Solar Orbiter: SPICE

Data Product Description Document

[SPICEFITS]

Issue 1.8

	Name	Signature / Date
Prepared by	**Terje Fredvik and Stein Haugan**	
Reviewed by		
Approved by		
Approved by		

Institute of Theoretical Astrophysics
University of Oslo
P.O. Box 1029 Blindern
0315 Oslo
Norway

UiO : University

EXTERNAL DISTRIBUTION

Name	Organisation
spice-calibration@ias.u-psud.fr	*e-mail list*
https://spice-wiki.ias.u-psud.fr/	*SPICE team internal wiki*
https://idoc-projets.ias.u-psud.fr/redmine/	*SPICE operations redmine issue tracking tool*

CHANGE LOG

Date	Issue	Revision	Pages	Reason for change
28.06.2018	0	1	All	Document created based on SOL-SGS-OTH-0004-DPDDtemplate and [SPICELLFITS]
01.11.2018	0	9	All	Document rewritten and extended.
07.11.2018	0	91	All	Added comments based on IAS/UiO telecon.
07.11.2018	0	92		Fixed date, added READMODE in Section 4.7.1, removed an accidentally pasted random number
04.03.2019	0	93	All	– Updated sections that are copied from LLDPDD (from v1.7). Noticeable changes include a new paragraph on “Scanned Time Series” in Section 4.1, a new header example in Section 4.7.1 with some new LL01+ keywords, clarifications on the relationship between `OBS_TYPE` and the SPICE on-board Observation ID in Section 4.7.1.1.1, a new section 4.7.1.2.6 on keywords describing on-board processing steps, and a new table in Appendix B on the keywords that are derived from the value of `STUDYFLG`. See change log in LLDPDD for further details. – 3.3.3, 3.3.4.3, 4.6: the L2 data cubes will be corrected for geometric distortions, only S/C roll will be described by the `PCi_j` matrix. – 3.3: some restructuring of the text, changed temporary placeholder names of calibration routines to better match input from MPS. – 4.7.1: added L1+ keywords `UCD`, `OBS_ID`, `TIMEDER`, `TIMESYER`, `DATE-AVG`, `SOOPNAME`, `TELAPSE`, `LONPOLE`, `SPECSYS`, `VELOSYS`, `INFO_URL`, and

				Solar ephemeris data keywords. Renamed SOOP_ID to SOOPTYPE. – 4.7.1.1.1: added explanation of Solar Orbiter-wide keyword OBS_ID – 4.7.2: brief mention of L2+ keywords PRxxxxn, CRDERi, CSYERi, solar ephemeris data keywords. – 4.7.2.1: New section on keywords describing the calibration processing steps (L1 to L2) – 4.7.4: New section on keywords describing the processing steps when creating derived data products (L2 to L3) – Appendix C rewritten – Multiple minor changes, reformulations and fixing of typos
21.03.2019	0	94	All	– 4.7.1 and 4.7.3.1: corrected value of BUNIT – 3.2, 3.3.4.1, 3.3.4.3, 4.1, 4.2, 4.7.1.5, 4.7.3.2, and 4.7.4: Modified descriptions of intensity-windows to make it clear that such windows are not necessarily observed in line/background pairs. – 4.7.2.1: Added keyword PRPVERn giving the version of the processing function given by PRPROCn
03.05.2019	1	00	All	– Multiple clarifications, restructuring and moving of text, fixings of typos and minor modifications. – Updated sections that are copied from LLDPDD (v1.8). Noticeable changes include new section 4.7.1.1.7 on FILE_RAW, and rewording on intensity-windows in 4.7.1.5 – Keywords TIMEXOBT and TIMEXUTC renamed TIMAQOBT and TIMAQUTC.
02.12.2019	1	1	All. 3.3.1, 4.3.1, 4.4.1.2, 4.4.1.3.1, 4.4.2, 4.4.3.1, 4.4.4.2	– Updated sections that are copied from LLDPDD (v.1.10). Noticeable changes include new header example with new COMPRESS and COMP_RAT replacing COMPDESC. COMPPARA set to 0 for SHC and uncompressed data. New keyword SPIOBSID giving the SPICE Observation ID, no longer equal to OBS_MODE. Simplified sections by removing all examples of header differences between the example header and headers of HDUs

				with other window types and study types. Added new Section 4.4.4.2. – All: Removed references in figures, tables and text to the non-existing study/window type combination sit-and-stare intensity-windows. – Removed old Section 4.4.1.4.2 on description of processing steps in L1 files – 3.3.1: Added additional processing steps and new sub-section 3.3.3.1 describing the handling of telemetry loss. – 4.3.1: Added description of representation of space craft roll and geometrical distortions. – 4.4.1.2: Brand new L1 FITS header example. – 4.4.1.3.1: Rewritten. – 4.4.2: Added description of L1 to L2 change of window size. – 4.4.3.1: Use SI units All: Multiple minor corrections, additions, and deletions.
17.12.2019	1	2	1.3, 4, 4.4.1.2, 4.4.1.3.4, 4.4.1.4, 4.4.2	– 1.3: Added [SPICELOST] to Reference Documents. – Updated sections that are copied from LLDPDD (v.1.11). Noticeable changes include added description of HDU layout for two-exposure mode (4.1), clarification of text and fixing of typo regarding the representation of lost telemetry in binary table extensions, and new Section 4.4.1.4 explaining the value of the `BLANK` keyword and the data type of the HDU's data arrays – 4.4.1.2: New header example – 4.4.2: New default behaviour of the calibration routines is to crop the windows to preserve the number of pixels instead of padding the windows to preserve the data

01.09.2020	1	3	1.3, 4.3, 4.4.1.2, 4.4.2, 4.4.1.3.3, 4.4.1.4.1, 4.4.4.3, 4.4.5, Appendix B, Appendix C	– Multiple sections: Updated paragraphs that are copied from LLDPDD (v1.13), see changelog therein. Important changes are found in paragraphs describing the scan direction of the slit: 4.3, 4.4.1.3.3, and Appendix B.
				– 1.3: Added [DISPERSION] to Reference Documents
				– 4.4.2: reverted to the previous way of treating windows, i.e. with padding to preserve the data. Added information on changes in plate scale when going from L1 to L2
				– 4.4.1.2, 4.4.1.4.1, 4.4.4.3, 4.4.5: Renamed some of the FITS keywords used for lost telemetry book keeping to make the syntax more similar to the syntax of the SOLARNET keyword `NLOSTPIX`:
				`NCHKLOST -> NLOSTCHK`
				`NCHKFAIL -> NFAILCHK`
				`NPLNAPRX -> NAPRXPLN`
				`NPLNLOST -> NLOSTPLN`
				`PKTLOST -> LOSTPKTS`
				`BINLOST -> LOSTBINS`
				– 4.4.1.2: New FITS keyword `XSTART`. Removed `TIMSYER` and `TIMRDER` for the time being, may be reintroduced in the future
				– Appendix C: New paragraphs on IORs and Study Sets, fixed errors in Study Set table.
				– Multiple sections: future simple -> simple present…

