DOCUMENT

# Herschel/HIFI Spectral Scans Highly Processed Data Product: Release Notes

| | |
|---|---|
| Prepared by | David Teyssier (ESAC) |
| Reference | HERSCHEL-HSC-DOC-2198 |
| Issue | 1 |
| Revision | Release 1.0 |
| Date of Issue | 2017-06-30 |
| Status | For release |
| Document Type | Release Note |
| Distribution | HSC |

**Document approval**

| | | |
|---|---|---|
| Prepared by: | D. Teyssier | Date: 2017-06-30 |
| Approved by: | | Date: |

**Change log**

| **Version** | **Date** | **Change** | **Author** |
|---|---|---|---|
| 1.0 | 2017-06-30 | Initial version of document | David Teyssier |

# 1 Introduction

The Heterodyne Instrument for the Far-Infrared (HIFI, (7)) was the high-resolution spectrometer on-board *Herschel*, offering a continuous spectral coverage in the ranges [480–1270] and [1430–1910] GHz with a resolving power in excess of $10^7$. One of the richest output of the HIFI scientific database lies in the hundreds of spectral line surveys it conducted using its Spectral Scan observing mode. Overall, HIFI collected around 500 observations in this mode, over about a hundred of different lines-of-sight, in partial or full spectral coverage.

Although the HIFI pipeline (12) generally did a very good job in calibrating out the instrument bandpass and its drift with time, residual baseline artefact may still be present in the final products (see also Section 5.3 of the HIFI Handbook[1]). Because the prime product from Spectral Scan observations are the Level 2.5 deconvolved spectra resulting from the processing of all individual Level 2 products, the removal of baseline artefacts on the whole deconvolved data is not a trivial task. Instead, the best approach consists in cleaning all individual Level 2 spectra, and feed them into the deconvolution algorithm. This Highly-Processed Data Product data-set provides the outcome of such a systematic baseline cleaning exercise performed by instrument experts from the Herschel Science Centre.

This document describes the method used to generate and validate the baseline cleaning of the Spectral Scan residual baseline artefacts. Owing to the semi-automatic approach followed for that purpose, we describe in particular the possible caveats of the resulting data-set, as well as the ancillary and product header information added into the delivered data-sets.

# 2 The HIFI Spectral Scans

## 2.1 Archive content and science readiness

We consider here the 500 Spectral Scan observations obtained in standard observing mode (in contrast with those performed in engineering mode) with one of the three possible referencing scheme for this mode (namely Double Beam Switching, Load Chop and Frequency Switching). Of those, we ignored the 18 observations stamped as non-public, as they usually made used of non-standard observing modes, or instrument configuration settings. The standard product generation pipeline provides calibration deconvolved spectra in each observed polarisation for each observation, and takes care of masking all spurious spectral features still present in the Level 2 products (see e.g. the HIFI Product Explained document[2], or the Spectral Scan cookbook[3]). What the pipeline does not take care of, though, is the treatment of any residual baseline distortion resulting from an imperfect bandpass calibration[1].

[1] http://herschel.esac.esa.int/twiki/pub/Public/HifiDocsEditableTable/hifi_handbook.pdf

[2] http://herschel.esac.esa.int/twiki/pub/Public/HifiDocsEditableTable/HIFI_Products_Explained_v1.pdf

[3] http://herschel.esac.esa.int/hcss-doc-15.0/load/hifi_um/html/hcb_ssc.html

## 2.2 Baseline cleaning method

### 2.2.1 Baseline removal of heterodyne spectra: general concepts

The main challenge of baseline cleaning resides in two fundamental steps: (i) the identification of spectral channel which holds spectral line information (in contrast with those holding only instrument – essentially noise – contribution, and (ii) finding out what is the best representation of the baseline shape one tried to correct from. The first step is basically a line masking allowing the best possible fit of the baseline during the second step. Common practice in radio-astronomy is to approximate the baseline by polynomial models, or sine wave models when well defined standing wave ripples are present in the baselines.

Two dedicated tasks were designed in HIPE[4] for that purpose:

- `fitHifiFringe`[5]: a task optimised to fit sine wave baseline component, based on a Fourier analysis of the input spectra. This task offers the option to also subtract a cubic spline baseline which is fitted automatically during the process
- `fitBaseline`[6]: a task dedicated to removing polynomial baseline fits to the input spectra. Unlike the previous task, the ”best” polynomial order is not assessed automatically by the task but must be set by the user

Both tasks can work in non-interactive mode, making use of a powerful automatic line masking algorithm (so-called `smoothBaseline`, see the task description in HIPE manual[7]). On top or instead of those automatic masks, users can provide their own *ad hoc* line windows.

