Mon. Not. R. Astron. Soc. **000**, 1–10 (2001) Printed 28 March 2001 (MN LATEX style file v1.4)

Effects of a gap filling method on p-mode parameters

D. Fierry Fraillon and T. Appourchaux
Space Science Department of ESA, ESTEC, 2200 AG Noordwijk, The Netherlands

Accepted 2001 February 9. Received in original form 2000 December 26

ABSTRACT

The quality of helioseismological ground based data strongly depends on the presence of gap in the observational window. In order to address that problem in the case of full disk low-degree p-mode velocity measurements, Fossat et al. (1999) proposed a gap filling method called 'Repetitive music'. The autocorrelation function of the velocity signal shows a correlation of more than 70% at about 4 hours due to the quasi-periodicity of p-mode peaks in the Fourier spectrum. The method then consists in filling gaps of the velocity signal by data, when they exist, located 4 hours before or after.

By using Monte Carlo simulations, we assess the effects of the gap filling method on p-mode parameters and their errors. A way to remove the modulation, resulting from the gap-filling method, in the power spectrum is proposed; its effects on p-mode frequencies, linewidths, amplitudes and asymmetries are discussed as a function of frequency and signal-to-noise ratio.

Key words: p-modes – Gap-filling: Sun.

1 INTRODUCTION

The presence of gap in full-disk velocity measurements strongly affects the quality of p-mode parameters estimation. For duty cycle lower than 100%, a part of the mode energy is redistributed in sidelobe peaks located at 11.57 μHz around each mode peaks. The superposition between both modes and sidelobes profiles, and the reduction of the signal-to-noise ratio, lead to a p-mode parameter determination with a lower accuracy than the one-hundred-percent duty cycle case (Fig. 1). For reducing systematic bias in p-mode parameters measurements the presence of sidelobes can be taken into account in a fitting procedure (Gelly et al. 1988), but the corresponding error bars are still strongly increased by the presence of gap in the velocity signal.

In order to increase the duty cycle of the time series, a gap filling method called "Repetitive Music" (Fossat et al. 1999) is used (Fig. 1). Because of the quasi-perodicity of the p-mode peaks in the Fourier power spectrum, the p-mode velocity presents a correlation of 70 % at 4 hours (Gabriel et al. 1999). Then the idea of the Repetitive Music is to fill gap in the velocity signal by data, if they exist, located 4 hours before or/and after.

2 THE GAP FILLING METHOD

A way to assess the gap filling effects on the p-mode parameters is to use Monte Carlo simulations. Each realization is generated from the Fourier spectra having a χ^2 with 2 degree-of-freedom distribution and a prescribed mean value (Anderson et al. 1990, Toutain and Appourchaux 1994, Fierry Fraillon et al. 1998). Simulations of the velocity signal $V(t)$ are then computed using the inverse Fourier transform of the spectra. On each simulation, the gapped temporal set $V_g(t)$ is obtained by applying a window function $W(t)$:

$$V_g(t) = V(t)W(t) \tag{1}$$

The gap filling method is then applied on $V_g(t)$ as follows:

- a gap is not filled when there is no data 4 hours before and after.
- a gap is filled by data located 4 hours before or after when only one exist
- a gap is filled by the average of data located 4 hours before and after when they both exist

The resulting gap filled velocity $V_f(t)$ is given by:

$$\begin{aligned} V_f(t) &= \sum_{t_k} V_g(t_k) + \sum_{t_i} V_g(t_i - \tau) + \sum_{t_j} V_g(t_j + \tau) \\ &\quad + \frac{1}{2} \sum_{t_h} (V_g(t_h - \tau) + V_g(t_h + \tau)) \end{aligned} \tag{2}$$

where τ is the filling period ($\tau \simeq 4h$), $t{=}t_i$ is the case where there is data only four hours before the gap, $t{=}t_j$ when there

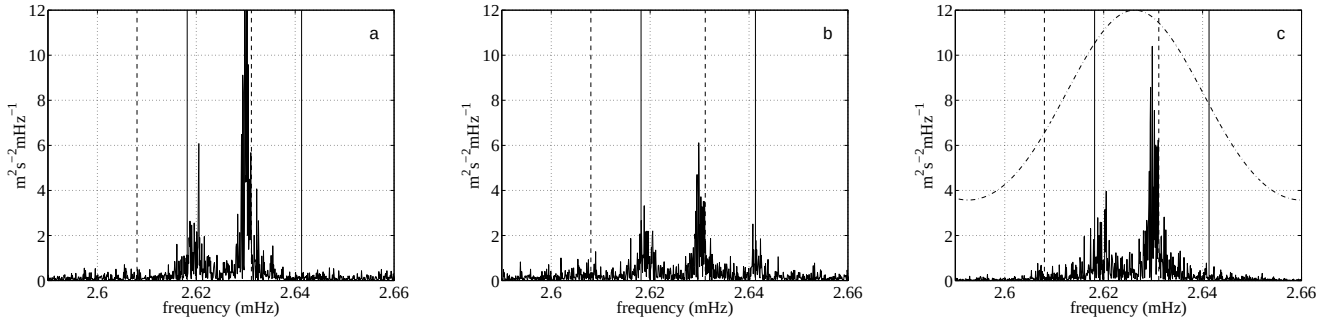

Figure 1. Observational window effects and gap filling results. a) Monte Carlo realization of an l=0-2 group, with a duty cycle of 100 %. b) same realization after applying an observational window on mode velocity, with a duty cycle of 55 %. c): same realization after applying the gap filling method in the case of a one-hundred-percent recovery. The solid lines show the location of the l=0 sidelobes and the dashed lines are for the l=2 sidelobes. In the case of the gapped spectrum (b) a part of the mode energy is redistributed in sidelobes peaks located at 11.57 μHz (because of the 24 hours periodicity in ground based observational window) around each mode peak. The superposition between both modes and sidelobes, and the reduction of the signal-to-noise ratio, lead to a p-mode parameter estimation with a lower accuracy than the one-hundred-percent case. For the gap filled spectrum (c) the sidelobes disappear, and the power spectrum looks like the original one. In fact, the gap filled spectrum is modulated by a $\cos^2(\pi\nu\tau)$-type function which must be taken into account in the p-mode parameters determination. The noise is lower in gap filled spectrum (c) than in the ungapped one (a) because this modulation (dash-dotted curve, in arbitrary units) is indeed smaller than 1 away from the modes it matches.

is data only four hours after the gap and t=t_h when there is data both before and after the gap.

On each simulated power spectrum, $D(\nu)$=$|\widehat{V}(\nu)|^2$ (ungapped) and $D_f(\nu)$=$|\widehat{V}_f(\nu)|^2$ (gap filled), the p-mode parameters are estimated using a maximum likelihood fit with a model mode profile. Using 500 realizations, the probability density function of each parameter is computed for estimating both its mean value and its error bar. In order to valid the gap filling method, we compute and compare the statistical mean of p-mode parameters coming from the two spectra (ungapped and gap filled).

3 MODELIZATION OF THE GAP FILLING EFFECTS

The gap filled spectrum $D_f(\nu)$ differs from the original one $D(\nu)$ by a modulation function (Fossat et al. 1999) that must be computed in order to estimate correctly gap-filled p-mode parameters. Equation (2) can be rewritten as a sum of different time series using $V_g(t) = \sum_k V_k(t)$ where a partial velocity set $V_k(t)$ is composed by a part of the data of $V_g(t)$ and settled by how its data points are used in the gap filling method. In that case the gap filled spectrum depends on cross products between the Fourier transform of the different velocities and therefore on the correlation of data in time serie:

$$\widehat{V}_k(\nu).\widehat{V}^*_{k'}(\nu) = \mathcal{F}\left[\int V_k(u).V_{k'}(t-u)du\right] \qquad (3)$$

where $\mathcal{F}$ denotes the Fourier transform and $\widehat{V}_k(\nu)$ the Fourier transform of the partial velocity set $V_k(t)$, and $\widehat{V}^*_k(\nu)$ its complex conjugate.

