Molecular fingerprints of star formation throughout the Universe : a space-based infrared study

Lahuis, F.

Chapter 2

ISO–SWS Signal Capture

Abstract

A description of the signal extraction of the ISO-SWS instrument is presented. It starts with the flux falling onto the ISO-SWS detectors up to the extracted slope signal. All known artifacts associated with the detectors and the cold and warm read out electronics are explained. The artifacts are described as well as their correction through software and calibration. The improvement of the deglitching for SWS low resolution data is described in more detail since this is an artifact which had a large impact on some science data.

Lahuis, F., Feuchtgruber, H., Golstein, H., Kester, D., Luinge, W., Shipman, R. F., & Wieprecht, E. 2003, "SWS Signal Capture: from Flux to Signal" in The Calibration Legacy of the ISO Mission (ESA SP-481), 387

Lahuis, F., Kester, D., & Shipman, R. F. 2003, "Deglitching of SWS Low Resolution Data" in The Calibration Legacy of the ISO Mission, poster contribution

2.1 Introduction

Starting with pre-flight laboratory tests (Feuchtgruber et al., 2003) during the ISO mission (Roelfsema et al., 2003; Shipman et al., 2003; Vandenbussche et al., 2003) and after the mission a continuous effort has been made to model known artifacts of the detectors and the electronics. Pre-flight laboratory tests were made to characterize individual components in as much detail as possible. Also tests under controlled circumstances have been made on the integrated system to monitor its behavior and optimize the instrument settings. The laboratory tests are at the base of the model presented in this chapter describing the integrated electronic system. The many in-flight calibration observations have helped to improve on the characterization of most instrumental artifacts. For some artifacts it has been possible to develop algorithms to correct for them. As much as possible these corrections are implemented in the SWS pipeline and or the SWS OSIA[1] (Wieprecht et al., 2003).

In the following sections the effects of the electronics are described from a data point of view. The electronics are viewed as being composed of single components or blocks of components which have a clearly identifiable impact on the observed data. Figure 2.1 shows a simplified scheme of the SWS electronics. Included are all components or component blocks which have been identified to affect the integrated signal in one way or another. All the artifacts observed in SWS data which can directly or indirectly be ascribed to the electronics are discussed. §2.3 describes how the artifacts are corrected for in the SWS calibration and data reduction of the SWS pipeline and the SWS Interactive Analysis (OSIA) (Lahuis et al., 1998; Wieprecht et al., 2003). §2.4 describes in more detail an advanced glitch detection ' for the SWS low resolution data where the standard glitch detection in many cases fails. This correction is included in the last SWS pipeline used to populate the ISO archive.

2.2 SWS detectors and detector read-out

The SWS instrument uses five types of detector materials. For the four grating arrays of 12 detectors each, InSb is used from 2.38 to 4.08 μm, Si:Ga from 4.08 to 12.0 μm, Si:As (BIBIB) from 12.0 to 29.0 μm and Ge:Be from 29.0 to 45.2 μm. For the Fabry Perot two 2-detector arrays are used with Si:Sb from 11.4 to 26.0 μm and Ge:Be from 26.0 to 44.5 μm. All detectors are used in combination with integrating pre-amplifiers with non-destructive read-out, employing heated JFET's.

2.2.1 Detectors artifacts

2.2.1.1 detector cross-talk

The SWS grating detectors exhibit a very clear crosstalk on the level of 10 to 15 percent. The main source of this is electrical crosstalk over the contacts to the detectors. The crosstalk matrix used to correct the observed data is derived from data taken from charged particle radiation tests in the laboratory and verified in orbit.

[1] OSIA is a joint development of the ISO–SWS consortium. Contributing institutes are SRON, MPE, KUL and the ESA Astrophysics Division. `http://sws.ster.kuleuven.ac.be/osia/`