### 2.2.2 Semi-automatic baseline cleaning in the Spectral Scan HPDP

The main shortcomings of automatic line masking manifest in the case of complex line profiles (especially for broad and/or absorption lines), and for very line rich spectra where the algorithm will struggle to find sufficient line-free spectral channels and discriminate real lines from noise contribution. Because a large fraction of the HIFI Spectral Scan will not be affected by such circumstances, it was expected that automatic baseline subtraction could work successfully “out-of-the-box” in main cases. Depending on the shape and nature of the residual artefacts, however, one of the two above-mentioned tasks could do a better job than the other, or alternatively, they could be use in sequence, with different specific purpose. We have performed a series of combination of task calls and options on each considered observation, and decided on which of these schemes would offer the best output. The decision was based both a on visual inspection of the resulting baseline-corrected deconvolved spectrum, as well as the generation of spectrum of the noise rms over the deconvolved data (it is expected that the best baseline-corrected spectrum will yield the lowest and most homogeneous noise spectrum).

[4] https://www.cosmos.esa.int/web/herschel/hipe-download

[5] http://herschel.esac.esa.int/hcss-doc-15.0/load/hifi_um/html/sw_removal_hifi2.html

[6] http://herschel.esac.esa.int/hcss-doc-15.0/load/hifi_um/html/hifium_bl1.html

[7] http://herschel.esac.esa.int/hcss-doc-15.0/load/hcss_urm/html/herschel.ia.toolbox.spectrum.standingwaves.SmoothBaselineTask.html

Section 4.1 describes the six different schemes considered in the study. Once the best scheme got selected based on the simple criterion explained above, a fine granularity check was performed by eye on each and every observation, especially in order to look for residual narrow features in the baseline residual. Those indeed reveal narrow frequency intervals around lines where improper masking results in over- or under-estimate of the baseline level under the line of interest. The outcome of this analysis was as follows:

- the baseline correction is relatively poor for spectrally-complex sources because the line masking fails in identifying all spectral features. As a consequence we decided to not cover the concerned ObsIDs in the final delivery. They amount to 174 observations. Section 2.2.3 gives further details about the affected data.
- for the other 326 observations where the automatic baseline correction seemed to do a decent job (listed in Table 3), we identified 128 observations where isolated lines had suffered from imperfect line masking. We explain in Section 2.2.4 the measures taken in order to optimise the line masking in those areas.
- A handful of observations was affected by particular spectral ghosts that needed a specific treatment. They are described in Section 2.2.5.

In case residual baseline-correction artefacts are still present in the end product despite of all efforts described above, dedicated flags are then added into the products to warn the user (see Section 3.3).

#### 2.2.3 Observations discarded from the final delivery

We noticed that a sub-sample of particularly spectrally-complex sources had relatively poor baseline correction result, whereby numerous lines would be either missed by the masking, or even worse, confused with baseline structured and consequently removed to a large extend in the final product. Not surprisingly, those sources have often been the topic of Guaranteed Time observing (Key) programmes and have therefore allocated very dedicated resources in their data post-processing and analysis. It seemed to us that those dedicated data curation were more fit than our automatic to provide the best possible baseline-corrected spectra. Among those one can highlight e.g. Orion KL (e.g. (5)), the SgB2 lines-of-sight (e.g. (4), (14)), $\eta$ Carina (e.g. (10)) or line rich evolved stars like IRC+10216 (e.g. (3)), VY CMa (e.g. (1)) or OH231.8+4.2 (e.g. (2)). Out of the 500 Spectral Scan observations initially contemplated, we decided to discard a total of 174 observations (i.e. about 1/3 of the sample). Tables 4 and 5 gives an inventory of the omitted observations, together with references to past or future work expected to provide the missing information in the the baseline-corrected data-set provided by the HSC. Essentially, the expectation is that those products (both baseline corrected and Spectral Line Catalogue) should become available as User-Provided Data Products[8].

#### 2.2.4 Manual line masking and other baseline manipulation

In case the automatic baseline-correction would work sub-optimally around isolated lines, it is most often enough to manually add dedicated line windows around the line of interest to recover a

[8] https://www.cosmos.esa.int/web/herschel/user-provided-data-products

representative baseline under those. This was particular true for absorption lines. This additional step provides very satisfactory output for almost all cases considered. In a few exceptions, even this manual masking still fails to provide perfect baseline fitting, and as a consequence the final baselines could exhibit apparent slightly positive or negative residual continuum level around zero.

