# SIMULATING A FIRST SIGHT: SPECTRAL LINE SURVEYS AT THZ FREQUENCIES

**C. Comito and P. Schilke**

Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany

Abstract

Complete unbiased line surveys will form an important part of the HIFI instrument on board of *FIRST*'s operation, and much observing time will be devoted to them. To be properly analyzed, spectral surveys require a process of deconvolution from double-sideband to single-sideband spectra. The quality of the reconstructed SSB spectrum, achieved through the Maximum Entropy Method, varies depending on several factors, namely i) the spacing between DSB scans, ii) the pointing errors, and iii) the sideband gain ratio. The aim of our work is to optimize the general procedure for line surveys observations, by minimizing the errors introduced by those factors. The analysis consists simulating DSB observations over the 960-1000 GHz band while letting the selected parameter(s) vary. The reconstruction of the SSB spectrum is than quantitatively compared to the model spectrum. In this preliminary report, only the spacing and the pointing offsets effects have been considered.

Key words: Stars: formation – Missions: FIRST – Data processing techniques: sideband deconvolution

## 1. Introduction

The *Far InfraRed and Submillimeter Telescope* will produce copious amounts of data, which we must be prepared to process and analyze properly. With this in mind, we are currently generating synthetic spectra of high-frequency molecular line emission towards star forming regions, in order to simulate observations of these objects with the *FIRST* HIFI instruments, and implement an efficient data reduction software.

Complete line surveys will form an important part of *FIRST*'s operation, and a significant amount of time will be dedicated to them. Our main aim, at this stage of development, is to **optimize the general procedure for line surveys observations**. During a line survey, a selected frequency band is covered by a series of double-sideband (DSB) spectra of given bandwidth. Once the data are collected, an appropriate analysis can be carried out only after the DSB scans have been deconvolved into a single-sideband (SSB) spectrum. The reconstruction method our analysis is based on is a variant of the *Maximum Entropy Method* (MEM), and has been used for processing the 607-725 GHz line survey of Orion-KL by Schilke et al. (2001). The algorithm for sideband deconvolution is described by Sutton et al. (1995); an extensive explanation of the principle of Maximum Entropy is provided by Gull & Skilling (1984).

Perfect DSB spectra can be deconvolved perfectly, without introducing any additional noise. In the following, we will describe the effects introduced by non-perfect DSB spectra. We must point out that, due to the non-linear nature of the MEM reconstruction, its properties cannot be derived analytically, and simulations have to be performed.

The reconstructed SSB spectrum is mainly affected by three factors:

a) the **spacing**, in frequency, between the scans. This parameter determines:
   i) the degree of redundancy of the collected information, the redundancy being minimum when the spacing corresponds to the bandwidth of the spectra. An optimal reconstruction would require the degree of redundancy to be as high as possible (i.e. the spacing should be as small as possible);
   ii) the time spent on source and the dead time spent on tuning the receiver (the smaller the spacing, the higher the number of tunings during the coverage), hence the r.m.s. noise in the scans. Here the ideal case is obviously that by which the least time is lost in tuning, i.e. the noise is as low as possible, which can be achieved by having the longest possible spacing.
b) the **pointing offset** for each DSB scan: every on-source scan is affected by a pointing error, and this causes spectral lines to show different intensities in different spectra. Because of the ambiguity of the information contained in the DSB scans, the deconvolution algorithm might fail in assigning the right intensity to each SSB channel. As a result, spurious lines might exist in the deconvolved spectrum. This kind of error can be countered by high redundancy in the data. It should also be kept in mind that this effect is much stronger for sources structured on scales which are comparable to the beam size, since relative intensities *within* a DSB spectrum can vary depending on pointing. Here we will not investigate this last aspect further, but note

Proc. Symposium *'The Promise of the Herschel Space Observatory'* 12–15 December 2000, Toledo, Spain
ESA SP-460, July 2001, eds. G.L. Pilbratt, J. Cernicharo, A.M. Heras, T. Prusti, & R. Harris