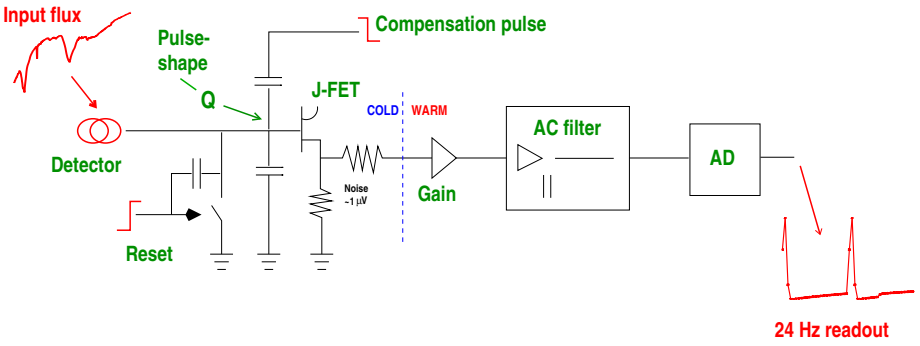

Figure 2.1 Top level scheme of the SWS Electronics showing all components known to affect the integrated signal from the detectors. The left-hand side of this scheme depicts the cold detector readout electronics located within the satellites cryostat. On the right-hand side the 'warm' electronics, the signal amplifier, filters and analog-to-digital converters.

2.2.1.2 memory effects

The SWS Si:Ga, Si:Sb and Ge:Be detectors all suffer from memory effects or transients. The InSb and Si:As detectors are both free from transient effects.

Already in the laboratory transient effects were observed. The existence of transient effects has carefully been taken into account in the design of the observing templates used for all general observations. The templates were designed such that transient effects were minimized and sufficient redundancy was present in each observation to be able to recognize transient effects in the time domain and possibly correct for them during or after the mission.

For the Si:Ga detectors a successful correction has been found in the model derived by Fouks and Schubert. This model and the application for the SWS Si:Ga detectors is discussed in detail in Kester (2003).

2.2.1.3 de-biasing

During the integration of the detector current, the capacitance at the input of the J-FET is charged and the voltage over the input to the J-FET increases. As a consequence the effective bias voltage over the detector is reduced as the integration continues and thus the responsivity of the detector changes over the reset. This effect is noticeable only at very high input signals (¿ few 10^4 Jy) and most conspicuous for the Ge:Be detectors.

For SWS this effect was studied but no correction was defined. For LWS de-biasing is also observed and an empirical correction algorithm was developed by the ISO-LWS instrument team.

2.2.1.4 particle impacts

All SWS detectors suffer from the impact of charged particles. When a particle impact occurs this is observed as a jump in the 24 Hz signal as a consequence of the additional charge deposited in the detector. These jumps are recognized and corrected for in the DSPD processing part of the pipeline (Wieprecht et al., 2000, and §2.4).

A secondary effect of the particle impact is the transient behavior after the impact. On time-scales of a few seconds transient tails are observed in particular for the Ge:Be and the FP detectors. In the pipeline no correction for this is available. In the SWS Interactive Analysis an experimental tool is available for the correction of the glitch tails (Wieprecht et al., 2000).

2.2.2 Reset of 24 Hz read-out signal

At the end of each integration ramp the system is reset by grounding the connection between the detectors and the input gate of the J-FET. This removes practically all charge from the input gate. This is directly followed by a compensation pulse, a positive or negative step depending on the bias direction. The amplitude of the compensation pulse depends on the voltage at the end of the previous integration ramp. The compensation pulse decays over the filter to the ground level centering the amplified ramps close to the ground level.

2.2.2.1 saturation

The application of the reset/compensation pulse is never perfect. A fraction of the compensation pulse is fed into the first few samples and these may go into saturation in the amplifier chain, e.g. in the operational amplifier in the high-pass filter. The reset time of this saturation is fast, less than a quarter of a second. The direct consequence however is that the first samples of the integration ramp cannot be used in the slope fit. Therefore for the default processing, for all resets the first six samples are not included in the determination of the slope.

2.2.2.2 reset-pulse aftermath

After the reset/compensation pulse has been applied a charge can be left before the input gate of the J-FET. This charge will decay over the combined capacitance of the cold electronics and is added to the charge built up by the integration of the detector current. This is seen as an additional exponentially decaying signal on the integrated 24 Hz signal. This is usually referred to as the pulse-shape effect. The pulse-shape is an additive effect and only important for low flux cases.

The pulse-shape signal is seen to change from observation to observation and often also within one observation. These can be slow or sudden changes. These changes are believed to result from variations in the deposited charge due to changes in the electronics.