The noise correlation depends strongly on the time scales involved. For a white noise, there is no time correlation whereas for a $1/\nu$ noise the correlation can be somewhat larger. The mode correlation depends stronlgly on the mode lifetime: the modulation is then significantly different for modes than it is for noise. To a first approximation, the gap-filling technique is mainly driven by the result of the time correlation term in Eq. (3). The amount of correlation at the main shift time τ determine the complexity of the modulation function that we wish to calculate. This is a simplified view of the calculation that we explain in the following sections.

3.1 Decomposition of the gap velocity in partials sets

A tricky way to compute the modulation function is to write the gap velocity as a sum of partial sets of velocity which do not overlap with each others:

$$\begin{cases} V_g(t) = \sum_k V_k(t) \\ V_k(t_1)V_{k'}(t_1) = 0 \qquad k \neq k' \end{cases}$$

There are four differents basic ways of using a datum for filling a gap (Eq. (2)): $V_g(t+\tau)$, $\frac{1}{2}V_g(t+\tau)$, $V_g(t-\tau)$ and $\frac{1}{2}V_g(t-\tau)$. We can use combinations of the basic moves in order to respect an empty intersection between the differents sets, and also to include the data not used for filling gap, namely $V_0(t)$.

$$\begin{aligned} V_f(t) \;=\; & V_0(t) \\ & + (V_p(t) + V_p(t+\tau)) \\ & + (V_m(t) + V_m(t-\tau)) \\ & + (V_{pm}(t) + V_{pm}(t+\tau) + V_{pm}(t-\tau)) \\ & + \left(V_u(t) + \frac{1}{2}\left[V_u(t+\tau) + V_u(t-\tau)\right]\right) \\ & + \left(V_{um}(t) + V_{um}(t-\tau) + \frac{1}{2}V_{um}(t+\tau)\right) \\ & + \left(V_{up}(t) + V_{up}(t+\tau) + \frac{1}{2}V_{up}(t-\tau)\right) \\ & + \left(V_{umm}(t) + \frac{1}{2}V_{umm}(t+\tau)\right) \end{aligned}$$

Table 1. Classification of how a data point belonging to the gapped velocity V_g belongs to the partial velocity V_k. The symbol • denotes the presence of a data at the current position, and ∘ a gap.

V_k	t_0-2τ	t_0-τ	t_0	t_0+τ	t_0+2τ
V_0	•	•	•	•	•
V_p	∘	∘	•	•	•
V_m	•	•	•	∘	∘
V_{pm}	∘	∘	•	∘	∘
V_u	•	∘	•	∘	•
V_{um}	•	∘	•	∘	∘
V_{up}	∘	∘	•	∘	•
V_{umm}	•	∘	•	•	•
V_{upp}	•	•	•	∘	•

$$+\left(V_{upp}(t)+\frac{1}{2}V_{upp}(t-\tau)\right) \quad (4)$$

Equation (4) is an another expression of Eq. (2) using the decomposition of the gap velocity.

Each datum in the gap velocity is classified by how it is used in the gap filling method and how one kind of data belongs to one set of partial velocity. A data point located at t_0, which belongs to the velocity set $V_k(t)$, is used for filling a gap located at $t_0+\tau$ and/or $t_0-\tau$. The way the data point is used (by average it or not) depends on the presence (or not) of data at the position $t_0 \pm 2\tau$ (Table 1).

Then:

$$\begin{aligned}
\widehat{V}_f(\nu) &= \sum_k f_k(\nu,\tau)\widehat{V}_k(\nu) \\
&= \widehat{V}_0(\nu) \\
&\quad +\widehat{V}_p(\nu)\left(1+e^{2i\pi\nu\tau}\right) \\
&\quad +\widehat{V}_m(\nu)\left(1+e^{-2i\pi\nu\tau}\right) \\
&\quad +\widehat{V}_{pm}(\nu)\left(1+e^{2i\pi\nu\tau}+e^{-2i\pi\nu\tau}\right) \\
&\quad +\widehat{V}_u(\nu)\left(1+\frac{1}{2}\left[e^{2i\pi\nu\tau}+e^{-2i\pi\nu\tau}\right]\right) \\
&\quad +\widehat{V}_{um}(\nu)\left(1+\frac{1}{2}e^{2i\pi\nu\tau}+e^{-2i\pi\nu\tau}\right) \\
&\quad +\widehat{V}_{up}(\nu)\left(1+e^{2i\pi\nu\tau}+\frac{1}{2}e^{-2i\pi\nu\tau}\right) \\
&\quad +\widehat{V}_{umm}(\nu)\left(1+\frac{1}{2}e^{2i\pi\nu\tau}\right) \\
&\quad +\widehat{V}_{upp}(\nu)\left(1+\frac{1}{2}e^{-2i\pi\nu\tau}\right) \quad (5)
\end{aligned}$$

where $\widehat{V}_k(\nu)$ is the Fourier transform of $V_k(t)$. The gap filled power spectrum $D_f(\nu)$ is simply given by:

$$D_f(\nu) = \widehat{V}_f(\nu)\widehat{V}_f^*(\nu) \quad (6)$$

3.2 Case of noise

The solar noise in the power spectra can be typically modeled by a superposition of stochastic processes having a given lifetime or time scale (Harvey, 1985). Such processes are common place in electronics for instance (du Pré, 1950), and lead to the so-called $1/\nu$ noise ($1/\nu^2$ in the power spectrum). The solar p modes are usually at high frequencies, and a proper high pass filter remove the contribution from the low-frequency part of the solar noise. In this case, the noise in the p-mode frequency range can be assumed to be locally white, or at least weakly dependent upon frequency. By filtering out the low frequency part of the noise, we explicitly remove the long-term correlation of the noise, i.e. the filtered noise correlation is close to zero (apart from the auto-correlation value at $t = 0$). Then, considering that there is no overlap between the partial sets of velocity and assuming that the filtered noise has no significant correlation, Eq. (3) gives:

$$\begin{cases} < \left|\widehat{V}_k(\nu)\right|^2 >= b_k \\ \_ \\ < \widehat{V}_k(\nu)\widehat{V}_{k'}(\nu) > \simeq 0 \qquad k \neq k', \end{cases}$$

where <> denotes the statistical mean and b_k the noise value in the power spectrum $D_k(\nu)$. If the noise were to be indeed correlated, there would be no real difficulty in calculating the modulation extracted below. It would simply add unnecessary complications that we can alleviate by a high pass filter.

Then, from Eqs. (5) and (6) the noise in the gap filled spectrum b_f can be written as a sum of noises b_k in the partial spectra multiplied by a modulation function $M_{b,k}(\nu,\tau)$:

$$< b_f >= \sum_k < b_k > M_{b,k}(\nu,\tau) \quad (7)$$

In the case of noise, the effect of the window function in the power spectrum, as a consequence of the theorem of Parseval, is to modify the noise amplitude according to the duty cycle of the temporal set. Then each b_k is proportional to the noise in the ungapped power spectrum b:

$$< b_k >= \alpha_k < b > \quad (8)$$

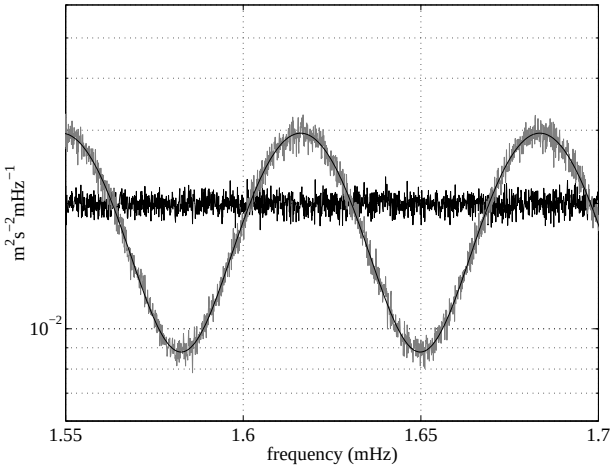

Figure 2. Effects of gap filling on noise. Black: Mean over 500 realizations of an ungapped white noise power spectrum. Gray: Resulting mean power spectrum after application of the gap filling method. Black solid line: analytical model of modulated noise in the case of gap filled spectra computed using the initial value of noise used in the simulation.