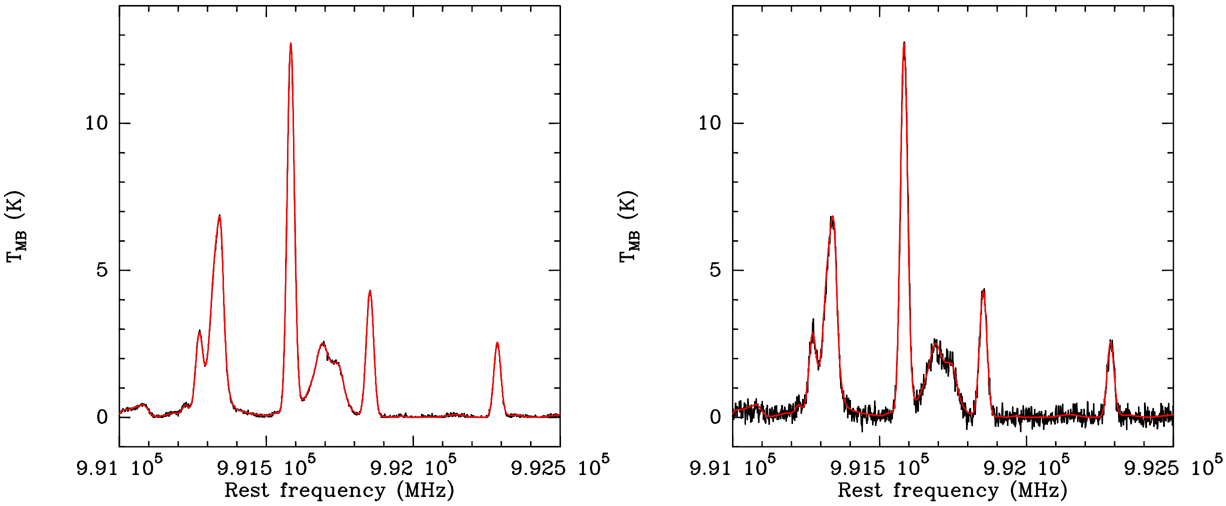


*Figure 1.* *(***left***) model spectrum (red) superposed to the SSB spectrum (black) resulting from the MEM deconvolution of 74 DSB spectra (minimum redundancy and minimum r.m.s. noise); (***right***) the model spectrum (red) is superposed to the SSB spectrum (black) resulting from the MEM deconvolution of 1000 DSB spectra (spacing = bandwidth/14, high redundancy but high noise as well).*

that ground-based surveys, such as that of Schilke et al. (2001) and Comito et al. (in prep.), suffer from it;

c) the **sideband gain ratio**: the sideband balance of the receiver is unlikely to be perfect, and the value of the sideband ratio (ideally = 1) will be different for every tuning. The effects of a sideband imbalance will be similar to those produced by pointing errors (see above), and again they can be reduced by a high redundancy of information in the DSB spectra.

The preliminary results shown here only take into account the variations of the deconvolution output as the *spacing between scans* and the *pointing offset* were (separately) varied.

We proceeded as follows:

– using the *XCLASS* software developed by us, we have produced a model spectrum representing the 960-1000 GHz band. The spectrum was created on the basis of the molecular abundances inferred by Schilke et al. (2001) for Orion-KL, and of the molecular data from the JPL catalog. The diameter of the telescope (3.5 m) and its half power beam width (21″ at these frequencies, *The HIFI Science Homepage*[1]), were also taken into account in the simulation. A series of DSB observations was then simulated, on the basis of the model spectrum, using the following setup:
  → the band was covered by a number of DSB spectra, each of bandwidth 500 MHz and spectral resolution 1 MHz;
  → the total integration time for covering the 40-GHz band (on source + tuning) was set to 1 hour, which gives a minimum r.m.s. noise of 0.03 K/channel;

– we repeated the coverage several times, each time varying the selected parameter (that is, reducing the spacing or adding a pointing offset).

## 2. Spacing effects

As already mentioned, a reduction of the frequency spacing produces, in the DSB spectra, an increase in the redundancy of information (which, by reducing the ambiguity of the data, makes the deconvolution more reliable), and a growth of r.m.s. noise (which makes the deconvolution results more uncertain). In order to determine whether these opposite effects can be balanced in a sort of optimal configuration, we have produced a coverage of our 40-GHz band using eight different values for the spacing, the initial value being 500 MHz (band coverage achieved with 74 scans, minimum redundancy) and the final being 500/14 MHz (coverage achieved with 1000 scans, high redundancy of information). Accordingly, r.m.s. noise on DSB spectra increased, following the

$$\sigma_{\rm rms} = \frac{T_{\rm sys}}{\sqrt{\Delta\nu \times t}}, \tag{1}$$

[1] http://www.sron.rug.nl/hifiscience/

where $\Delta\nu$ is the spectral resolution (Hz), $t$ is the total integration time per scan[2] (sec) and $T_{\rm sys}$ is the system temperature. At these frequencies, the system temperature is expected to be around 210 K (*The HIFI Science Homepage*). Gaussian random noise was added to each scan.