2.2.3 Read-out electronics

The current from the SWS detectors is measured using integrating pre-amplifiers with heated J-FET's. The pre-amplifiers are read-out with non-destructive read-outs in a sampling of 24 Hz. The integration is done over integration resets of 1, 2 or 4 seconds depending on estimated source flux and detector. At the end of each reset a destructive read-out is applied to reset the electronic system (see §2.2.2).

The integrated signal is passed through the warm electronics where it is amplified, going through a high-pass filter, digitized and then sent down with the telemetry data.

2.2.3.1 gain amplification

SWS has three switchable gain settings, 1, 4 and 16, for all detectors. The electronic amplification factor is 225 for the grating detector and 450 for the FP detectors. The actual gain applied is the selected gain times the electronic amplification factor.

2.2.3.2 high-pass filter

A high pass filter is included to reduce the large DC voltage which would result from the amplification of the J-FET output offset. The typical RC time-constant is 2 seconds. In the laboratory individual time-constants have been derived for all 52 detectors. In orbit only for a few detectors small corrections were applied but for most detectors the time-constants derived in the laboratory are used.

2.2.3.3 A-D conversion

At the end the analogue signal is digitized. The amplified analogue signal ranges from ± 12 V. It is converted to a bitrange from 0 to 4095 corresponding to ± 10 V. So any analogue signal outside the ± 10 V range is set to 0 or 4095 and flagged as out of limit in the first stage of the processing and excluded from any later processing.

This sampling was chosen to just sample the system noise (§2.2.4). Any lower or higher sampling would have meant either a loss of sensitivity or a loss in the dynamic range of the instrument.

2.2.3.4 saturation

At very high signal levels (10^5 V/s) saturation effects come into play. First partial ramps become saturated and at even higher signal levels the complete ramp can become saturated. In the last case the ramp will be outside the measured range already after the first sample.

2.2.4 System noise

The effective root mean square (r.m.s.) noise as measured in the laboratory was typically in the range of $0.5 - 1.0 \mu \mathrm{Vs}^{-1}$. With these noise values for the highest gain setting the noise is just sampled by two bits. In orbit the observed r.m.s. noise was higher by a factor of 1 to 10 depending on the detector band.

2.3 Calibration and data reduction

For as far as possible all the artifacts described above are corrected for in the Derive-SPD stage of the pipeline processing. The corrections are applied in the reverse order in which they physically occur. Below is a description of the processing steps which together with routines for the wavelength calibration make up the Derive-SPD stage. These are the top level routine names in the SWS Interactive Analysis or in OSIA. All the routines make use of the standard pipeline modules. In between brackets the calibration files relevant for this processing step are listed. The descriptions are given on the basis of the implementation in pipeline version 10.

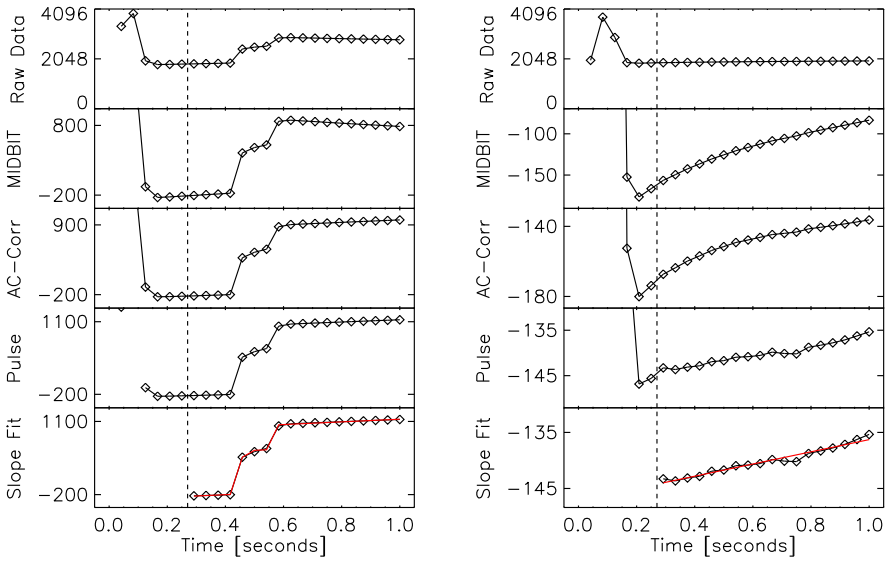

Figure 2.2 Examples of ramps at different stages of the processing for two one second resets. Shown left a reset with two glitches of which the first one is a double glitch. The right plot is a ramp without problems in which in particular the effect of the reset-pulse after effect is clearly illustrated. From top to bottom in both plots we see the raw data samples, the data after DRANGE and MIDBIT, the data after AC-correction, after pulse-shape correction and finally the glitch detection and slope-fit (including identified glitches). The fitted slope is overplotted on the data.