31.05.2021	1	4	3.2, 3.3.3, 4, 4.1, 4.4.1.2, 4.4.1.3.7, 4.4.1.3.9, 4.4.1.4, 4.4.1.5, 4.4.1.6, 4.4.1.7, 4.4.2, 4.4.2.1, 4.4.3.1, 4.4.3.3, Appendix A, Appendix B, Appendix C	– 3.2: Rewritten text, updated all figures, added new figures 3-6 and 3-7. Updated all tables and added new table 3-4. – 3.3.3: Corrected names of calibration subroutines, added intensity units – 4: Updated Table 4.2. Added info on sit-and-stare studies and repeated full detector studies, highlighted these changes in the text – 4.1:`SPIOBSID-RASTERNO` is a part of the filename. – 4.4.1.2: Updated L1 header example with new keywords extracted from IORs and Study Sets, data statistics keywords, and updated keywords on data processing. – 4.4.1.3.7: New section on FITS keywords describing data processing – 4.4.1.3.9: New sectionon FITS keywords describing linear fit to scan mirror position. – 4.4.1.4-7: Updated tables – 4.4.2: Added description of adjacent and near-adjacent windows. New value for `UCD`. – 4.4.2.1: Updated `PRxxxxn` keywords – Appendix A: Removed some descriptors that should not be there, e.g. [int] for sit-and-stare studies. – Appendix B: `RASTNUM` renamed to `RASTERNO`. – 4.4.3.1: New formatting of `BUNIT` values following recommendations given in the FITS standard – 4.4.3.3: Updated `PRxxxxn` keywords – Appendix C: Added new keywords extracted from the IORs: `STP` and `WINSHIFT`, and from the Study Sets: `MISOSTUD`, `XSTART`, and `MISOWIN`.

01.10.2021	1	5	4.1, 4.4.1.2, 4.4.2, 4.4.2.1, 4.4.5, Appendix C	– Multiple sections: Updated paragraphs that are copied from LLDPDD (v1.14), see changelog therein. Most changes are found in sections describing lost telemetry, e.g. 3.3.1, 3.3.1.1 and 4.4.1.3.5, – 4.1: Updated Table 4-2 – 4.4.1.2: Updated L1 header example – 4.4.2: Updated paragraphs on adjacent and near-adjacent windows in L2 files. Adjacent windows are now concatenated. – 4.4.2.1: Added example of `PRxxxxn` keywords describing window concatenation in L2 – 4.4.5: Rewritten – Appendix C: removed `DARKSPID` (the `SPIOBSID` of the dark that has been subtracted is given by `PRPARAn` (`n` being the processing step number of the `'DARK-SUBTRACTION'` processing step). Added `SETFILE` and `SETVER`. – A few minor changes in wording
13.10.2021	1	6	4.4.1.7, 4.4.2	– 4.4.1.7: `IWINBKG` is now the HDU *number* of the HDU storing its background window, not the HDU `EXTNAME`. `IWINLINE` gives a comma separated list of HDU numbers of all HDUs having this background window defined as background. – 4.4.2: Neither Intensity-windows nor dumbbells are merged.

03.01.2023	1	7	2.3.1, 2.3.2, 3.2, 3.3.2, 3.3.3, 3.3.4, 3.3.6, 3.4, 3.4.1, 3.4.2, 4, 4.1, 4.4.1.2, 4,4,1,3,3, 4.4.1.5, 4.4.1.3.5, 4.4.1.3.6, 4.4.1.3.9, 4.4.1.3.10, 4.4.2, 4.4.2.1, 4.4.3.2, 4.4.4.3, 4.4.5, Appendix A, Appendix C	–2.3.1, 2.3.2: Added information on calibration – 3.2 and throughout the text: emphasise that most user should use L2 files, not L1. – 3.2 and throughout the text: emphasise that dumbbells and intensity-windows are not used regularly – 3.3.2: Added information on SPICE-S/C coalignment and S/C pointing variations. Added information on pixel level offset subtraction, new variable keyword `RADCAL`, and the removal of lost telemetry binary table extensions in L2. – 3.3.3: specify correct physical units for wide-slit observations. Add merging of adjacent windows as L2 processing step – 3.3.4: added warning about outdated description of Level 3 FITS files – 3.3.6, 3.4, 3.4.1, 3.4.2: Replaced placeholder text with some very basic information – 4, Figure 4-1: new figure to illustrate the Y-shift of the spectra on the two detectors – 4, Table 4.2: Added information on actual usage of the different study type/window type combination. 4.4.1.2: New L1 header example – 4.4.1.3.3 and throughout the text: added information on the FITS keyword consequences of merging of adjacent windows in L2. Removed incorrect description of reversal of wide-slit images – 4.4.1.3.5: Added description of all keywords describing the telemetry – 4.4.1.3.6: New section coordinate distortions –4.4.1.3.9: New description of how to determine the study type – 4, 4.4.1, 4.4.1.5, etc: Added information on the lack of separate

				dumbbell HDUs due to the SW/LW detector y-offset 4.4.2.1: Updated processing steps – 4.4.3.2: IWINBKG/IWINLINE are HDU numbers, not the EXTNAME – 4.4.1.3.6: New section on coordinate distortions due to S/C pointing instability – 4.4.1.3.8: Added information on pixel level offset subtraction – 4.4.1.3.10: rewritten to explain that the study type is now determined from the IORs and study definition files instead of mirror positions. 4.4.2: Updated description on L2 FITS keywords 4.4.4.3: New section describing the variable keyword RADCAL – 4.4.5: Added information on telemetry completeness keywords – Appendix A: added information on actual data product type usage – Appendix C: Added FITS keyword SETVER. Modified description of EXTNAME in the case of merged windows

02.11.2023	1	8	1.3, 3, 3.2, 4, 4.4.1.2, 4.4.1.3.6, 4.4.2.1, 4.4.3, 4.4.3.1	– All: added information on the infrequent usage of intensity-windows and dumbbells. Removed some detailed descriptions of these types of observations. Simplified and shorted text a few places. Removed green colouring of text that indicated sections identical to sections in [SPICELLFITS] – 1.3: added references – 3: added information on IDL analysis tools – 3.2: removed detailed descriptions and illustrations of data flows of seldomly obtained observations. Updated Figure **3-1** – 4,: Updated Table 4-2 – 4.4.1.2: new L1 header example – 4.4.1.3.6: added information on new SolarSoft support of coordinate distortions. Updated Figure 4-2 including caption. – 4.4.2.1: Updated processing steps describing calibration, including correction for burn-in and time dependent detector response – 4.4.3: Rewritten from scratch, now describes the L3 R files being produced by the pipeline – 4.4.3.1: Rewritten from scratch

TABLE OF CONTENTS

1. INTRODUCTION

1.1 Purpose and Scope

This Data Product Definition Document describes the format and content of the SPICE Science data. It includes descriptions of the data products and associated metadata, including the data format, content, and generation pipeline. These products will be stored and distributed from the Solar Orbiter Science Archive (SOAR) of the SOC.