For each coverage, the DSB spectra were deconvolved into a SSB spectrum by using the Maximum Entropy Method. The resulting spectrum was then compared to the original model. A qualitative comparison between the deconvolved spectrum that best matches with the model, and the worst-matching one, is shown in Fig. 1.

In order to obtain a quantitative estimate of the "goodness" of a deconvolved spectrum with respect to the model, we have performed a channel-by-channel difference between each deconvolved spectrum and the model. We then took the area of the absolute of the difference spectra (hereafter $A_{|\rm diff|}$) as a "fidelity parameter" - the smaller the area, the better the deconvolution relative to the model. We have also considered different receiver tuning times, from 0 up to 3 sec. The results are summarized by the solid curves in Fig. 2, which represent the variation of $A_{|\rm diff|}$, expressed in K $\times$ channel$^{-1}$, as a function of the number of tunings per unit bandwidth (indicated with $N$).

From these plots, we find that the negative effect of the increasing noise quickly overwhelms the positive one of the increasing amount of information in the data: the optimal setup seems to be that with a minimum redundancy of data. We then repeated the modeling with a reduced amount of noise (dashed curve in Fig. 2). Note that the "bump", shown by the dashed curve between $N = 4$ and $N = 10$, is due to "numerical noise" introduced by the deconvolution processes, by an as yet unknown cause. However, a general trend can still be inferred: the curve reaches a minimum at $N = 2$ and then starts to grow again, so that the best result is achieved with a frequency spacing of half a bandwidth (250 MHz).

## 3. Pointing offset effects

The second parameter we have decided to take into account is the pointing offset. As already explained in section 1, pointing errors are likely to cause, during the deconvolution from DSB to SSB, a misassignment of intensities to some (or many) of the lines. While a reduced, or increased, line intensity might not, in principle, prevent from identifying a transition, problems will arise if some flux is assigned to a frequency at which no line exists at all. It could, in most cases, be very difficult to distinguish between a spurious line produced in this fashion, and a real one belonging to some unknown molecule. In the examples we have considered so far, no strong spurious lines were found, and the effect manifested itself only through incorrect line strengths. However, we expect spurious lines to appear more frequently once sideband imbalance and structured sources are investigated.

In order to determine the optimal strategy to minimize these errors, we have repeated the band coverage, again each time reducing the frequency spacing, and at the same time assigning to every DSB spectrum a random pointing offset (with all offsets belonging to a Gaussian of HPW = 5 arcsec). At this stage, we have not included the r.m.s. noise in this simulation. After the deconvolution, the fidelity of the SSB spectrum to the model was estimated using the same method as in section 2. A plot of $A_{|\rm diff|}$ as a function of $N$ is shown in Fig. 3. Again, we consider the "bump", for $4 < N < 10$, and the rise, for $N > 10$, in this curve not to be realistic (see section 2), and the general trend shows the accuracy of the deconvolution, with respect to the model, to increase with the number of tunings, as expected.