Figure 2.2 illustrates the effects of these corrections on two individual integration ramps.

DRANGE (CAL04)
The binary data limits are checked and data outside the limits are flagged as invalid. In the calfile the limits are listed for each detector and each gain setting.

MIDBIT (CAL02A)
The binary data is centered around the bit value corresponding to the electronic zero level of the system.

ACCOR (CAL02, CAL03)
Correction for the high pass filter. The exponential decay with a time-scale off approximately 2 seconds is approximated with a linear decay over each 1/24 second sample. At these time-scales this is an excellent approximation which allows for an efficient algorithmic implementation.

PULSE_SHAPE (CAL03, CAL05, CAL06)
A time-dependent reset-pulse aftermath correction derived from the observation data itself is applied. The applied correction is an exponential decay model with a single time-constant for the observation and a time-varying amplitude. This is subtracted from the individual ramps.

Table 2.1. Scan parameters for AOT-1 low resolution observations.

speed	reset length	dwell time	# steps	step size
1	24	3	6 (8)	4
2	48	3	14 (16)	2
3	48	3	14 (16)	1
4	48	6	7 (8)	1

Reset length and dwell time are in readout samples of $1/24$ sec and grating steps in LVDT (grating scanner position). The number of readouts in between brackets is the total number in a reset, the first number the actual numbers of steps normally used in the slope calculation.

CROSSCOR (CAL01)
The crosstalk matrix for the grating arrays is multiplied over the twelve detectors at each read-out sample.

"GLITCH" (CAL03, CAL06)
Jumps in the integrated ramp are detected. In the pipeline the glitch detection and correction are integrated in the SLOPE sub-routines.

SLOPE (CAL03, CAL05, CAL06)
As a last step slopes are fit to the fully corrected ramps. For SWS low resolution data glitched samples are first corrected on the ramps (see Sec. 2.4) after which single slope is fit. For full resolution data the slopes and glitches are fit together including step functions at the identified glitch locations.

2.4 Deglitching of SWS low resolution data

All SWS science observations were performed with one of four pre-defined observation templates depending on the purpose of the observation. These templates are called AOT's, Astronomical Observing Templates. One of these, the AOT-1, operated the instrument in a reduced or low resolution mode.

During an observation in an AOT-1 low resolution configuration the grating scanner moved during resets. At each scanner position (LVDT) the scanner remained fixed for n read-outs of 1/24 sec and then moved to a new position (see Table 2.1 for details). There exist four speeds with which AOT-1 observations were performed. The difference between these four speeds is the duration for which the scanner remained fixed (the dwell time) and the number of scanner positions the scanner moved (step size). The resulting resolution varies from approximately 15 % of the maximum grating resolution for speed 1 to approximately 50 % for speed 4 (Lorente, 1998). Table 2.1 summarizes the scan parameters for all four AOT-1 speeds.

The resulting integrated ramp can thus contain varying slopes depending on the underlying spectral structure. The most extreme changes within a reset occur when the instrument scanned over an unresolved emission line. Considering that the instru-