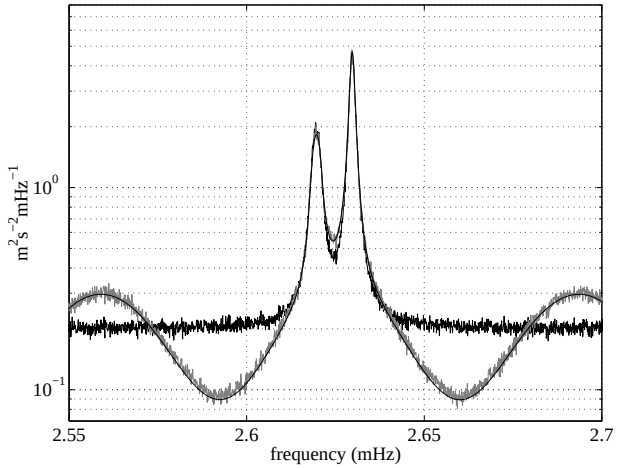

Figure 3. Effects of gap filling on modes, case of a signal-to-noise ratio of 20 for the l=0 mode, average spectrum over 500 realizations. Black: ungapped spectrum for a l=0-2 group, computed using an initial set of parameter $\mathbf{a_0}$. Gray: corresponding gap filled spectrum. Black solid line: analytical model of modulated mode profile, computed using the same $\mathbf{a_0}$. In that case, the analytical model recovers the simulated one.

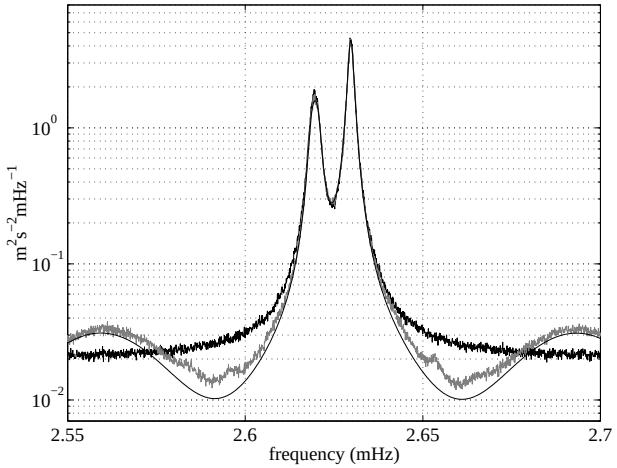

Figure 4. Effects of gap filling on modes, case of a signal-to-noise ratio of 200 for the l=0 mode, average spectrum over 500 realizations. Black: mean ungapped spectrum for a l=0-2 group, computed using an initial set of parameter $\mathbf{a_0}$. Gray: corresponding gap filled spectrum. Black solid line: analytical mode profile for gap filled spectra computed using the same $\mathbf{a_0}$. In that case, the analytical model does not recover the simulated one but the differences are located in the wings of the mode peak.

where α_k is the duty cycle of the partial time serie. Then, from Eqs (5), (6) and (8), the theoretical modulation function for the noise is given by:

The mean noise in the gap filled spectrum $<b_f>$ is then proportional to the mean noise in the original spectrum $<b>$:

$$< b_f >= M_b(\nu,\tau) < b > \tag{9}$$

And,

$$\begin{aligned} M_b(\nu,\tau) &= \alpha_0 + (\alpha_p+\alpha_m)\left[4\cos^2(\pi\nu\tau)\right] \\ &+\alpha_{pm}\left[1+2\cos(2\pi\nu\tau)\right]^2 \\ &+(\alpha_{um}+\alpha_{up})\left[\frac{9}{4}+3\cos(2\pi\nu\tau)+\cos(4\pi\nu\tau)\right] \\ &+(\alpha_{upp}+\alpha_{umm})\left[\frac{1}{4}+2\cos^2(\pi\nu\tau)\right] \\ &+\alpha_u\left[1+\cos(2\pi\nu\tau)\right]^2 \end{aligned} \tag{10}$$

$M_b(\nu,\tau)$ can be obtained without the help of the original spectrum $D(\nu)$ and only with the knowledge of the window function. Figure 2 shows that the noise modulation $M_b(\nu,\tau)$ recovers the effects of the gap filling method on noise.

3.3 Case of modes

The case of the modes is different that of the noise in the sense that their time correlation has to be taken into account. A p mode, with a lifetime varying from 3 days (at $\nu \simeq 4$ mHz) to 24 days (at $\nu \simeq 2$ mHz), presents a correlation of about 80 % at 4 hours and therefore the velocities cross terms can not be assumed equal to 0:

$$\begin{cases} < \left|\widehat{V}_k(\nu)\right|^2 >= L_k(\nu) \\ \\ < \widehat{V}_k(\nu)\widehat{V}^*_{k'}(\nu) >= L_{kk'}(\nu)e^{i\phi_{kk'}(\nu)} \quad k \neq k' \end{cases}$$

where $L_k(\nu)$ is the mode profile in the power spectrum computed from the partial velocity $V_k(t)$. The partial sets of velocity are coming from the same velocity $V(t)$, then the cross-terms $<\widehat{V}_k(\nu)\widehat{V}^*_{k'}(\nu)>$ are complex with a modulus $L_{kk'}$ proportional to $|\widehat{V}(\nu)|^2$, and a phase $\phi_{kk'}(\nu)$.

Because of the mode lifetime, $<\phi_{kk'}(\nu)>$ is not supposed to be equal to zero. In fact, the value of $<\phi_{kk'}(\nu)>$ is strongly dependent on the noise value: the addition of a stochastic noise to the whole velocity time serie lead to a loss of correlation between two points coming from two different partial velodity sets. We can expect from $<\phi_{kk'}(\nu)>$ to be larger when the noise is equal to zero; and to tend toward zero when the signal to noise ratio is decreased. However, considering $<\phi_{kk'}(\nu)> \approx 0$ in the computation of the mode modulation function is a good approximation as it is shown hereafter (See Appendix A).