The specifications described in this document apply to all SPICE Science products submitted to ESA's Solar Orbiter SOC for further archival and exploitation. The specifications described in this document apply to all SPICE data products generated by the Science Data Pipeline running at the SPICE premises in Oslo. It does not address the Low Latency data (see [LLData]) delived by the Low Latency Pipeline run at SOC, as these data products are described in [SPICELLFITS]. However, the LL01 data products produced by the Low Latency Pipeline are identical to the Level 0 data products created by the Science Data Pipeline, save for some minor FITS keyword differences.

1.2 Applicable Documents

APPL. DOC. SHORT	APPLICABLE DOCUMENT TITLE	DOCUMENT ID	ISSUE
[DPICD]	*Data Producer to archive ICD*	*SOL-SGS-ICD-0002 ?*	*Will be rewritten from scratch*

Table 1-1: Applicable Documents

1.3 Reference Documents

REF. DOC. SHORT	REFERENCE DOCUMENT TITLE	DOCUMENT ID	ISSUE
	Template for SoLO Data Product Description Document	SOL-SGS-OTH-0004-DPDDtemplate	
[ACRONYMS]	SPICE Acronym List	SPICE-RAL-LI-0001	1.0
[LLFITSICD]	Solar Orbiter Interface Control Document for Low Latency Data FITS Files	SOL-SGS-ICD-0005	1.4
[FITSpaper]	Definition of the Flexible Image Transport System (https://fits.gsfc.nasa.gov/standard40/fits_standard40aa-le.pdf)		4.0
[METADATA]	Metadata Definition for Solar Orbiter Science Data	SOL-SGS-TN-0009	2.5
[IOR ICD]	Solar Orbiter Instrument Operations Request ICD	SOL-SGS-ICD-0003	1.1
[SEGU]	SOC Engineering Guidelines for External Users	SOL-SGS-TN-0006	1.2
[SOAR]	Solar Orbiter Archive Plan	SOL-SGS-PL-0009	2.0
[MAN]	SPICE Instrument User Manual	SPICE-RAL-MAN-0001	13.0
[DATAICD]	SPICE Data Interface Control Document	SPICE-RAL-ICD-5003	10.0
[RAW]	SPICE FM Raw Science Data Handling Scheme	SPICE-RAL-TN-5202	1.0
[DECOMP]	SPICE Science Data Decompression Recipe	SwRI EM 17489-017	0
[SUNSPICE]	The SunSPICE Ephemeris Packeage for Solar Missions, William Thompson, 2018 (https://hesperia.gsfc.nasa.gov/ssw/packages/sunspice/doc/sunspice.pdf)		
[DISTORTIONS]	Simulated Image Distortions in SPICE Data, William Thompson, 2019 (https://spice-wiki.ias.u-psud.fr/lib/exe/fetch.php/documents:internal:simulate.pdf)		
[DISPERSION]	SPICE spectral dispersion function and detector geometry, Thompson and Young, 2020 (unpublished, under development)		
[COALIGN]	SPICE pointing offset, Pelouze, 2022 (unpublished)		
[LLdata]	Solar Orbiter Low-Latency Data: Concept and Implementation	SOL-SGS-TN-0003	1.2
[SPICELLFITS]	SPICE Low Latency Data Product Description Document	SPICE-UIO-DPDD-0001	1.14
[SPICELOST]	Solar Orbiter SPICE Reconstructing Data with Lost Telemetry Packets (https://spice-wiki.ias.u-psud.fr/lib/exe/fetch.php/documents:internal:spice-uio-dpdd-0003-0.2-reconstructing_data_with_missing_telemetry_-_comressed_images_hd.docx)	SPICE-UIO-DPDD-0003	0.2
[S-META]	SOLARNET Metadata Recommendations for Solar Observations (http://sdc.uio.no/open/solarnet/ , https://arxiv.org/abs/2011.12139,)		2.1-live
[FITSCOORD]	Representation of spectral coordinates in FITS, Greisen & al., 2006, A&A, 446, 747-771 (https://www.aanda.org/articles/aa/pdf/2006/05/aa3818-05.pdf)		

[WCSDISTORTIONS]	Representations of distortions in FITS world coordinate systems, Calabretta et al., 2004, draft (https://fits.gsfc.nasa.gov/wcs/dcs_20040422.pdf)		
[SSTRED]	SSTRED: Data- and metadata-processing pipeline for CHROMIS and CRISP, Löfdahl et al., 2021, A&A 653, A68 (https://www.aanda.org/articles/aa/pdf/2021/09/aa41326-21.pdf)		
[CFIT]	The Component Fitting System (CFIT) in IDL	CDS Software Note No. 47	2
[DATAMAN]	SPICE Data analysis user's manual (https://spice-wiki.ias.u-psud.fr/doku.php/data:data_analysis_manual)		
[IDLANA]	SPICE IDL Quicklook and Data Analysis Software (https://github.com/ITA-Solar/solo-spice-ql/wiki/)		

Table 1-2: Reference Documents

1.4 Abbreviations and Definitions

FITS	Flexible Image Transport System
JPEG	Joint Photographic Experts Group; a lossy image format
MPEG	Moving Picture Experts Group; a lossy format for encoding and compressing video images.
LL01	Low Latency Level 1
LL02	Low Latency Level 2
L0	Level 0
L1	Level 1
L2	Level 2
L3	Level 3
SHC	Spectral Hybrid Compression
SOAR	Solar Orbiter Archive
SOOP	Solar Orbiter Observing Program
WCS	World Coordinate System

A complete list of all acronyms used in the SPICE project can be found in [ACRONYMS].

2. SPICE INSTRUMENT DESCRIPTION

SPICE is a high-resolution imaging spectrometer operating at ultraviolet wavelengths. It will address the key science goals of the Solar Orbiter mission, by providing quantitative knowledge of the physical state and composition of the plasmas in the solar atmosphere, in particular investigating the source regions of outflows and ejection processes which link the solar surface and corona to the heliosphere. [MAN]

2.1 Science Objectives

See Table 2 in Section 2.1 of [MAN].