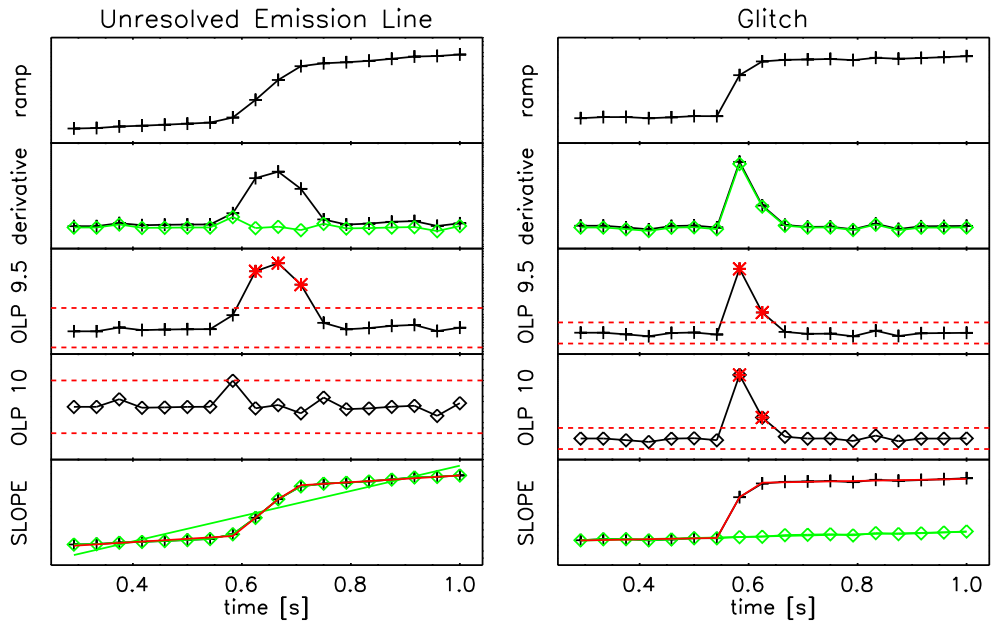

Figure 2.3 Glitch detection in low resolution data in OLP 9.5 and OLP 10 with examples of a reset scanning an unresolved line (left) and a glitched continuum reset. From top to bottom: – the input ramp – the uncorrected ramp derivative (black asterisks) and the corrected ramp derivative (green/gray diamonds)– glitch detection for OLP 9.5 on the uncorrected ramp derivative. Samples outside the limits are flagged as glitches – glitch detection for OLP 10 on the corrected ramp derivative. The emission line is not identified as containing glitches and the real glitch is properly identified. – slope fits. The OLP 9.5 fit includes the identified glitch samples (false on the left and real on the right) and follows the integrated ramp. The OLP 10 fit is done on the ramp after glitch correction (none on the left and two on the right). For the emission line reset (left) the OLG 9.5 fit clearly underestimates (by 75 %) the average signal across the reset. On the glitched continuum reset both signals agree.

mental resolution is approximately 4 to 5 LVDT's it is clear that especially in the case of speed 1 or 2 this may cause problems. When the instrument scanned over an unresolved line in a speed 1 or 2 observation only in one or possibly 2 of the reset positions the signal from the line is seen, in the rest the continuum signal is seen. When the contrast between the continuum and the line is large enough the signal from the line may be identified as a glitch and the inferred slope will be underestimated (see figure 2.3 for an example). It may be that only some resets within a line are affected, but also half of a line or a complete line can be removed in this way (see figure 2.4 for examples).

Here a solution is presented which has been introduced in the SWS Interactive Analysis environment and which will be used in OLP 10, the final archive version. The general problem is described in §2.5 and a description of the modifications are given in §2.5.2 and §2.5.3. In §2.5.4 examples and comparisons between the two deglitching methods are given with plots and by comparing line fluxes and FWHM of emission lines.

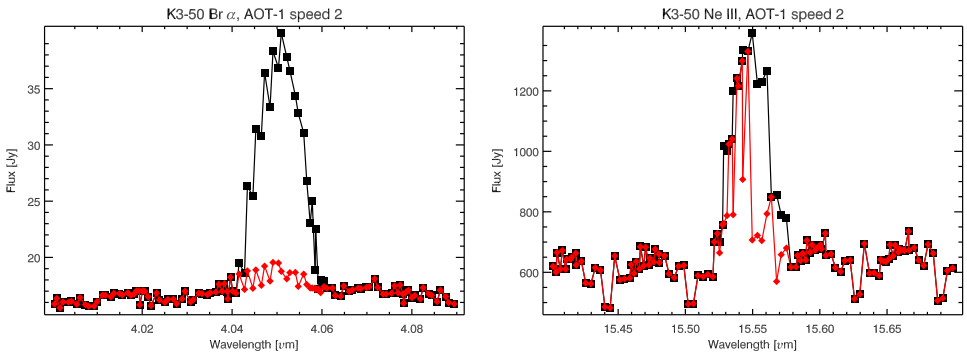

Figure 2.4 Br α and [Ne III] emission lines in an AOT-1 speed 2 of K3-50. Plotted in gray/red is the data as produced by OLP 9.5 and in black the data using the updated deglitching and slope fitting, as used in IA and OLP 10. The OLP 9.5 results clearly illustrate the problems with the old deglitching method. In the case of Br α the complete line is removed and in the case of [Ne III] half of the line disappears. In both cases the results from OLP 9.5 will be quite misleading.