In some cases, additional manipulation is needed on the baseline in order to account to residual imperfection that cannot be resolved by a line mask readjustment. We should indeed bear in mind that the line masking is effectively an input to the baseline correction of the individual Level 2 spectra, which come with their own baseline shape and noise. On the other hand, the delivered products is the result of the deconvolution of all those Level 2 spectra, which mixes information from potentially plenty of tunings. As a consequence, in more rare cases final baseline cosmetics have been necessary directly on the deconvolved data in order to repair obviously unruly baseline model in isolated frequency ranges. *This aspect of the baseline-correction adds an unavoidably subjective component to the final product, that end users should keep in mind when exploiting the data.* **As a general rule, it is recommended to always have a close look at the subtracted baseline spectrum also provided with the delivery products (see Section 3.2).**

#### 2.2.5 Bright line and Frequency Switching ghosts

Within the observations that needed further massaging was the particular case of observations with either very strong lines, or taken in frequency switching mode (FSW). In those cases, the deconvolved spectrum would often suffer from spectral ghosts: in case of strong line, this is a side effect of the deconvolution algorithm; in case of FSW this originates from the "negative ghost" resulting from the folding (see e.g. the Frequency Switching cookbook[9]. In both cases the solution consists in flagging the line responsible for the ghosts as `Bright Line`[10] and run the deconvolution doubly, following the recipes described in Section 2.7.4.3 of the Spectral Scan Cookbook[3]. Figure 2 illustrates such a case for a FSW observation.

# 3 Content of the Spectral Scan HPDP

## 3.1 Delivered Spectra and Postcards

The delivered baseline-corrected Spectral Scan data-sets actually contain more than just the baseline-corrected deconvolved spectra in each polarisation (for a couple of exceptions only one polarisation was available[11]). The product made available as HPDP will host the following spectra:

- `dataset`: this is the unmodified (original) spectrum provided as Level 2.5 by the standard pipeline generation

[9] `http://herschel.esac.esa.int/hcss-doc-15.0/load/hifi_um/html/hcb_pfsw.html`
[10] `http://herschel.esac.esa.int/hcss-doc-15.0/load/hifi_um/html/cflags.html`
[11] Observation IDs: 1342196475 (H only), 1342181163 and 1342253144 (V only)

- `baseline-corrected`: this is the baseline-subracted spectrum, obtained from the deconvolution of all cleaned Level 2 spectra
- `baseline`: this is the baseline effectively subtracted from the original spectrum, i.e. it is computed as (`dataset` – `baseline-corrected`)

All three spectra are grouped into one single so-called "SimpleSpectrum" structure saved as a FITS file, named `<obsid>_WBS-<polar>_SpectralScan_HPDP.fits.gz` (where `<polar>` is either H or V). A dedicated postcard is also created in order to illustrate this whole content, and named `<obsid>_Postcard_SpectralScan_HPDP.png`. An example of such postcard is shown in Figure 1. Both FITS files and the postcard are then bundled into one single tar ball file named `<obsid>_SpectralScan_HPDP.tar.gz`. The `baseline` spectrum is particularly important for users to figure out the reliability of the applied baseline correction in case of spectra showing a particularly high line density.

Although the above files will be provided by the Herschel Science Archive[12], they can also be fetched from a dedicated legacy repository: `http://archives.esac.esa.int/hsa/legacy/HPDP/HIFI/HIFI_spectral_scans/`. This directory provides a separate links to the file bundles (`Data`), as well as one allowing to check postcards individually (`Postcards`). The total size of the HPDP data-set is 3.8 Gb.

## 3.2 How to use the isolated baseline spectrum

As explained in the previous section, the delivered data-sets come with a separate spectrum of the subtracted baseline. One should bear in mind that this spectrum is actually the deconvolved residual of all baselines subtracted from the individual Level 2 products. As such this spectrum will be a combination of all removed baseline artefacts, and the potential intrinsic source continuum (see also Section 5.8.5 of the HIFI Handbook). As such this spectrum can be used in order to make an estimate of the source continuum that got removed by the baseline correction process. It is up to the user to decide what is the best fit to the separate baseline spectrum in order to derive this continuum.

The separate baseline spectrum should also be used to inspect any dubious baseline correction especially in spectrally-crowded areas, where the line masking might have performed sub-optimally (Section 2.2.2). We encourage users interested in a particular line to review how the baseline compares with the original data (also provided in the data-set) to confirm that no important line information (e.g. on potential line wings) might have been altered inadvertedly.

## 3.3 Flags

A collection of flags have been used in order to pass on additional information to the users of the quality-enhanced Spectral Scan products. They are most often warnings about peculiarities in the baseline-correction quality data, or simply ancillary information about the way observations were taken which are relevant for the data interpretation. Table 1 lists all possible flags added in the headers of the FITS products, and their meaning.