2.2 Operational Modes

Data from the SPICE instrument is read out in wavelength regions, see [MAN] Section 2.2.2.2. Depending on the operational mode (study definition), the data from each of these readout windows may result in 1, 2 or 3 data arrays. It is also possible to make a full-frame read-out, resulting in one data array per detector array.

2.3 Calibration

2.3.1 On-ground Calibration

The observations are calibrated on-ground by the Science Data Pipeline running in Oslo. Calibrated files are of Level 2, see Section 3.3.3. Most users should use Level 2 files in their data analysis.

2.3.2 In-flight Calibration

Dark current subtraction is the only form of calibration that may be applied in-flight. If not applied on-board, the dark current is subtracted on-ground by the Science Data Pipeline when creating Level 2 files.

3. DATA GENERATION PROCESS AND ANALYSIS PROCESS

The SPICE Science Data products are produced by the Science Data Pipeline running in Oslo. The data generation and analysis process are described in this section.

The science data products produced by the Science Data Pipeline are immediately available as to the SPICE consortium through the Oslo SPICE archive, `http://astro-sdc-db.uio.no/vol/spice/fits/`. The data are made available to the broad solar community through the Solar Orbiter archive following the policies described in the Archiving Plan [SOAR].

For practical hints and tips regarding data analysis, please see the SPICE Data analysis user's manual [DATAMAN]. IDL users may use the tools developed at UiO for quicklook and analysis purposes, see the team's wiki pages for details [IDLANA].

3.1 Scientific Measurements

Top-level description of the data acquired by the instrument (RAL).

3.2 Data Flow Overview

The SPICE team in Oslo retrieves the telemetry data from IAS, passing it as input to the Science Data Pipeline running in Oslo. The Science Data Pipeline produces science data files of Level 1, 2 and 3[1].

[1] In June 2023 the pipeline started creating simplified L3 files containing gaussian fit parameters.

Most users should use Level 2 files for their data analysis. Level 1 files are meant for advanced users. L3 files should only be used for quicklook purposes.

Level 1 and Level 2 files are in the FITS format (Sections 3.3.2 and 3.3.3). Level 3 files will come in three flavours: Regular Level 3 FITS (L3 R, Section 3.3.4.1), Concatenated Level 3 FITS (L3 C, Section 3.3.4.2), and Quicklook Level 3 JPEG/MPEG (L3 QL, Section 3.3.4.3). For the time being only L3 R files are being produced.

The two SPICE detectors are read in wavelength regions, or windows. Each window is of a specific *window type*, see Table 4-1. A study is of a given *study type* (Table 4-2) and may consist of multiple window types. The study may be repeated. The dataflow differs for the multiple combinations of window types and study types, and for repeated vs non-repeated studies. See Section 4 for further information regarding window types, study types, and data products.

The most frequently obtained SPICE science observation is narrow-slit spectral rasters (see also Table 4-2). The dataflow inside the Science Data Pipeline for such an observation is illustrated in Figure **3-1**. The file formats resulting in narrow-slit spectral rasters are summarised in Table 3-1.

Wide-slit rasters are also obtained on a regular basis. Other observations are either mostly used for calibration purposes (full-detector single exposure observations), are not that frequently used (sit-and-stare observations), or have hardly been used at all (e.g. Intensity-windows and dumbbells). We therefore omit a thorough description of these types of data products in this document.

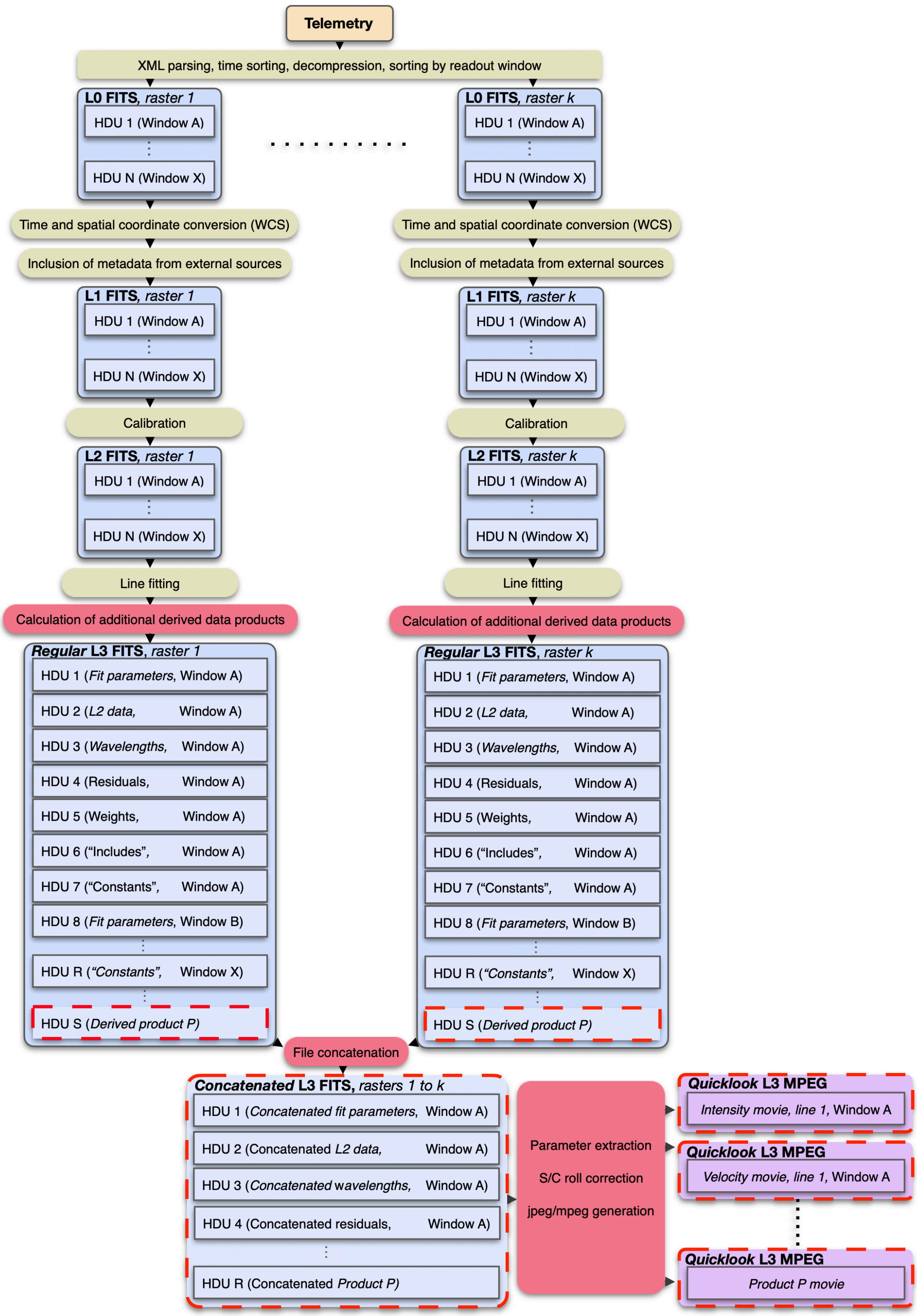

Figure 3-1 Pipeline dataflow: narrow-slit spectral-profile. Red colour signals planned processing steps and HDUs/file types that are not yet part of the pipeline. Concatenated L3 files will only be made for repeated rasters.