In the case of modes, the effect of the window function on the mode profile $L(\nu)$ (Lorentz or asymmetric) is to redistribute a part, proportional to the duty cycle, of the mode energy in sidelobes located around $L(\nu)$. Then a good approximation of the mode profile $L_k(\nu)$ in gapped spectra is given by:

$$L_k(\nu) \simeq \alpha_k^2 L(\nu) \tag{11}$$

where α_k is again the duty cycle of the partial time serie. Since the partial velocities $V_k(t)$ and $V_{k'}(t)$ are different sampling of the same velocity $V(t)$, the cross terms $<\widehat{V}_k(\nu)\widehat{V}^*_{k'}(\nu)>$ are proportional to $<|\widehat{V}(\nu)|^2>$ when $<\phi_{kk'}(\nu)> \approx 0$. Then, for the mode profile:

$$\begin{aligned} < \widehat{V}_k(\nu)\widehat{V}^*_{k'}(\nu) > &= \sqrt{L_k(\nu)}\sqrt{L_{k'}(\nu)} \\ &= \alpha_k\alpha_{k'}L(\nu) \end{aligned} \tag{12}$$

Eq. (12) is only valid for the computation of the mode modulation function and for low signal to noise ratio. Then, from Eqs. (5), (6), (11) and (12), we have:

$$
\begin{aligned}
< D_f(\nu) > &= \sum_k |f_k(\nu,\tau)|^2 < |\widehat{V}_k(\nu)|^2 > \\
&\quad + \sum_{k,k'} [f_k(\nu,\tau) f^*_{k'}(\nu,\tau) < \widehat{V}_k(\nu).\widehat{V}^*_{k'}(\nu) > \\
&\qquad + f^*_k(\nu,\tau) f_{k'}(\nu,\tau) < \widehat{V}^*_k(\nu).\widehat{V}_{k'}(\nu) >] \\
&= [M_l(\nu,\tau) + M_{lc}(\nu,\tau)] L(\nu) \qquad (13)
\end{aligned}
$$

where $M_l(\nu,\tau)$ is the modulation function due to the direct terms $|\widehat{V}_k(\nu)|^2$, and $M_{lc}(\nu,\tau)$ the modulation function due to the cross terms $\widehat{V}^*_k(\nu).\widehat{V}_{k'}(\nu)$. $M_l(\nu,\tau)$ is obtained from Eqs (5), (6) and (11):

$$
\begin{aligned}
M_l(\nu) &= \sum_k \alpha_k^2 |f_k(\nu,\tau)|^2 \\
&= \alpha_0^2 + \left(\alpha_p^2 + \alpha_m^2\right)\left[4\cos^2(\pi\nu\tau)\right] \\
&\quad + \alpha_{pm}^2 \left[1 + 2\cos(2\pi\nu\tau)\right]^2 \\
&\quad + \left(\alpha_{um}^2 + \alpha_{up}^2\right)\left[\frac{9}{4} + 3\cos(2\pi\nu\tau) + \cos(4\pi\nu\tau)\right] \\
&\quad + \left(\alpha_{upp}^2 + \alpha_{umm}^2\right)\left[\frac{1}{4} + 2\cos^2(\pi\nu\tau)\right] \\
&\quad + \alpha_u^2 \left[1 + \cos(2\pi\nu\tau)\right]^2 \qquad (14)
\end{aligned}
$$

$M_{lc}(\nu,\tau)$ is obtained from Eqs (5), (6), (12) and (13):

$$M_{lc}(\nu,\tau) = \epsilon_1(\nu,\tau) + \epsilon_2(\nu,\tau) \qquad (15)$$

where:

$$\epsilon_1(\nu,\tau) = 2\sum_{k,k'} \alpha_k \alpha_{k'} \mathcal{R}(f_k(\nu,\tau) f^*_{k'}(\nu,\tau)) \cos(\phi_{k,k'}(\nu)) \qquad (16)$$

And,

$$\epsilon_2(\nu,\tau) = -2\sum_{k,k'} \alpha_k \alpha_{k'} \mathcal{I}(f_k(\nu,\tau) f^*_{k'}(\nu,\tau)) \sin(\phi_{k,k'}(\nu)) \qquad (17)$$

where $\mathcal{R}$ and $\mathcal{I}$ means respectively the real and the imaginary part of a complex number. By considering $<\phi_{k,k'}(\nu)> \simeq 0$, $\epsilon_2(\nu,\tau)$ is neglected (Appendix A2). The function $M_{lc}(\nu,\tau)$ is described in more details in appendix A3. Then:

$$M_{lc}(\nu,\tau) \simeq 2\sum_{k,k'} \alpha_k \alpha_{k'} \mathcal{R}(f_k(\nu,\tau) f^*_{k'}(\nu,\tau)) \qquad (18)$$

The mean mode peak in the gap filled spectrum $<L_f(\nu,\tau)>$ is then proportional to the mode profile in the original spectrum $<L(\nu)>$:

$$
\begin{aligned}
< L_f(\nu,\tau) > &= [M_l(\nu,\tau) + M_{lc}(\nu,\tau)] < L(\nu) > \\
&= M_L(\nu,\tau) L(\nu) \qquad (19)
\end{aligned}
$$

Figures 3 and 4 shows the efficienty of the gap filled spectrum modelisation for two signal-to-noise ratios in the case of a one-hundred-percent duty-cycle recovery. The analytical mode modulation $M_L(\nu,\tau)$ has been approximated for low signal-to-noise ratio. We have neglected some terms in the analytic expression of $M_{lc}(\nu,\tau)$ because theses terms are minimum close to the mode frequency (Appendix A) and tend to 0 when the signal-to-noise ratio is decreasing. Then the gap filled mode model recovers the simulated gap filled spectrum for low signal-to-noise ratio (Fig. 3) while it does not totally for higher signal-to-noise ratio (Fig. 4). In addition, as shown hereafter, neglecting some terms in $M_L(\nu,\tau)$ does not lead to a measurable bias on p-mode parameters.

3.4 p-mode parameters determination

The fact that the modulation is different for modes and noise, Eqs. (9) and (19), implies that there is no way to compute a demodulated power spectrum, from the gap filled one, without the knowledge of p-mode parameters that we wish to determine. Therefore a way to estimate p-modes parameters on gap filled spectra is to fit a modulated mode profile using the modulation function $M_L(\nu,\tau)$ and $M_b(\nu,\tau)$.

The model of the mode profile $L_m(\mathbf{a},\nu,\tau)$ in gap filled spectra, where $\mathbf{a}$ is the set of mode parameters, depends on the gap distribution in the window function and on the value of the filling period τ. The noise modulation $M_b(\nu,\tau)$ and the mode modulation $M_L(\nu,\tau)$ does not depend on the mode parameters. The modulated mode profile is given by:

$$L_m(\mathbf{a},\nu,\tau) = L_0(\mathbf{a},\nu) M_L(\nu,\tau) + b M_b(\nu,\tau) \qquad (20)$$

where $L_0(\mathbf{a},\nu)$ is the usual mode profile (Lorentz or asymmetric profile) and b is a noise constant value.

The gap filled p-mode parameters $\{\mathbf{a_f}\}$ are estimated by a maximum likelihood fit of gap filled spectrum with the mode model $L_m(\mathbf{a_f},\nu,\tau)$. We have checked, using a Kolmogorov-Smirnov test, that the statistical distribution of the gap-filled spectrum was indeed a χ^2 with 2 degrees of freedom. In addition, we have checked that adjacent frequency bins had a level of correlation no greater than 1% after filling the gaps; it was no larger than 8% before doing so (this latter level is comparable to the one given for LOI and GONG data by Appourchaux et al, 2000).

We may then expect some bias in the measurement of $\{\mathbf{a_f}\}$ for high signal to noise ratio cases because of the approximation of $M_L(\nu,\tau)$. For low signal-to-noise ratio a maximum likelihood fit by the modulated model should produce a correct estimation of the parameters (Fig. 3), but for higher signal-to-noise ratio the estimation of mode parameter by using the analytical gap filled profile might produce some bias which will mainly affect asymmetry and noise (Fig. 4).

When the gap filling method does not produce a one-hundred-percent duty cycle recovery on the gap filled spectrum, the computation of the modulation function remains the same. The differences are in the values of the partial duty cycle ratios α_m and in the mode model used which must include sidelobes (Gelly et al. 1998).