[12] `http://archives.esac.esa.int/hsa/whsa/`

| **Flag name** | **Description** |
|---|---|
| highLineDensityFlag | Warns against spectra were the average line density is above 1 line/GHz. This can be indicative of less accurate line fit parameters due to sub-optimal baseline correction |
| baselineCorrectionFlag | Warns against possible sub-optimal baseline correction that could lead to less accurate line intensities |
| absLineFlag | Indicates that absorption lines are present in the data. Line masking and baseline correction on such features could be less accurate and should be checked with caution |
| manualBaselineMasking | When set to true is indicates that additional line masks where added by the instrument experts in order to optimise the baseline correction. In those cases, the masks are added in the header (see next rows). Otherwise set to false (fully automated baseline correction) |
| manualMaskMin_n | Lower limit of manual line mask #n (optional) |
| manualMaskMax_n | Upper limit of manual line mask #n (optional) |
| fswFlag | Indicates that the input data were obtained in Frequency Switching mode. Such spectra often had to be cleaned up for residual spectral ghosts |
| freqGroupingFlag | Indicates that the input data were obtained with the frequency grouping option. Such data usually may suffered from more pronounced residual baseline artefacts in the standard pipeline products, making automatic baseline correction more challenging |
| brightLineFlag | Indicates that spectral ghosts due to bright lines are present in the standard pipeline products, and that dedicated bright line flags were involved in the generation of the HPDP |
| offContaminationFlag | Indicate that some lines are affected by OFF position contamination. This will lead to either missed line detection, or inaccurate line flux computation from the line fitting. See (13) for details about the HPDP data-set dedicated to isolated OFF-position spectra for HIFI |

Table 1: List of possible flags added in the FITS file headers.