The table below lists the formats of the files that the Science Data Pipeline produces:

	Narrow-slit spectral-profile			
	Repeated raster	Single raster	Single sit-and-stare	Repeated sit-and-stare
L1	FITS	FITS	FITS	FITS
L2	FITS	FITS	FITS	FITS
L3 R	FITS	FITS	FITS	FITS
L3 C	FITS	–	–	–
L3 QL	MPEG	JPEG	JPEG	JPEG

Table 3-1 Narrow-slit spectral-profile file formats. Only repeated studies result in data arrays in L3 C FITS files.

3.3 Data Generation

The Science Data Pipeline takes XML telemetry as input and produces files of all file levels described in Sections 3.3.1 to 3.3.4 sequentially. The pipeline is written in IDL, with the use of java objects and C executables.

Most users should use Level 2 files for their data analysis. Level 1 files are meant for advanced users who either require full control of the calibration process or need access to the extended information on lost telemetry packets that these files contain.

3.3.1 Level 0 – Raw Data

Level 0 FITS files (L0) are used as temporary files within the Science Data Pipeline and will normally not be made available to the scientific community[2]. L0 files contain uncalibrated data expressed in engineering units (counts).

The processing steps performed by the Science Data Pipeline to create an L0 file:

1. The Science Data Pipeline regularly checks its input directory for processing requests.
2. If a request is found, the XML contents of the request is read, and the telemetry data is extracted.
3. The telemetry packets are sorted by time of packet creation.
4. A processing request may contain telemetry from multiple observations. The following steps are repeated for each observation found in the sorted telemetry stream:
5. Each telemetry packet is inspected, and the data and metadata of a packet is extracted according to Sections 4.2.6.1 to 4.2.6.3 of [DATAICD]. Data and metadata in telemetry packets are accumulated into collections called raster segments. The data array of a raster segment is a 1-dimensional 8-bit array.

[2] L0 files were made available to the SPICE consortium during the commissioning phase. The Science Data Pipeline up to (and including) the generation of L0 files is identical to the Low Latency Pipeline running at SOC. L0 files are indistinguishable from Low Latency Level 01 FITS files (LL01), with the exception of the values of a few descriptive FITS keywords (e.g. `FILENAME`, `LEVEL`, `CREATOR`, etc.).

6. When a raster segment is complete, or there are no packets left in the telemetry that belongs to the segment, it is decompressed according to [RAW] and [DECOMP]. Special care must be taken when telemetry packets are lost, see Section 3.3.1.1.
7. The decompressed raster segment data array is translated into a 3-dimensional 16-bit data cube with dimensions `(exposure,y,dispersion)`, according to [RAW]. If the telemetry was compressed using Spectral Hybrid Compression (SHC), this translation also involves a inverse Fourier transform as described in [DECOMP].
8. The decompressed raster segment data cubes are assembled into window data cubes. Window data cubes may consist of multiple raster segments, but each data cube only contains data stemming from a single readout window.
9. When all window data cubes of the observation are complete, or there are no more packets left in the telemetry, the data from each window data cube is reformatted into `(x,y,dispersion,time)` data cubes, with a singular `x` dimension for sit-and-stare observations and singular `time` dimension for raster scans.
10. A temporary L0 FITS file is created, with the data for each window stored in separate Header/Data Units (HDUs). The data from the first readout window are stored in the primary HDU, data from additional readout windows are stored in image extension HDUs.
11. If the telemetry contains dumbbell data that are downlinked in separate segments, i.e. not part of a spectral window, the data cubes are stored in the two last image extension HDUs (or in the last image extension HDU if only one dumbbell was recorded and/or received). Note that dumbbells are very seldomly downlinked separately, see Section 4 for additional information on dumbbell data.
12. Each L0 HDU that contains data from readout windows also contains all relevant metadata present in the telemetry packets, see Appendix B. Such HDUs also contain metadata that have been derived from the telemetry, in some cases using formulas and conversion tables stored in the virtual machine.
13. In the case of multi-exposure observations, the L0 file also contains a binary header extension storing the individual values of keywords that vary with exposure number, see Section 4.4.4.1.
14. In the case of multi-segment observations, the binary header extension mentioned in bullet 13 will also store the individual values of keywords that vary with segment number, see Section 4.4.4.2.
15. If telemetry packets are lost, the lost packet indices are stored in a binary table extension. If SHC compression was applied to the data, this binary table extension also stores the lost FFT Bin indices, see Section 4.4.4.4.
16. If telemetry packets are lost and JPEG or SHC compression was applied to the data, one or more image planes will either have approximated values or be completely lost (i.e. be set to the `BLANK` value). The coordinates of these image planes (or image plane ranges in the case of multi-segment observations) are stored in a binary table extension, see Section 4.4.5.[3]
17. GOTO 5

3.3.1.1 Reconstructing data with lost telemetry packets

During the first years of the mission, lost telemetry has not been a pronounced problem. We might not be that lucky in the future, and we have therefore implemented mechanisms for storing

[3] For technical reasons the flagging of lost or approximated image plane ranges currently does not work if all packets of a segment are missing.

information on lost telemetry in the L1 FITS files. L2 files contain a simplified set of FITS keywords describing the completeness of the telemetry.

The implications of lost telemetry are discussed in [SPICELOST]. Missing telemetry packets affect the reconstructed data cube depending on the type of compression applied to the data:

- **No compression:** pixel columns in the dispersion-Y plane are lost, i.e., they have the `BLANK` value.
- **JPEG compression:** entire X-Y or dispersion-Y image planes of a single segment[4] are affected. *All* pixels in affected planes have *either*:
  - the `BLANK` value if the missing packets contained JPEG header information, or
  - approximated values if the missing packets did *not* contain JPEG header information
- **SHC compression:** *all* pixels stemming from the segment that the lost packets belonged to have approximated values. The degree of approximation depends on which FFT Bins are lost, see discussion in [SPICELOST].

In Level 0 and Level 1 files, detailed information concerning lost telemetry is stored in binary table extensions and in FITS header keywords of the observational HDUs, see Sections 4.4.4.4.and 4.4.5. These binary table extensions are not present in Level 2 files.