2.5 Requirements

A usable deglitching algorithm for low resolution data had to fulfill the following requirements:

it should: *not* identify emission line signal as a glitch.

it should: perform identical to the old algorithm for continuum data.

it should: have *little or no* impact on the derived signal in case of error.

it should: be efficient so it can be implemented in the pipeline

To achieve this the following changes with respect to the previous standard glitch detection and correction algorithm were made:

1. identification of low resolution data in the observation. Information needed for this is contained in the SWS housekeeping data and readily accessible in OSIA or the pipeline.

2. correction of the ramp derivative (§2.5.2) used in the glitch detection for low resolution resets (note that the actual glitch detection method, as described in §2.5.1, is the same for low resolution resets and normal resets).

3. for low resolution data the slope and glitch fitting are decoupled. Instead the ramps are corrected for glitches using the glitch height as derived in the glitch detection before slope fitting (see §2.5.3).

2.5.1 SWS glitch detection

As illustrated in figure 2.3 the problem for the deglitching of low resolution data is the confusion between signal from a spectral feature and signal from a glitch. In the deglitching algorithm, glitches are identified using the integrated ramp derivative.

The standard deglitching is implemented by taking the derivative of the integrated ramp and calculating the median and the median width of its distribution. Two calibration parameters, ALPHA and MINWID, are then used to determine glitched samples. A sample is considered as glitched if the deviation of the derivative with respect to the median is larger than ALPHA times the median width and in absolute terms larger than MINWID. In the third and fourth row in figure 2.3 the limits, ALPHA times the median width and MINWID, are indicated with red and green horizontal lines. The samples identified as glitched are plotted in red. In §2.5.2 the modification to the glitch detection for low resolution data in OLP 10 is discussed.

2.5.2 Correcting the ramp derivative

To suppress false glitch detection due to scanner movement a distinction has to be made between a change in slope due to source signal and a change due to a glitch. In doing this care should be taken that the problem is not reversed and real glitches are not properly identified and corrected anymore.

To achieve this a source signal is estimated and used to correct the ramp derivative before the glitch detection. The source signal is estimated with the knowledge that source signal, contrary to the signal from a glitch, is always correlated over n samples (the dwell time, table 2.1). The estimated source signal is subtracted from the ramp derivative. Then the derivative is corrected for signal noise using a scaling based on the estimated source signal. The noise scaling is required to prevent the detection of spurious glitches as a result of signal noise.

After the correction of the ramp derivative the standard glitch detection algorithm is used to identify glitches. When a glitch is identified the glitch height is estimated from the corrected derivative without the noise scaling. In the second row of figure 2.3 the derivative (in black) and the corrected derivative (in green and in red with the signal noise correction) are plotted.

In figure 2.3 two examples illustrating the difference between the deglitching and slope fitting (see §2.5.3) of low resolution data in OLP 9.5 and OLP 10 are given. Shown are an example of a reset scanning over an unresolved emission line and an example of a glitched reset.

2.5.2.1 Problems

This correction works well in almost all cases. It only fails when there is a glitch which is seen in two samples and which in both samples has the same height or at least with a difference smaller than the noise on the integrated ramp. In such a case the glitch can be interpreted as being real signal with a negative glitch resulting in a larger signal. Any such point is however easily be removed using standard spectral filtering methods.

Since these glitches are rare and are easy to remove from the final spectrum it has been decided to not try and cater for them. In view of the use in the pipeline this is also preferred to ensure maximum robustness and highest possible efficiency.