4 RECOVERING THE CORRECT VALUE ?

Monte-Carlo simulations were performed to assess whether we can recover the p-mode parameters without significant bias. Five hundred time series lasting 2 months and sampled at 45 s were simulated. The window function used is a genuine IRIS window function (David Salabert private communication) modified in order to obtain a one-hundred-percent duty cycle after gap filling. When the gap filling method is applied with the original IRIS window function the duty cycle reachs from 55% to 96%. For reaching a 100 % duty

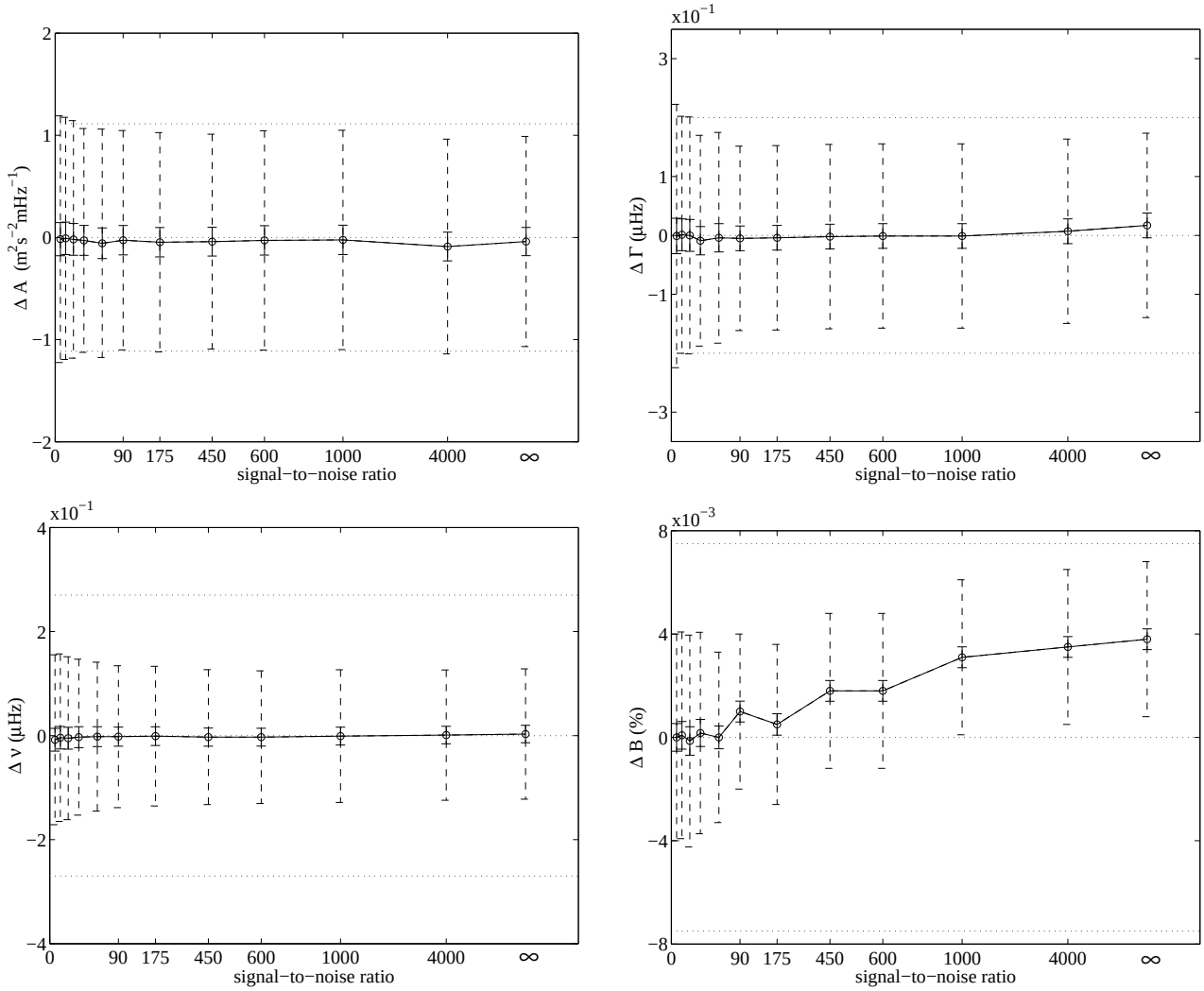

Figure 5. Recovery as a function of signal-to-noise ratio, N=500 simulations. Differences between the gap filled parameters a_f and the target value a_0 used in the simulations; top left: amplitude ($\Delta A = <A_f> - A_0$), bottom left: frequency ($\Delta\nu = <\nu_f> - \nu_0$), top right: linwidth ($\Delta\Gamma = <\Gamma_f> - \Gamma_0$), bottom rigth: asymmetry ($\Delta B = <B_f> - B_0$). The dashed error bars are error bars corresponding to one single realization (σ_{a_f}), and the solid ones are the error bars at 3σ on Δa_f ($3\sigma_{a_f}/\sqrt{500}$). The upper and the lower dotted lines correspond to the relative dispersion of the parameter: $\Delta A/A = \pm 25\%$, $\Delta\Gamma/\Gamma = \pm 25\%$, $\Delta B/B = \pm 25\%$ and $\Delta\nu/\nu = \pm 0.01\%$. As expected from Fig. 4 the asymmetry is the most sensible mode parameter to signal-to-noise ratio.

cycle the window function has been modified by substituting a 0 by a 1 on 4% of the time serie length, it corresponds to the presence of gap without data located 4 hours before and after. After filling the gaps the spectra are fitted as described above. The goal is to make the comparison between the parameters fitted on a spectrum of ungapped data and the parameters fitted on a spectrum of gap-filled data.

4.1 Variations with signal-to-noise ratio

Monte Carlo simulations of a single asymmetric mode (Nigam et al. 1998) are computed for several signal-to-noise ratios in order to validate the approximation in the mode modulation function $M_{lc}(\nu, \tau)$. The gap filled parameters $\{\mathbf{a_f}\} = \{A_f, \nu_f, \Gamma_f, B_f\}$, and their error bars, are estimated by modeling their probability density function using N=500 statistical realizations.

Figure 5 shows the recovery of single mode parameters as a function of signal-to-noise ratio. The approximation of the mode modulation function does not lead to a measurable bias for frequency. For linewidth and amplitude, the bias becomes significant for signal-to-noise ratio greater than 1000 whereas this limit decreases to 80 for asymmetry. As expected from Figs. 3 and 4 the approximation in the analytical mode profile produces a measurable bias for high signal-to-noise ratio affectying mainly asymmetry. However, the bias for asymmetry is well within the error bar of a single realization for signal-to-noise ratio lower than 600.