3.3.2 Level 1 – Engineering Data (uncalibrated)

Level 1 FITS files (L1) contain uncalibrated data expressed in engineering units (counts).

Note that most users should use Level 2 files for their data analysis. Level 1 files are meant for advanced users who either require full control of the calibration process or need access to the extended information on lost telemetry packets that these files contain.

The FITS headers include additional keywords giving times converted from on-board time (OBT) to UTC, and the spatial coordinates are converted from being given relative to the spacecraft boresight to coordinates relative to the Solar disc. In this respect L1 files are identical to the Low Latency Level 02 files generated by SOC. However, L1 files also include additional metadata gathered from the Study Set files and the IORs. These additional metadata describe all available information about:

a) the study definition (e.g. name of the study, name of the readout windows, name of the author of the study, etc.), and
b) this particular instance of the study (e.g. the Solar Orbiter-wide ID of the observation, name of the SOOP, etc)

The L1 FITS files also include all metadata needed by the calibration routines in order to convert the file from L1 to L2, including Solar ephemeris data. L1, L2 and L3 FITS files will be stored in the SOAR.

The processing steps performed by the Science Data Pipeline to create an L1 file:

[4] Note that an observation may consist of multiple segments. Telemetry loss may in such cases affect only parts of the data cube image planes, see Section 4.4.5.

1. An L0 file is given as input to the Level 1 FITS file generator.
2. The on-board time is converted to UTC using the SunSPICE IDL implementation of the SPICE[5] toolkit, see [SUNSPICE]. Additional FITS standard keywords `DATE-BEG`, `DATE-END` and `DATEREF` given in UTC are added to the L1 FITS headers.
3. Solar ephemeris data keywords are added using the SPICE toolkit.
4. IAS produces XML files with metadata that describe the specific studies and the specific instances of each study, i.e. the study definitions and the IORs. Oslo hosts a local database storing all this metadata. Based on this database and the FITS keywords in the input L0 file, additional metadata are added to the L1 file.
5. The keywords describing the pointing and FOV are converted from being relative to the spacecraft boresight to being relative to the solar disc (helioprojective) using SPICE toolkit routines and the XML metadata files obtained from IAS. The offset between the SPICE and spacecraft boresight is taken into account, based on the equations given in [COALIGN].
6. The spacecraft roll and pointing are calculated using the SPICE toolkit. The `PCi_j` transformation matrix explained in Section 4.3 is updated in order to describe the resulting rotation of the field-of-view. The distortions of the Solar X and Solar Y coordinates due to any spacecraft pointing variations during the observation are stored in separate image extensions, see Section 4.4.1.3.6.
7. The L1 file inherits any binary table extensions present in the L0 file. An additional binary table column storing the acquisition time converted to UTC is added to the binary table extension. See Section 4.4.4 and 4.4.5 for details.
8. If a pixel level offset was added on-board, this offset is subtracted from the data array.

3.3.3 Level 2 – Science Data (calibrated)

Level 2 FITS files (L2) are calibrated L1 files that are ready for scientific analysis:

– The data are corrected for dark current (only if a dark frame hasn't already been subtracted on-board).
– The data are corrected for flat-field.
– The data are corrected for burn-in. Note that only sections of the detectors where strong lines have reduced the sensitivity significantly are corrected for burn-in.
– The data are corrected for spatial and spectral distortions. Corrections for slit tilts, spectral slant, detector misalignments, non-uniform dispersion, and other geometric distortions are applied to the data cubes by interpolation onto a linear grid.
– The data are calibrated to physical units: `W/m2/sr/nm` for narrow-slit observations and `W/m2/sr` for wide-slit observations. This step also includes correcting the data for the time-dependent response of the detectors (i.e. the reduction of detector sensitivity with time).
– The data arrays of adjacent windows are merged

The rotation of the field-of-view due to spacecraft roll is not corrected for, instead the rotation of the field-of-view is described by the `PCi_j` transformation matrix.

The processing steps performed by the Science Data Pipeline to create an L2 file:

1. An L1 file is given as input to L2 FITS file generator, and it is sent to the calibration wrapper routine `spice_prep.pro`
2. `spice_prep.pro` calls the following calibration subroutines:
   o `spice_prep_dark_offset_correction.pro` (dark current subtraction)

[5] NASA's Observation Geometry System for Space Science Missions.

- `spice_prep_flat_field_correction.pro` (flatfield correction)
- `spice_prep_burnin_correction.pro` (burn-in correction of selected lines)
- `spice_prep_distortion_correction` (correct for geometrical distortions)
- `spice_radiometric_calibration.pro` (ADU to physical units. Correct for the time-dependent decrease of detector sensitivity)

3. Image planes that contain approximated values in the L1 file are set to `NaN` in the L2 file, see Sections 4.4.4.4 and 4.4.5.
4. The data arrays of adjacent windows are merged, see Section 4.4.2.
5. The output from `spice_prep.pro` is a calibrated L2 FITS file. The data type of a L2 data array is floating point.
6. The L2 file inherits any variable keywords binary table extensions present in the L1 file, with an additional variable keyword `RADCAL`. Any binary table extensions storing information on lost telemetry are removed.

3.3.4 Level 3 – Higher Level Data

NOTE: As of November 2023 the pipeline produces simplified Level 3 R files containing Gaussian fit parameters. The format of future Level 3 files containing secondary derived parameters is yet to be finalised.

Level 3 files (L3) contain higher-level data products. L3 can be either FITS files (Sections 3.3.4.1 and 3.3.4.2), or JPEG images or MPEG movies (Section 3.3.4.3).

3.3.4.1 Regular Level 3 FITS (L3 R)

Regular Level 3 FITS files (L3 R) contain derived data products like e.g. line peak intensity, line shift and line width obtained by fitting Gaussians to the line profiles. L3 R FITS files are created from narrow-slit raster and sit-and-stare observations (i.e. windowed spectral-profile HDUs, see Table 4-1 and Table 4-2).

Note that line fitting is sensitive to the initial values of the Gaussian parameters and in some cases the derived data products produced by the pipeline may therefore be less trustworthy than if the line fitting had been performed with human supervision. Also, the algorithm that automatically detects emission lines may fail, and it may report false positives or miss some weaker emission lines. For scientific purposes, we therefore strongly encourage users to perform the line fitting manually, starting with L2 files.

The processing steps performed by the Science Data Pipeline to create an L3 R file:

1. An L2 file is given as input to the L3 FITS file generator, which calls the `::create_l3_file` method of the `spice_data` IDL object (see [IDLANA]).
2. This method selects HDUs with data stemming from narrow-slit spectral-profile windows and attempts to automatically detect and identify up to 10 emission lines in each readout window.
3. An initial guess of the amplitude, position, and width of each line is given as input to the line fitting routine `cfit.pro` (see [CFIT]), which returns the fitted gaussian line parameters peak intensity, line shift and line width.
4. Each observational HDU in the input L2 file results in multiple HDUs in the L3 files, containing e.g. the Gaussian fit parameters, L2 data, residuals, etc. See Section 4.4.3. for details.