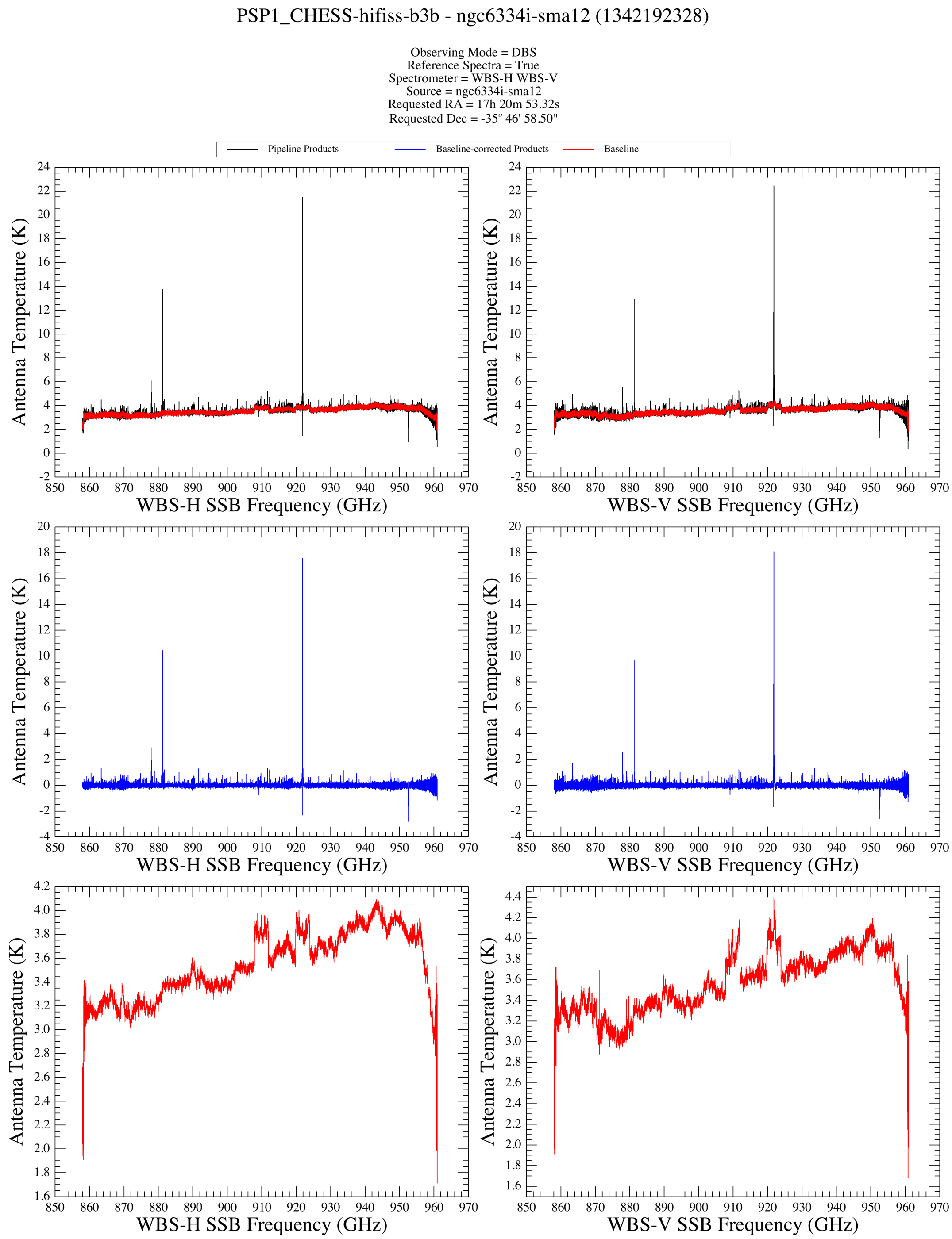


Figure 1: Illustration of a baseline-corrected HPDP for obsid 1342192328. The upper panel shows the products from the standard pipeline generation (black) with the demoed baseline in red. The middle panel (blue spectra) shows the resulting baseline-corrected spectra. Finally the lower panel shows the removed baseline. Note that the removed baseline is composed of the intrinsic source continuum and the residual baseline artefacts.

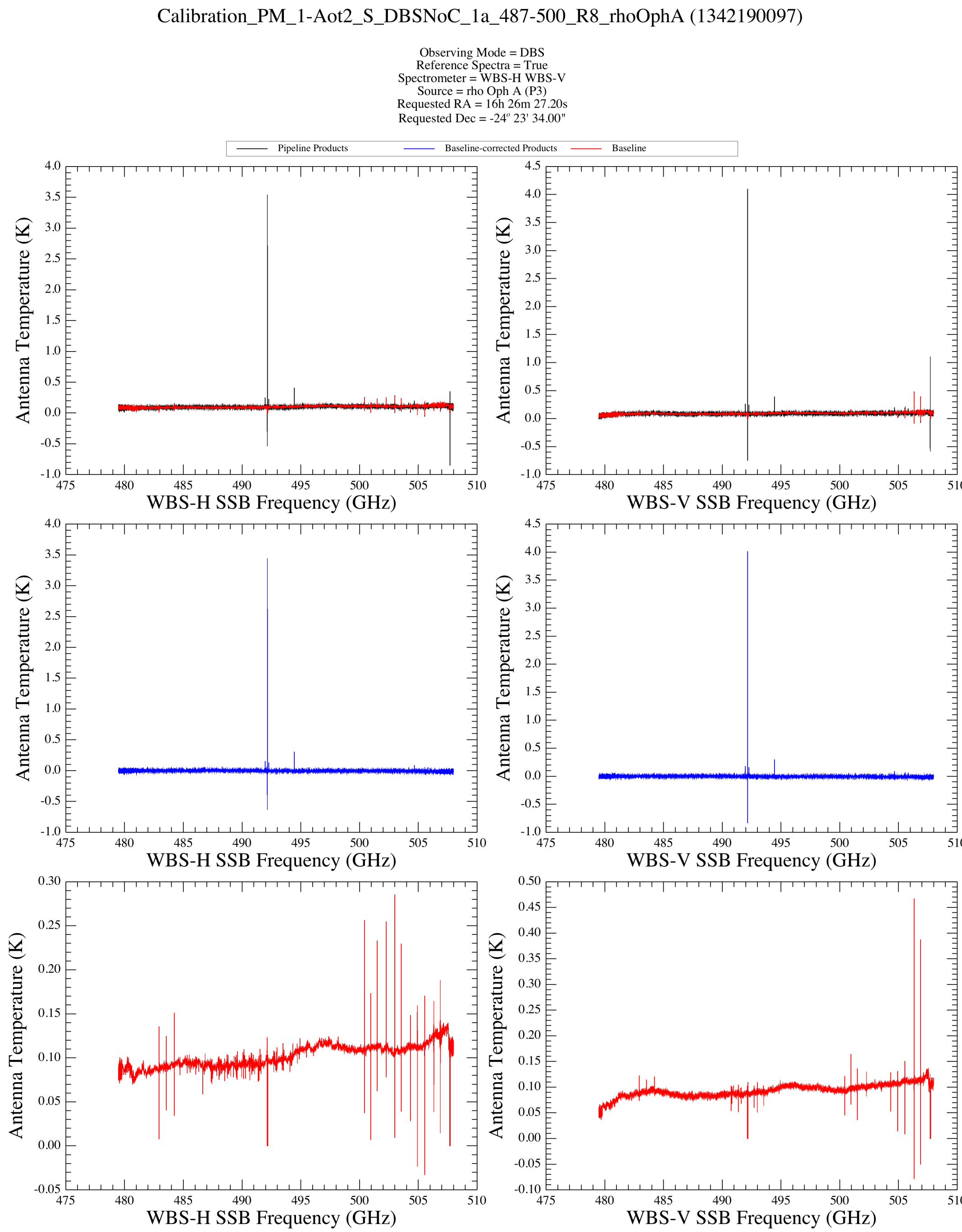


Figure 2: Postcard illustrating a prominent case of ghost spectral artefacts resulting from the presence of a bright line, here in ObsID 1342190097. Note the disappearance of the spectral features above 500 GHz in the baseline-corrected spectra (middle panel), which are also evidenced in the baseline spectrum (lower panel).

