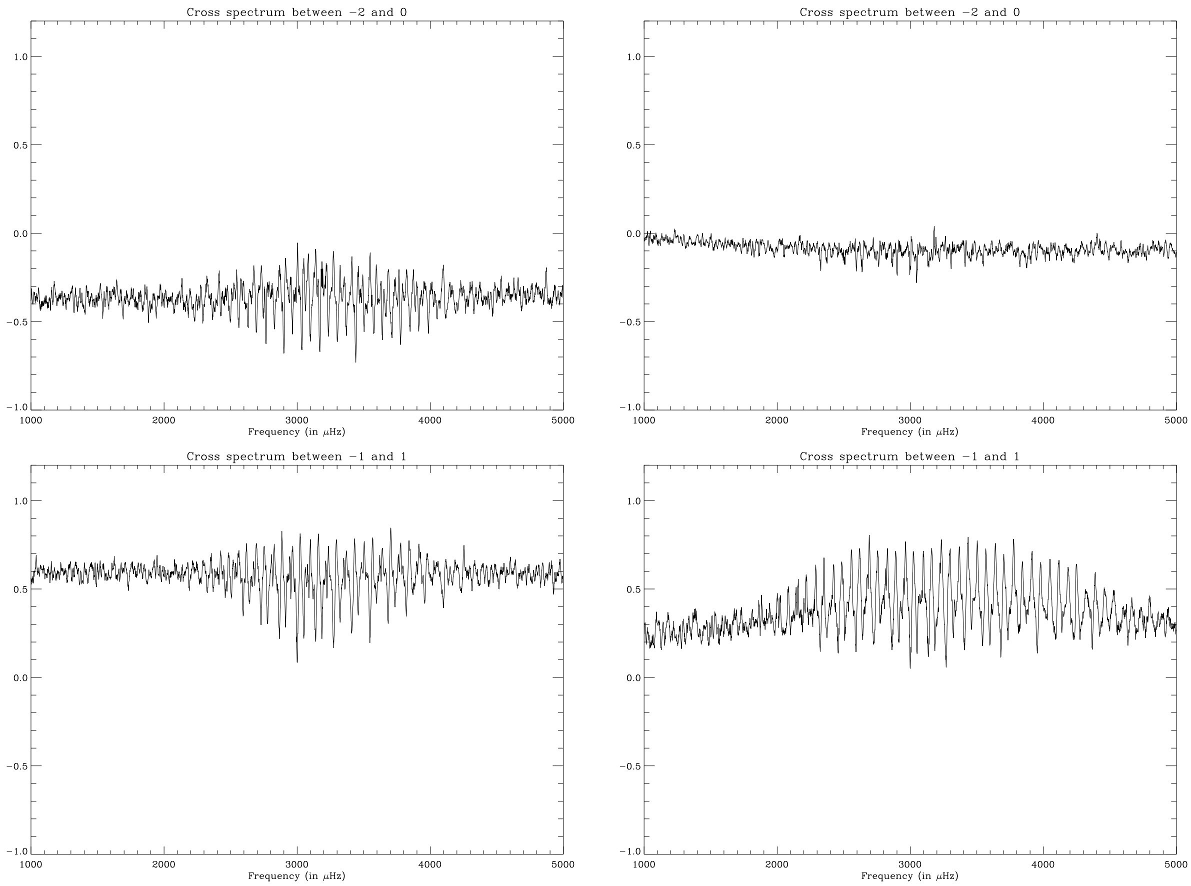


Fig. 15. Real part of the cross spectrum ratios of $l = 2$ smoothed to 10 μHz. (Left) For the LOI data: top, for $m = -2$ and $m' = 0$. The noise correlation is about -0.45; bottom, for $m = -1$ and $m = 1$. The noise correlation is about 0.60. (Right) For the GONG data: top, for $m = -2$ and $m' = 0$. The noise correlation is rather small and about -0.10; bottom, for $m = -1$ and $m' = 1$. The typical value is about 0.4.