Table 2.2. Lineflux and FWHM comparison.

speed	flux	FWHM	flux	FWHM
	OLP 9.5		OLP 10	
K3-50 Br α				
2	4.9 (-15)	325	4.9 (-14)	362
4	4.5 (-14)	1237	4.5 (-14)	1237
K3-50 [Ne III]				
2	1.1 (-13)	975	3.1 (-13)	499
4	2.6 (-13)	1452	2.6 (-13)	1452
NGC 7027 Br α				
1	3.4 (-14)	285	7.9 (-14)	465
2	7.4 (-15)	348	9.0 (-14)	355
3	8.1 (-14)	717	8.2 (-14)	699
4	9.2 (-14)	1200	9.2 (-14)	1200
NGC 7027 [Ne III]				
1	1.4 (-12)	569	1.4 (-12)	512
2	1.5 (-13)	4169	1.5 (-12)	458
3	1.4 (-12)	895	1.5 (-12)	828
4	1.1 (-12)	1400	1.1 (-12)	1391

2.5.3 Slope fitting

In OLP 9.5 the slope fitting in all cases meant an integrated fit of slope and glitches. For a reset containing glitches a fit is made of a slope plus a step function for each glitch.

This procedure works perfect for normal slopes and optimizes the number of samples used in the fit. However in the case of low resolution data an integrated fit of the slope and glitches will lead to an underestimate of the flux when signal of an unresolved line is present in the reset.

For low resolution data therefore the glitches are not fitted but corrected. Before fitting, the ramps are corrected for the glitches using the glitch-height derived in the glitch detection (see example in Fig. 2.3).

2.5.4 Deglitching results

AOT-1 observations of K3-50 and NGC 7027 are used to compare the deglitching as it is in OLP 9.5 and as it is in IA and OLP 10.

Figure 2.4 shows plots of two unresolved emission lines in K3-50 showing typical examples of line-removal problems with in red the OLP 9.5 result and in black the OLP 10 result. One line which is almost completely removed and one line which is partly removed and appears to be much narrower than the instrumental resolution.

Table 2.2 gives a summary of the derived line fluxes and FWHM of Br α and [Ne III] in K3-50 and NGC 7027 at different AOT speeds. In particular for the speed 1 and 2 observations the results are dramatically improved. For the speed 3 the results are

marginally different and for speed 4 observations there is no difference seen for these observations. The derived line fluxes with the improved AOT-1 deglitching compare well between the observations at the different speeds. The observed differences are within the nominal uncertainties of SWS in particular taking into account that e.g. no correction has been made for fringing in the case of the [Ne III] line which may have a small impact on the derived line fluxes.

During the scientific validation of OLP 10 similar exercises have been done for other lines and other sources. In all cases where lines were affected in the OLP 9.5 data they were properly reduced in OLP 10. The deglitching on continuum data was always comparable.

2.6 Conclusions

Two important improvement in the SWS data reduction after the nominal ISO mission have been the development of a pulse-shape correction and the implementation of a dedicated deglitching scheme for SWS low resolution data. The major difference of these two correction compared to the other correction such as ac-correction is that these two corrections are derived using the data itself, while the other are all corrections parametrized through calibration files.

The low resolution deglitching has been an effect which has had a direct and major impact on the science results for the affected data. Therefore a more detailed description of the implementation plus examples of this are given.

The pulse-shape correction results in a more stable zero level of the final spectrum leading to more reliable spectra for low signal sources.

For low resolution data false glitch detections on unresolved emission lines are highly suppressed and real glitches are still caught in the improved AOT-1 glitch detection algorithm. In rare cases glitches are detected where there is none, however these have little effect on the derived signal. Therefore no modifications are pursued to correct for these special cases in order not to make the algorithm overly complicated and thereby inefficient.

The pulse-shape and low resolution deglitching algorithms have been implemented efficiently to allow them to be incorporated in the SWS pipeline. This meant a significant improvement in the overall quality of the SWS data in the final ISO archive. Most importantly the archive will be free of data with suspect unresolved line features increasing the accessibility and reliability of the SWS data for the general observer and reducing the risk of misinterpretation of the spectra.

Acknowledgements

We would like to thank all the people who have been involved in ISO-SWS over the years. From the people involved in the very early development and laboratory work, the members of the SIDT in VILSPA and the support teams at the home institutes to the members of the SWS data centers and all the SWS observers in the world.