# References


[1] Alcolea, J., Bujarrabal, V., Planesas, P. et al., 2013, A&A 559, A93

[2] Sanchez-Contreras, C., Velilla Prieto, L., Agúndez, M. et al., 2015, A&A 577, A52

[3] Cernicharo, J.; Waters R., Decin L. et al., 2010, A&A, 521, L8

[4] Comito C., Schilke, P., Rolffs, R. et al., 2010, A&A 521, L38

[5] Crockett, N.; Bergin, E., Neill, J. et al., 2014, ApJ, 787, 112

[6] Crockett, N.; Bergin, E., Neill, J. et al., 2015, ApJ, 806, 239

[7] de Graauw, T., Helmich, F. P., Phillips, T. G et al., 2010, A&A, 518, L6

[8] Goldsmith P., Liseau R., Bell T., et al., 2011, ApJ 737, 96

[9] Hartogh, P., Jarchow, C., Lellouch, E. et al., 2010, A&A 521, L49

[10] Morris P., Gull, T., Hillier, D. et al., 2017, ApJ, 842, 79

[11] Neill J., Bergin E., Lis, D. et al., 2014, ApJ, 789, 8

[12] Shipman R., Beaulieu S., Teyssier, D. et al., 2017, A&A, in press

[13] Teyssier, D., 2016, *HIFI Reference Position Spectra Data Products: Release notes*, HERSCHEL-HSC-DOC-2111.

[14] Zernickel, A., Schilke, P., Schmiedeke, A. et al., 2012, A&A 546, A87

# 4 Appendix

## 4.1 Baseline correction processing scheme considered in the HPDP generation

| Scheme # | fitBaseline options | | fitHifiFringe options | | |
|---|---|---|---|---|---|
| | doGlue | add Cont.[13] | doGlue | add Cont.[13] | subBase |
| 1 | NA | NA | Yes | Yes | Yes |
| 2 | NA | NA | No | Yes | Yes |
| 3 | No | No | NA | NA | NA |
| 4 | Yes | No | Yes | No | No |
| 5 | NA | NA | Yes | No | Yes |
| 6 | Yes | No | Yes | No | No |

Table 2: Task options considered for given automatic baseline correction schemes. ”NA” means that the task is not invoked. When both tasks are involved, `fitHifiFringe` is always run first. `fitBaseline` considers a polynomial fit of order 1.

[13] full task parameter name is `addMedianContinuum`

## 4.2 List of source considered in the Line Catalogue Generation, and the total frequency coverage of the corresponding observations