5. *To be implemented:* based on the fitted gaussian line parameters, secondary derived parameters will be estimated.

3.3.4.2 Concatenated Level 3 FITS (L3 C)

To be implemented. For the time being only L3 R FITS files are produced, see Section 3.3.4.1.

Concatenated Level 3 FITS files (L3 C) will consist of time series of derived data products from multiple Regular L3 files. Only repeated, narrow-slit spectral-profile windows will result in HDUs in L3 C files.

The processing steps performed by the Science Data Pipeline to create an L3 C file:

1. All L3 R files from a repeated study are identified
2. The data cubes of these files are concatenated in the time dimension.
3. A Concatenated L3 FITS file is created, each HDU containing a concatenated data cube.
4. The coordinates of an L3 C data cube may need to be specified using table lookup (see Section 6 of [FITSCOORD]), and some of the FITS keywords may need to be stored using the variable-keyword mechanism (Appendix I of [S-META]).

3.3.4.3 Quicklook Level 3 MPEG and JPEG (L3 QL)

To be implemented. For the time being only L3 R FITS files are produced, see Section 3.3.4.1.

Quicklook Level 3 files (L3 QL) will be either MPEG movies or JPEG images. L3 QL files are not suitable for scientific analysis, as they are not in an appropriate scientific file format, they are highly compressed, and they do not contain the necessary metadata required for a thorough analysis. They may, however, provide useful context information, and they can be used as a tool to find interesting observations that should be investigated further.

L3 QL files contain either an image or movie of a single derived data product. An L3 QL MPEG file is a movie of a derived data product based on repeated raster observations.
An L3 QL JPEG file is an image of a derived data product based on a single raster observation.

The processing steps performed by the Science Data Pipeline to create an L3 QL file:

1. An L2, L3 R or L3 C file is given as input to the L3 quicklook jpeg/mpeg file generator.
2. The spacecraft roll is adjusted for before the data cube is written to a movie/image file (in L2, L3 R and L3 C files rotation of the FOV is described by the `Pci_j` transformation matrix).
3. Movies are created using the `IDLffVideoWrite` object using the mpeg4 codec (TBD)
4. Images are created using the `write_jpeg.pro` procedure with `QUALITY` set to 75 (TBD).

3.3.5 CAL – Calibration data

The calibration routines and the necessary calibration data will be made available in SolarSoft (SSW) and in the SolarSoft Data Base (SSWDB).

3.3.6 HK – Housekeeping data

The Oslo pipeline does not process housekeeping telemetry, only science telemetry (both Low Latency Data and Science Data).

3.4 Validation

The FITS files produced by the Oslo pipeline are continuously validated by the Oslo team. Multiple validation procedures are run automatically every time new FITS files are created. The results of these tests are automatically sent to the Oslo team by e-mail. The team is also notified by e-mail and/or SMS if the pipeline encounters errors or crashes.

In earlier stages of the pipeline development the data products were compared to the output from RAL's *cube builder*.

3.4.1 Instrument Team Validation

The FITS files produced by the Oslo pipeline are automatically made available to the SPICE consortium for validation and analysis. Consortium members of IAS, MPS, RAL and Goddard are making invaluable contributions to the pipeline development and data product validation by reporting bugs, providing support, and participating in discussions.

3.4.2 SOC Validation

SOC validates the SPICE FITS files upon ingestion into the Solar Orbiter Archive.

4. DATA PRODUCT DESCRIPTIONS

SPICE data products stored in FITS files are formatted in accordance with the rules outlined in [METADATA]. This section provides details on the formats, metadata, and filenames for each of the products included in the SPICE data.

A SPICE FITS file contains a primary Header/Data Unit (HDU), and it may contain one or more image extensions. All primary and image extension HDUs containing observational data have a full header as described in [METADATA] – i.e. there is no distinction between primary HDUs and image extensions other than those required by [FITSpaper]. Each HDU is regarded as a data product.

In addition, Level 1 and Level 2 files may contain binary table extensions storing auxiliary data that have individual values for each exposure of the observation. Level 1 files may also contain binary table extensions describing lost telemetry, see Sections 4.4.4 and 4.4.5.

Finally, Level 1 and Level 2 files storing multi exposure observations always have two additional image extensions describing coordinate distortions due to spacecraft pointing instabilities, see Section 4.4.1.3.6.

The SPICE detector is read in wavelength regions, or windows, of a specified wavelength range. The data array of an observational HDU stem from one of the following ***6 window types***, described in the `WIN_TYPE` FITS keyword:

	Window type description	`WIN_TYPE`
1	Narrow-slit spectral-profile (2”, 4”, and 6” slits)	`'Narrow-slit Spectral'`
2	Dumbbell stack	`'Dumbbell (lower)'` or `'Dumbbell (upper)'`
3	Wide-slit (30” slit)	`'Wide-slit'`
4	Intensity-window (2”, 4”, and 6” slits)	`'Intensity-window'`
5	Full detector read-out (2”, 4”, and 6” slits)	`'Full Detector Narrow-slit'`
6	Full detector read-out (30” slit)	`'Full Detector Wide-slit'`

Table 4-1: All available window types and corresponding values of `WIN_TYPE`. Note that the vast majority of SPICE observations are of window type 1 and 5. Window type 2 and 4 have hardly ever been used, see Table 4-2.

Note that the spectra on the two detectors have a significant relative shift in the detector Y direction. Due to hardware limitations all type 1 windows within a study must have the same detector Y coordinates. Therefore, the readout windows must be significantly taller than the height of the slit to ensure that all windows on both detectors cover the full slit height, see Figure 4-1. This has two major consequences:

1) The upper and/or lower part of most windows of type 1 will contain dumbbell data. These pixels must be masked before any line fitting can be done to the spectral data of the window. The WCS keywords do not apply to the dumbbell pixels.
2) A multi-exposure Narrow-slit observation will normally only contain such extended windows of type 1, and no separate dumbbell windows.