The analytical model is not strictly valid for high signal-to-noise ratio, and differs from a perfect one in the wings of the mode profile, therefore frequency estimates are not affected by the variation of signal-to-noise ratio. The estimation on amplitude and linewidth for signal-to-noise ratio higher than 1000 gives a measurable bias: the amplitude is under-estimated ($\Delta A < 0$) and the linewidth overestimated

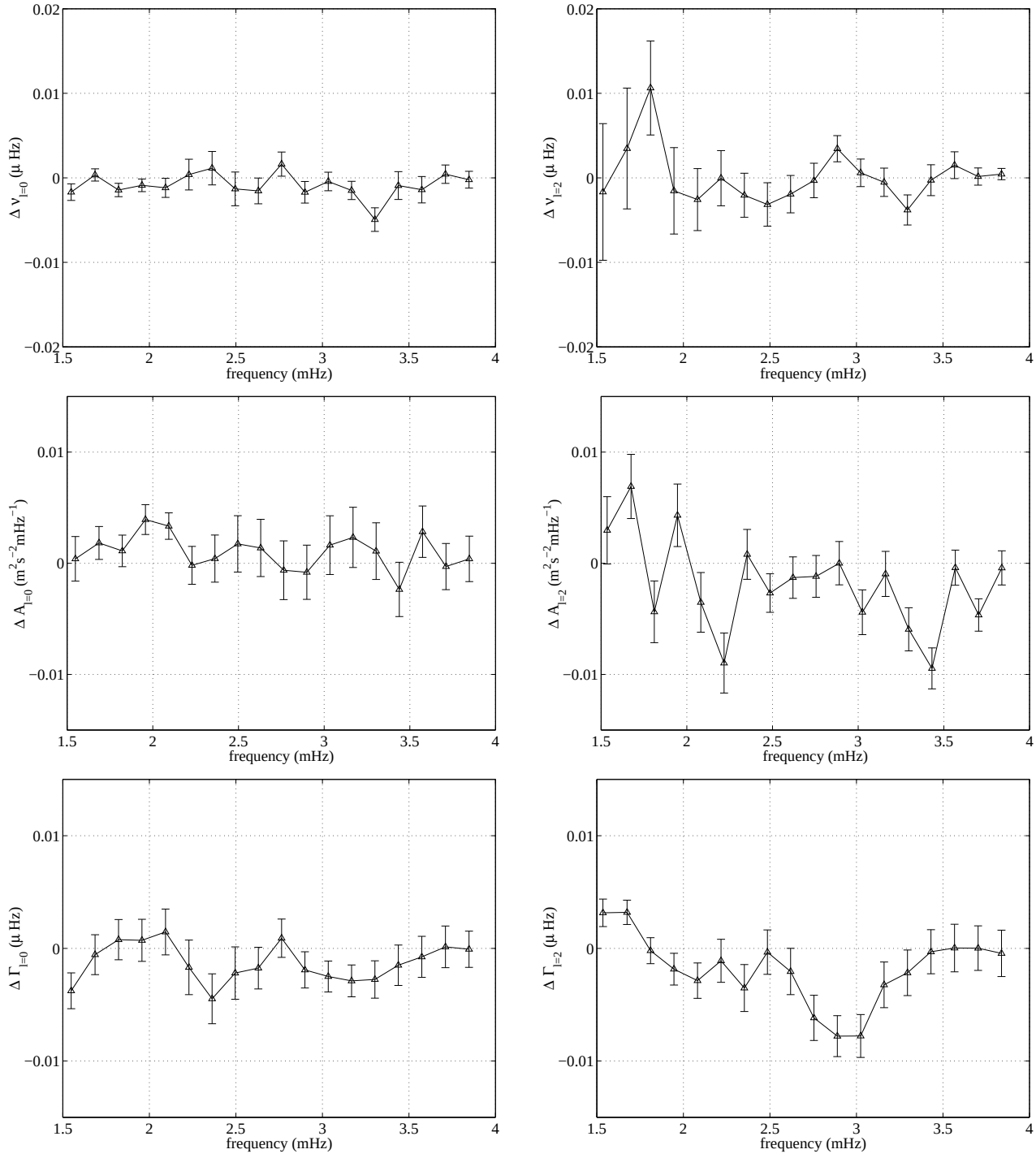

Figure 6. Recovery as a function of frequency, N=500 realizations. The values of $\Delta\mathbf{a} = <\mathbf{a_f} - \mathbf{a}>$ and of the uncertainties are estimated on their probability density function. For one Monte Carlo simulation, $\mathbf{a}$ is estimated on the ungapped spectrum using the mode profile $L(\mathbf{a}, \nu)$ and $\mathbf{a_f}$ is estimated on the gap-filled spectrum using the mode profile $L_m(\mathbf{a_f}, \nu, \tau)$. The error bars plotted are 1-σ error bars corresponding to $\sigma_{\Delta\mathbf{a}}/\sqrt{500}$. There is no evidence of a systematic bias as a function of the frequency, and therefore function of the mode lifetime. Even if the difference $\Delta\mathbf{a}$ does not always match 0 at a 1-σ level, 84 % of the points match 0 at a 2-σ level, the values of $\Delta\mathbf{a}$ are really small and are well negligible when compared to the uncertainty of a single realization ($\sigma_{\Delta\mathbf{a}}$).

($\Delta\Gamma > 0$). This is due to an over-estimation of the noise, as shown in figure 4 a fit on the gap filled spectrum gives a greater noise estimate than the target one.

4.2 Variations with frequency.

The data used are simulated velocity in full-disk measurement with a duration of observation of 2 months. The signal-to-noise ratio varies from 1 to 200. Only $l = 0$, 1 and 2 p modes from 1.5 mHz to 4 mHz are present.

Monte-Carlo simulations of ungapped and gap filled spectra are used for estimating the parameters and their error bars. The gap filled spectrum is computed using the value of τ corresponding to the mode group studied ($\Delta\tau/\tau$=$\pm$0.04%). The precise value of the filling period is determined by placing the maximum of the modulation on the mean frequency of a mode group ($\cos^2(\pi\nu\tau)$=1 for which $\nu = (\nu_{l=0} + \nu_{l=2})/2$).

Figure 6 shows that there is no systematic bias for all parameters but the linewidths. There is an indication that the linewidths might be underestimated by no more than 1 %. The differences between the gap filled and the ungapped parameters are clearly negligible when compared to the error bars for a single realization, for $\sigma_{\Delta\mathbf{a}} \ll \sigma_{\mathbf{a}}$ ($\sigma^2_{\Delta\mathbf{a}} = \sigma^2_{\mathbf{a_f}} + \sigma^2_{\mathbf{a}} - 2\sigma_{\mathbf{a_f}}\sigma_{\mathbf{a}}r$, where the correlation between the two parameters r is close to 1).

5 DISCUSSION AND CONCLUSION

An analytical mode profile has been tested for gap filled spectra obtained by the "Repetitive Music" method (Fossat et al. 1999). This modulated profile depends on the distribution of gap in time series and can be estimated only with the knowledge of the temporal window function.

We show that fitting gap filled spectra using this analytical model allows to recover the correct p-mode parameters, when the signal-to-noise ratio is lower than 600 and when the correct value of the filling period τ is used. The value of τ is settled by centering the $\cos^2(\pi\nu\tau)$-type function on the frequency of the group of modes studied, however we can imagine some refinement of the method by centering the modulation on a mode instead of a group of modes (David Salabert, private communication).

This gap-filling method can be used for any solar seismic data (either ground- or space-based) but also for any stellar seismic data. For example, stellar data acquired at one site have typically 16-hour data gap. For such gaps, the ideal candidate would be a star with a large separation corresponding to 68 μHz giving maximum temporal correlation at 8 hours (i.e. respectively 136 μHz and 4 hours in the solar case). This would put the duty cycle to 100%. Unfortunately, this 'best' scheme will produce a modulated spectrum that will look like what is looked for, e.g. regularly-spaced peaks. This gap-filled method is in any case rather suited when the data are partially replicated: two stations is better than one.

The interest of the method is then to increase the duty cycle of any gapped asteroseismic data without losing information on p-mode parameters. The signal-to-noise ratio is increased and the presence of sidelobes is limited leading to a p-mode parameters estimation with a better accuracy, and negligible systematic bias. This gap filling method, and the derived mode profile, can be applied in the p-modes frequency range on all kind of full disk data. It concerns mainly ground based experiment looking at the Sun (BISON, IRIS, LOWL-integrated, GONG-integrated) or Procyon (ELODIE) but also space data (ACRIM).

ACKNOWLEDGMENTS

The authors would like to thank Mr D. Salabert, from the IRIS team of the University of Nice, for the production of a genuine IRIS window function.