| Source Name | Covered frequency ranges (GHz) |
|---|---|
| IC1795-1 | [555.4, 631.2] [958.8, 1050.2] |
| W3(H2O) | [647.8, 676.3] [1139.9, 1168.3] |
| L1448MM-1 | [647.8, 676.3] [1139.9, 1168.3] |
| irc+50096 | [525.8, 551.9] [602.8, 629.9] [690.8, 720] [778.8, 821.1] [866.8, 898.2] [949.9, 987.2] [1053.9, 1089.1] [1141.9, 1179.2] |
| NGC1333 IRAS4A | [626, 800.9] [1139.9, 1168.3] |
| NML Tau | [479.5, 628.5] [641.7, 719.5] [746.4, 775.3] [799.1, 843.3] [858.1, 933.8] [966.2, 1042.2] [1058.7, 1220.2] [1434.6, 1535] [1608.5, 1695] [1713.2, 1797.4] [1832.3, 1906.8] |
| R Dor | [479.5, 560.2] [641.7, 719.5] |
| Orion S | [479.5, 1280] [1426.5, 1532.7] [1573.3, 1906.7] |
| OMC-1 (Peak 1) | [484.7, 501.9] [520.8, 545.1] [771.3, 788.5] |
| Orion Bar CO+ peak | [479.5, 1279.7] [1427.2, 1562.7] [1573.3, 1906.7] |
| OMC-2 FIR4-1 | [479.5, 1243.9] [1489.1, 1508.9] [1535.4, 1556.6] [1592.9, 1638.8] [1650.5, 1680.3] [1707.2, 1734.9] [1756.8, 1777.6] |
| CIT 6 | [525.8, 551.9] [602.8, 629.9] [690.8, 720] [778.8, 821.1] [866.8, 898.2] [949.9, 988.1] [1053.9, 1089.1] [1141.9, 1179.2] |
| CarinaN-point-I F | [958.8, 1050.2] |
| Garradd (C/2008 Q3) | [541.9, 559.5] |
| IRAS 12326-6245 | [514.8, 547.1] [969.9, 1001.2] [1020.4, 1039.9] [1146.8, 1181.1] [1205.8, 1226.5] |
| Y CVn | [525.8, 551.9] [602.8, 629.9] [690.8, 720] [778.8, 820.2] [866.8, 898.2] [949.9, 988.1] [1053.9, 1089.1] [1141.9, 1179.2] |
| G316.81-0.06 | [512.9, 538.2] |
| II Lup | [479.5, 1121.9] |
| G323.74-0.26 | [512.9, 538.2] |
| G327.3-0.6 | [514.8, 547.1] [969.9, 1001.2] [1020.4, 1040] [1146.8, 1181.1] [1205.9, 1226.5] |
| G331.28-0.19 | [512.9, 538.2] |
| rho Oph A | [479.5, 508] [771.1, 776.5] [783.1, 788.5] |
| iras16293-2422A B-1 | [479.5, 1238.4] [1481, 1510.9] [1573.2, 1798.6] |
| IRAS16293.2422 | [554.5, 628.5] |
| 16293E | [479.5, 636.5] |
| NGC6334i | [750.9, 771.4] [1223.9, 1241.5] [1840.1, 1857.1] |
| ngc6334i-sma12 | [479.5, 1279.8] [1575.1, 1906.7] |
| NGC6334I(N)-SMA 1-1 | [626.1, 801.9] |
| IRAS 17233-3606 | [514.8, 547] [969.9, 1001.2] [1020.4, 1039.9] [1146.8, 1181.1] [1205.8, 1226.5] |
| GCM+0.693-0.027 | [647.8, 676.3] [1140, 1168.3] |
| G0.55-0.85 | [512.8, 538.2] |
| G5.90-0.43 | [512.9, 538.2] |
| G5.90-0.44 | [512.8, 538.2] |

| | |
|---|---|
| G8.14+0.23 | [512.8, 538.2] |
| G9.62+0.19 | [512.8, 538.2] |
| G8.67-0.36 | [512.8, 538.2] |
| G10.47+0.03 | [514.8, 547] [969.9, 1001.2] [1020.4, 1039.9] [1146.8, 1181.1] [1205.8, 1226.5] |
| G10.30-0.15 | [512.8, 538.2] |
| G10.34-0.14 | [512.8, 538.2] |
| G10.32-0.16 | [512.8, 538.2] |
| G10.6-0.4 (W31 C) | [750.9, 771.4] [877.9, 956.1] [1108, 1232.3] [1232.5, 1238.2] [1645.1, 1678.9] [1840.1, 1857.1] |
| GAL 12.21-0.10 | [647.7, 676.3] [1139.8, 1168.3] |
| GAL 012.91-00.2 6 | [647.7, 676.3] [1139.8, 1168.3] |
| GAL 19.61-0.23 | [647.8, 676.3] [1139.8, 1168.3] |
| G23.44-0.18 | [512.8, 538.2] |
| G24.79+0.08 | [512.8, 538.2] |
| G25.83-0.18 | [512.8, 538.2] |
| GAL 31.41+0.31 | [647.8, 676.3] [1139.8, 1168.3] |
| GAL 034.3+00.2 | [647.7, 676.3] [1139.8, 1168.3] |
| W49N | [968.7, 986.9] |
| w51-e1/e2 | [1573.3, 1702.7] |
| AFGL2591 | [479.5, 1238.4] |
| GAL79.29+00.46 | [555.4, 636.0] |
| W75N | [647.8, 676.3] [1139.9, 1168.3] |
| DR21 | [968.7, 986.9] [1059.9, 1120.8] [1823.1, 1844.9] |
| LDN1157-B1 | [479.5, 1178.1] [1191.9, 1228.1] [1595.1, 1674.9] |
| V Cyg | [525.8, 551.9] [602.8, 629.9] [690.8, 720] [778.8, 821.1] [866.8, 898.2] [949.9, 988.1] [1053.9, 1089.1] [1141.9, 1179.2] |
| NGC7023 | [521.8, 568.4] [572.2, 580.4] [1050.6, 1121.2] |
| NGC7027 | [509.5, 545.1] [556.4, 591.2] [602.8, 635.7] [670.3, 722.2] [765.5, 800.9] [807, 850.1] [869.4, 900.1] [913.1, 944.1] [961.5, 996.4] [1022.8, 1056.5] [1227.1, 1279.9] [1459.6, 1510.9] [1577, 1605] [1633.1, 1662.9] [1713.1, 1739] |
| S CEP | [525.8, 551.9] [602.8, 629.9] [690.8, 720] [778.8, 821.1] [866.8, 898.2] [949.9, 988.1] [1053.9, 1089.1] [1141.9, 1179.2] |
| NGC7538 IRS1 | [1058.7, 1116] |
| CRL3068 | [525.8, 551.9] [602.8, 629.9] [690.8, 720] [778.8, 821.1] [821.1, 866.8] [866.8, 898.2] [949.9, 987.4] [1141.9, 1179.2] |

Table 3: Frequency range of available HIFI Spectral Surveys.