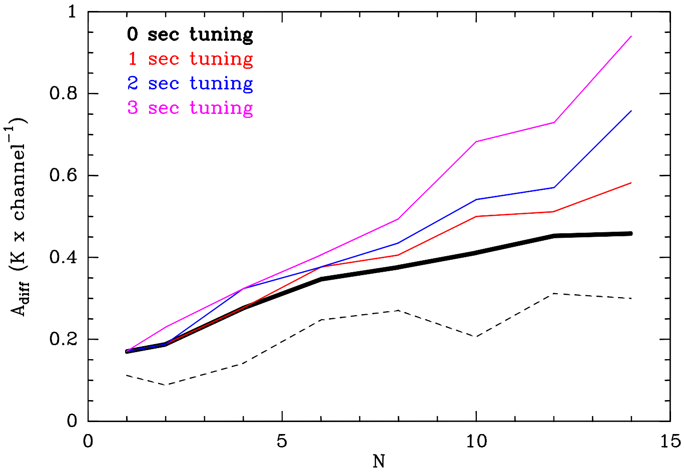


*Figure 2.* *Plot of $A_{|\rm diff|}$ as a function of the number of tunings per unit bandwidth (see text).* **Solid:** *the differently coloured curves represent different tuning times, as indicated in the legend. The best result (smallest area) is obtained with the minimum redundancy setup, which means that the trend is dominated by the increase of noise.* **Dashed:** *the dashed curve is obtained by adding to the DSB spectra a reduced amount of r.m.s. noise (50% less, i.e. 4 times the observing time, than the "0 sec tuning" curve represented by the black solid line). In this case, due to the increase of redundancy and to the low level of noise, the smallest area is obtained for $N = 2$. After this point, the noise dominates again. Note that the "bump" for $4 < N < 10$ is due to numerical problems in the DSB deconvolution. The general trend, anyway, is not affected.*

[2] If $n$ is the number of tunings needed to cover the band, $T$ is the total observing time (1 hr), and $\tau$ the time needed for a single tuning, the total integration time per scan is $t = (T - \tau)/n$.

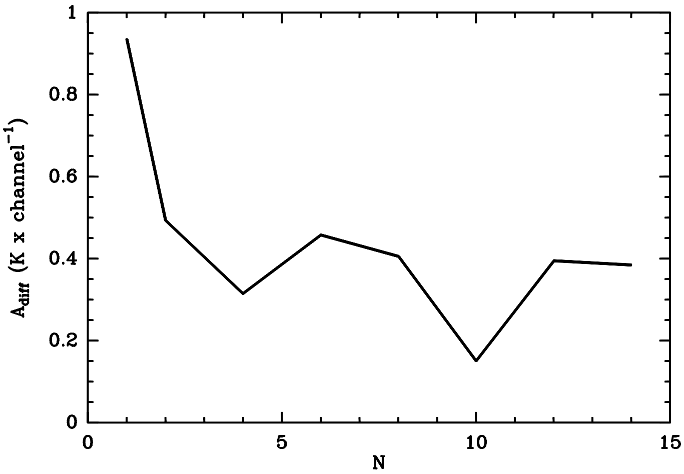


*Figure 3. Plot of $A_{|\mathrm{diff}|}$ as a function of the number of tunings per unit bandwidth, this time including pointing errors in the simulation (and leaving the r.m.s. noise out). Considering that, again, the "bump" for $4 < N < 10$ and the rise for $N > 10$ are not real but due to numerical noise (see also Fig. 2), the general trend shows an improvement of the quality of the deconvolution (decrease of $A_{|\mathrm{diff}|}$) as the redundancy of information in the DSB data grows (increase of N).*

## 4. Discussion

The analysis presented here is still at a very early stage, and a complete study on the optimization of line survey procedures is currently under way.

So far, we have concentrated on the effects that the *variation of the spacing between DSB scans* and the *pointing errors* separately have on the quality of the deconvolved SSB spectrum. Though partially affected by numerical noise introduced by the deconvolution algorithm, our results show that:

a) the quality of the deconvolved SSB spectrum, as the redundant information in the DSB data increases (i.e. as the spacing between the DSB scans decreases), does not improve indefinitely: it reaches a maximum (for a certain value of tunigs per unit bandwidth, $N$) and then falls again. In our example (dashed curve in Fig. 2), this happens for $N = 2$. For $N > 2$, the effects of the increasing noise, introduced by the reduction of the spacing, overcome the effects of the increasing redundancy, and the deconvolved SSB spectrum results deteriorated;
b) in the case in which pointing errors are included in the simulation (but r.m.s. noise is excluded, Fig. 3), the accuracy of the deconvolution increases, in general, with the number of tunings.

Both results confirm our expectations. The next steps towards the completion of this study will consist in:

– studying the effects sideband imbalances on the results of the deconvolution;
– considering *all* parameters (spacing, pointing error and sideband imbalances) together, in order to reproduce a situation as close as possible to "real-life" observations. The errors introduced by the observation of structured sources (see section 1, point b)) must also be taken into account;
– applying these results to currently ongoing ground-based line surveys, in order to verify them.

## References


Gull, S.F., Skilling, J. 1984, Proc. International Symposium on Indirect Imaging, ed. J.A. Roberts, Cambridge University Press, 267

Schilke, P., Benford, D.J., Hunter, T.R., Lis, D.C., Phillips, T.G., ApJSS, in press

Sutton, E.C., Peng, R., Danchi, W.C., Jaminet, P.A., Sandell, G., Russel, A.P.G. 1995, ApJSS, 97, 455S