REFERENCES

Anderson E.R., Duvall Jr T.L., Jefferies S.M., 1990, ApJ, 364, 699

Appourchaux, T., Chang, H. -.Y, Gough, D. O. & Sekii, T. 2000, MNRAS, 319, 365

du Pré, F.K., 1950, Phys. Rev., 78, 615

Fierry Fraillon D., Gelly B., Schmider F.X., Hill F., Fossat E., Pantel A., 1998, Astron. Astrophys., 333, 362

Fossat E., Kholikov Sh., Gelly B., Schmider F.X., Fierry Fraillon D., Grec G., Pallé P.L., Cacciani A., Ehgamberdiev S., Hoeksema J.T., Lazrek M., 1999, Astron. Astrophys, 343, 608

Gabriel M., Grec G., Renaud C., 1999, Astron. Astrophys., 338, 1109

Gelly B., Fierry Fraillon D., Fossat E., Pallé P.L., Cacciani A., Ehgamberdiev S., Grec G., Hoeksema J.T., Khalikov S., Lazrek M., Loudagh S., Pantel A., Regulo C., Schmider F.X., 1997, Astron. Astrophys., 323, 235

Harvey, J., 1985, in 'Future Missions in Solar, Heliospheric and Space Plasma Physics', Eds E. Rolfe and B.Battrick, ESA Publications Deprtment, Noordwijk, The Netherlands, 199

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

Toutain T. and Appourchaux T., 1994, Astron. Astrophys., 289, 649

APPENDIX A:

A1 Phase estimate

The main difficulty in the computation of the mode modulation function is the estimation of the phase of the velocity cross terms:

$$< \widehat{V}_k(\nu)\widehat{V}_{k'}(\nu) >= L_{kk'}(\nu)e^{i\phi_{kk'}(\nu)} \qquad k \neq k' \tag{A1}$$

Figure A1 shows the phase for four differents signal-to-noise ratio of the $< \widehat{V}_0(\nu)\widehat{V}_{upp}(\nu) >$ cross terms, which gives the most important contribution ($\alpha_0 = 51\%$ and $\alpha_{upp} = 15\%$), in the case of a single mode. As expected, the cross terms phasis tend toward their minimum when the signal-to-noise ratio is decreasing. Figure A3 shows that $\cos(\phi_{kn}(\nu))$ can be assumed equal to 1 in a frequency range close to the mode frequency. On the other hand, as shown in Fig. A2, $\sin(\phi_{kk'}(\nu))$ can be assumed close to 0 in the same frequency range.

A2 Neglecting $\epsilon_2(\nu,\tau)$

Let's note $\epsilon_2(\nu,\tau)$ the contribution to the modulation function of the imaginary part of the velocities cross terms:

$$\epsilon_2(\nu,\tau) = \sum_{k,k'} J_{kk'}\mathcal{I}(f_k(\nu,\tau)f_k^*(\nu,\tau)) \tag{A2}$$

where $\mathcal{I}$ means the imaginary part of a complex number and:

$$J_{kk'} = \alpha_k\alpha_{k'}\sin(\phi_{kk'}(\nu)) \tag{A3}$$

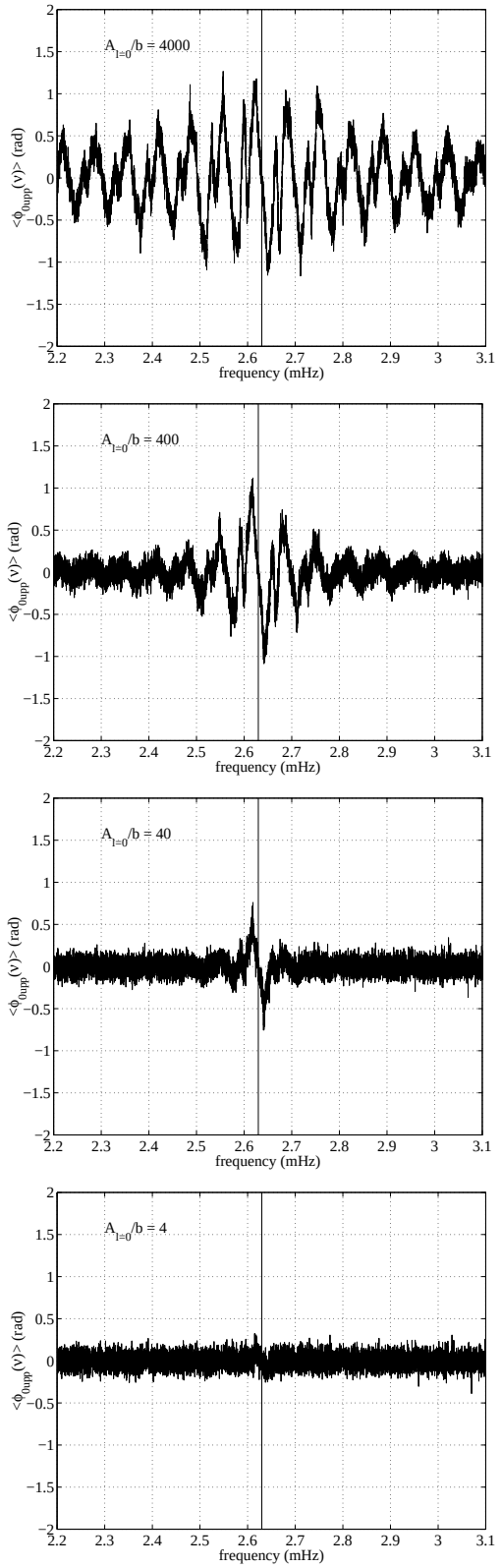

Figure A1. $<\phi_{0upp}(\nu)>$ for four differents signal-to-noise ratio in the case of a l=0 mode ($\nu_{l=0}$ = 2629.73μHz, solid line). Mean over 500 realizations of the phasis of the velocity cross-term $<\widehat{V}_0(\nu)\widehat{V}_{upp}(\nu)>$, which gives the main contribution to the cross term computation ($V_0(t)$ and $V_{upp}(t)$ correspond respectively to 51% and 15% of $V_g(t)$ for the window function used). When the signal-to-noise ratio is decreasing, the cross term phasis tend towards zero in the whole frequency range except close to the mode frequency.

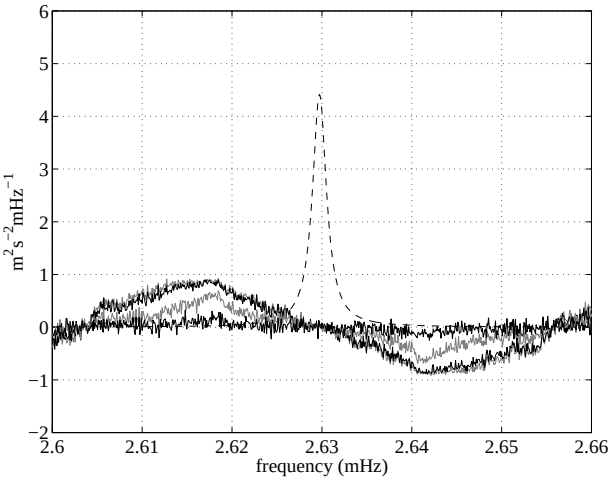

Figure A2. $<\sin(\phi_{0upp}(\nu))>$ for four different signal to noise ratios (same than figure A1). Dashed line: mode profile. Gray/black: mean over 500 realizations of $<\sin(\phi_{0upp}(\nu))>$, the value of $|<\sin(\phi_{0upp}(\nu))>|$ increase when the signal-to-noise ratio is increasing. Note that the values for $A_{l=0}/b$=4000 (gray) and for $A_{l=0}/b$=400 (black) are similar, whereas it is clearly lower for $A_{l=0}/b$=40 (gray) and almost zero for $A_{l=0}/b$=4 (flat black).