## 4.3 List of observations discarded for the generation of baseline-corrected Spectral Scan and Spectral Line Catalogues

| Source | Observation ID | Reference for supplementary material |
|---|---|---|
| Orion KL | 1342190871, 1342190872, 1342191504, 1342191592, 1342191601, 1342191649, 1342191725, 1342191727, 1342191728, 1342191755, 1342192220, 1342192329, 1342192562, 1342192563, 1342194176, 1342194178, 1342194540, 1342194732, 1342205334, 1342216387, 1342266895, 1342192673, 1342192674, 1342194733 | (5), (6), HEXOS UPDP[14] |
| SgrB2(M) | 1342191482, 1342191565, 1342191680, 1342192546, 1342204723, 1342204739, 1342205848, 1342206455, 1342206640, 1342215935, 1342216702, 1342218200, 1342243701, 1342243702, 1342251112, 1342192656, 1342206501, 1342266904 | (4), HEXOS UPDP[14] |
| SgrB2(N) | 1342204692, 1342204703, 1342204731, 1342204812, 1342204829, 1342205491, 1342205855, 1342206364, 1342206370, 1342218198, 1342266903, 1342206498, 1342206643, 1342215934, 1342216701 | (11), HEXOS UPDP[14] |
| SgrB2(S) | 1342190897, 1342191483, 1342191740, 1342190899, 1342190900, 1342191684 | HEXOS UPDP[14] |
| SgrA* | 1342230279, 1342230394, 1342239594, 1342239609, 1342243685, 1342243700, 1342243707, 1342251185, 1342230396, 1342243697, 1342243705, 1342251446, 1342252173, 1342253143, 1342253145, 1342266608 | Goicoechea et al. in prep. |
| IRC+10216 | 1342196414, 1342196423, 1342196473, 1342196475, 1342196483, 1342196514, 1342196516, 1342196518, 1342196541, 1342196543, 1342196566, 1342196574, 1342196590, 1342210102, 1342210742, 1342210754, 1342221429 | (3) and upcoming UPDP |

[14] http://archives.esac.esa.int/hsa/legacy/UPDP/HEXOS_HIFI/ReleaseNote/hexos_release_note.pdf and http://archives.esac.esa.int/hsa/legacy/UPDP/HEXOS_HIFI/Data/

| | | |
|---|---|---|
| VY CMa | 1342228611, 1342230402, 1342231504, 1342244486, 1342244491, 1342244512, 1342244631, 1342244789, 1342244791, 1342244945, 1342244960, 1342244962, 1342231467, 1342244537, 1342244610 | Quintana-Lacaci et al. in prep. and upcoming UPDP |
| OH231.8+4.2 | 1342231503, 1342231526, 1342231532, 1342244632, 1342244942, 1342244944, 1342244964, 1342245270, 1342245371 | Sanchez-Contreras et al. in prep., and upcoming UPDP |
| $\eta$ Carina | 1342181171, 1342180817, 1342180818, 1342180819, 1342181165, 1342181170, 1342232978, 1342232982, 1342235769, 1342235809, 1342235831 | (10) and upcoming UPDP |
| Mars | 1342194492, 1342194496, 1342194545, 1342194685, 1342194693, 1342194742, 1342194744, 1342194746, 1342194748, 1342194751, 1342194753, 1342235092 | Upcoming UPDP from HssO (9) |

Table 4: List of sources and observations being discarded for the generation of Spectral Line Catalogues (see text for details).

| **Source** | **Observation ID** | **Reference for supplementary material** |
|---|---|---|
| NGC6334I | 1342191481, 1342191561, 1342206085, 1342206087, 1342251671 | (14) |
| OMC-1 | 1342191503, 1342191754, 1342192215 | (8), HOP UPDP[15] |
| W49N | 1342229905 | Unknown |
| W3IC1795 | 1342190881 | Unknown |
| G327.3 | 1342238588 | Unknown |
| IRAS17233-3606 | 1342239610 | Unknown |
| G10.47 | 1342242817 | Unknown |

Table 5: List of sources and observations that were only considered for line identification, but for which no line fit parameter was provided due to imperfect baseline correction (see text for details).