Then:

$$
\begin{aligned}
\epsilon_2(\nu,\tau) \;=\; & \sin(2\pi\nu\tau)(2J_{0p} - 2J_{0m} - J_{0um} + J_{0up} - J_{0upp} \\
& + J_{0umm} - J_{pumm} + J_{mupp} + J_{mum} \\
& - J_{pup} - 4J_{pm} + 2J_{umup} + 2J_{umumm} \\
& - 2J_{upupp} - 3J_{mup} - 3J_{pum} + 3J_{mumm} \\
& - 3J_{pupp} - 2J_{uppumm}) \\
+ \; & \sin(4\pi\nu\tau)(J_{mum} - J_{pup} - 2J_{pm} + \frac{3}{2}J_{umup} \\
& + J_{umumm} - J_{upupp} + 2J_{mup} - 2J_{pum} \\
& + 2J_{mumm} - 2J_{pupp} - \frac{1}{2}J_{umupp} \\
& - \frac{1}{2}J_{uppumm} - \frac{1}{2}J_{upumm}) \\
+ \; & \sin(2\pi\nu\tau)(\cos^2(\pi\nu\tau))(J_{uup} - J_{uum} \\
& + J_{uumm} - J_{uupp} \\
& + 2J_{mu} - 2J_{pu}) \\
+ \; & \sin(2\pi\nu\tau)(4\cos^2(\pi\nu\tau) - 1)(2J_{mpm} - J_{pmum} \\
& - J_{pmupp} + J_{pmumm} \\
& - 2J_{ppm} + J_{pmup}) \quad \text{(A4)}
\end{aligned}
$$

In a frequency range close to the mode frequency $\epsilon_2(\nu,\tau)L(\nu)$ can be neglected because $\epsilon_2(\nu,\tau)$ depends on the product between $\sin(2p\pi\nu\tau)$ (p=1, 2) and $\sin(\phi_{kk'}(\nu))$. On one hand, the value of τ is chosen in order to have $\cos(2p\pi\nu\tau)$=1 for the mode frequency, $\sin(2p\pi\nu\tau)$ is then close to 0 in a frequency range close to the mode frequency. On the other hand, as shown in Fig. A2, $\sin(\phi_{kk'}(\nu))$ can be assumed close to 0 in the same frequency range. For high signal-to-noise ratio, neglecting $\epsilon_2(\nu,\tau)$ produce a bias in the p-mode parameter measurement: $\epsilon_2(\nu,\tau)L(\nu)$ gives a non-negligible contribution in the wings of the mode profile (Fig. A2) leading to an over-estimation of the noise and af-

fecting parameter estimation of asymmetry, amplitude and linewidth.

A3 Estimate of $\epsilon_1(\nu,\tau)$, the mode modulation function

Figure A3 shows that $\cos(\phi_{kk'}(\nu))$ can be assumed close to 1 in a frequency range close to the mode frequency. Therefore, the mode modulation function for the velocity cross terms $M_{lc}(\nu,\tau)$ is given by:

$$
\begin{aligned}
M_{lc}(\nu) \quad = \quad & \sum_{k,k'} \alpha_k \alpha_{k'} M_{lc,k,k'}(\nu,\tau) \\
& +2(\alpha_0\alpha_p + \alpha_0\alpha_m + \alpha_0\alpha_u + \alpha_0\alpha_{pm})\,[1+c_1] \\
& +(\alpha_0\alpha_{um} + \alpha_0\alpha_{up})\,[2+3c_1] \\
& +(\alpha_0\alpha_{upp} + \alpha_0\alpha_{umm})\,[2+c_1] \\
& +2(\alpha_p\alpha_m)\,[1+2c_1+c_2] \\
& +2(\alpha_p\alpha_u + \alpha_m\alpha_u)\,[1+c_1]^2 \\
& +2(\alpha_p\alpha_{pm} + \alpha_m\alpha_{pm})\,\left[1+3c_1+2c_1^2\right] \\
& +(\alpha_p\alpha_{um} + \alpha_m\alpha_{um})\,[3+3c_1+2c_2] \\
& +(\alpha_p\alpha_{up} + \alpha_m\alpha_{up})\,[4+5c_1+c_2] \\
& +(\alpha_p\alpha_{upp} + \alpha_m\alpha_{umm})\,[2+3c_1+c_2] \\
& +3(\alpha_m\alpha_{upp} + \alpha_p\alpha_{umm})\,[1+c_1] \\
& +\alpha_u\alpha_{pm}\,\left[1+3c_1+2c_1^2\right] \\
& +(\alpha_u\alpha_{um} + \alpha_u\alpha_{up})\,\left[2+5c_1+3c_1^2\right] \\
& +(\alpha_u\alpha_{upp} + \alpha_u\alpha_{umm})\,\left[2+3c_1+c_1^2\right] \\
& +(\alpha_{pm}\alpha_{up} + \alpha_{pm}\alpha_{um})\,\left[2+7c_1+6c_1^2\right] \\
& +(\alpha_{pm}\alpha_{upp} + \alpha_{pm}\alpha_{umm})\,\left[2+5c_1+2c_1^2\right] \\
& +\alpha_{um}\alpha_{up}\,\left[4+6c_1+\frac{5}{2}c_2\right] \\
& +\alpha_{upp}\alpha_{umm}\,\left[2+2c_1+\frac{1}{2}c_2\right] \\
& +(\alpha_{um}\alpha_{upp} + \alpha_{up}\alpha_{umm})\,\left[3+4c_1+\frac{1}{2}c_2\right] \\
& +(\alpha_{um}\alpha_{umm} + \alpha_{up}\alpha_{upp})\,\left[\frac{5}{2}+6c_1+c_2\right] \qquad \text{(A5)}
\end{aligned}
$$

where $c_1 = \cos(2\pi\nu\tau)$ and $c_2 = \cos(4\pi\nu\tau)$.

Considering $L(\nu)<\cos(\phi_{kn}(\nu))>=L(\nu)$ for high signal to noise ratio becomes wrong in the wings of the mode profile (Figs. A2, A3 and 4).

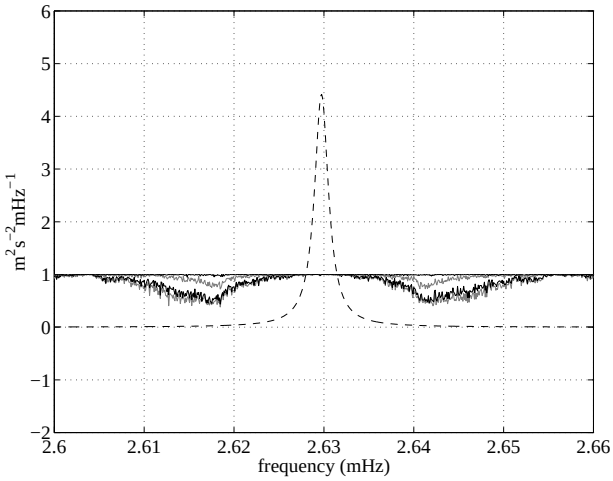

Figure A3. $<\cos(\phi_{0upp}(\nu))>$ for four different signal to noise ratios (same than figure A1). Dashed line: mode profile. Gray/black: mean over 500 realizations of $<\cos(\phi_{0upp}(\nu))>$, the value of $|<\cos(\phi_{0upp}(\nu))>|$ decrease when the signal-to-noise ratio is increasing. Note that the values for $A_{l=0}/b$=4000 (gray) and for $A_{l=0}/b$=400 (black) are similar, whereas it is clearly lower for $A_{l=0}/b$=40 (gray) and almost one for $A_{l=0}/b$=4 (flat black).