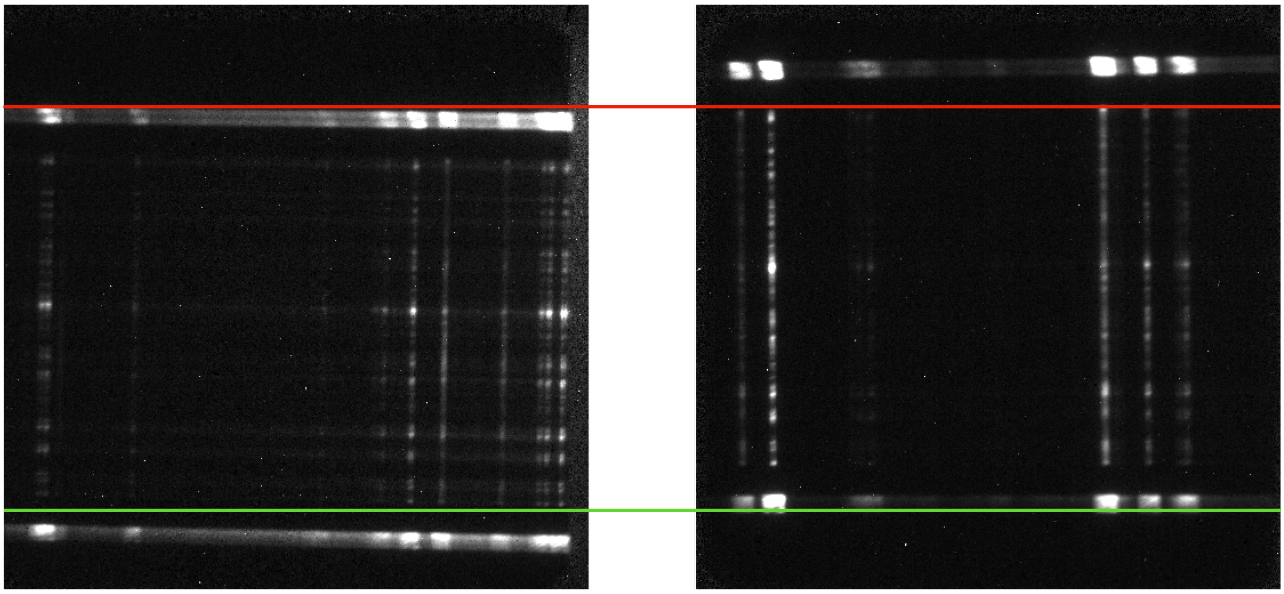

Figure 4-1: The spectra projected on the Short Wavelength detector (SW, left) and the Long Wavelength detector (LW, right) are shifted in the detector Y direction. The lower edge of the slit on the SW detector (green) is at a lower Y pixel value than the dumbbell region on the LW detector. The upper edge of the slit on the LW detector (red) is above the SW dumbbell region. In order to catch the full slit height on both detectors the readout windows must be covering the Y-range between the green and the red line, i.e., a spectral window will also contain dumbbell information. The figure shows SPIOBSID 150995892 corrected for flat field and dark current.

The window types may be of one the following ***3 study types***, described in the `STUDYTYP` FITS keyword:

	Study type	Possible window types	Actual usage, all files	Actual usage, science only	STUDYTYP	X	t
1	Sit-and-stare	1 2 3	1.1% - 0.1%	3.5% - <0.1%	'Sit-and-stare'	1	>1
2	Raster	1 2 3 4	26% <0.1% 3.5% <0.1%	**91%** - 5.6% -	'Raster '	>1	1
3	Single exposure	1 3 5 6	3.5% 0.3% **60%** 6.6%	- - - -	'Single Exposure'	1	1

Table 4-2: All available combinations of study types and window types, and their actual usage in percentage of the number of files obtained after commissioning up to November 2023. Note that for science studies most observations are narrow-slit spectral rasters. However, be aware that even if a file is not labelled as *Science* (i.e. it is labelled *Calibration*) it may contain scientifically interesting data. As an example, even if none of the single exposure full detector observations are labelled as *Science*, only 2% of these files are darks.

The very seldomly used window type 4 (Intensity-window) may not be of study type 1 (Sit-and-stare). Window types 5 and 6 (full-detector read-out) may only be of study type 3 (single exposure). However, it is possible to repeat full detector observations, with or without moving the scan mirror, thereby enabling the construction of L3 rasters and sit-and-stare observations also for full detector observations.

Note that a full-frame read-out results in a FITS file with two HDUs, one for each detector array[6].

Considering all possible combinations of window types and study types, SPICE FITS files may contain `9` distinct types of science data products or HDUs. The data product description is stored in the keyword `DATAPROD`, which is simply the concatenation of `WIN_TYPE` and `STUDYTYP`, e.g. '`Narrow-slit Spectral Raster`'. See also Table 4-8.

Note that in other SPICE documents and applications, e.g. [MAN] and in the SPICE Study Generator tool, the nomenclature is a bit different from what has been outlined above. As an example, instead of differentiating between study types and window types, the Study Generator only uses the concept of study type, which may have one of the following values: "*Full Spectrum*", "*Spatial Scan*", "*Time Series*" and "*Scanned Time Series*". In the pipeline a "*Scanned Time Series*" is treated identical to a series of repeated "*Spatial Scans*":

- in both cases the series of spatial scans are stored in multiple L1, L2 and L3 R FITS files – each file contains a single spatial scan per readout window
- In both cases all spatial scan repetitions will be gathered in a single L3 C file

4.1 Filename, L1, L2, L3 R, and L3 C

Following the specifications in [METADATA], the SPICE FITS file names have the following format:

```
solo_[level]_spice[-concat][-slit][-type][-db][-int]_[time]_V[version]_SPIOBSID-RASTERNO.fits
```

– `[level]` is `L1`, `L2`, or `L3`.
– `[-concat]` is `-concat` for Concatenated L3 files, otherwise empty.
– `[-slit]` is either `w` or `n`, for "wide" (30") or "narrow" (2"/4"/6") respectively.
– `[-type]` is either -`ras` (for "raster"), -`sit` (for "sit-and-stare") or -`exp` (for "single exposure").
– `[-db]` is `-db` for files that include separate dumbbell extensions, otherwise empty. Note that due to technical reasons HDUs storing narrow-slit spectral-profile windows may contain dumbbell data, see page 26 of Section 4. The `[-db]` descriptor of such files are empty.
– `[-int]` is `-int` for files that include intensity-windows, otherwise empty.
– `[time]` is the UTC time at the beginning of the study. For Concatenated L3 files it is the start time of the first observation in the concatenated file.
– `[version]` is an incremental number padded with '`0`' to two characters.
– The combination of `SPIOBSID` (SPICE Observation ID) and `RASTERNO` uniquely identifies a single observation. If a study is repeated all files in the series will have the same `SPIOBSID` but different `RASTERNO`.

[6] SPICE is capable of transmitting full-frame data in both compressed and uncompressed format. In this case the L0+ files contain *four* HDUs, *two* for each detector array, one with decompressed data and one with the uncompressed data.

– The combination of `level`, `version`, `SPIOBSID` and `RASTERNO` uniquely identifies a single FITS file

4.2 Filename, L3 QL

Quicklook L3 files contain an image or a movie of a single data product or a single detector plane readout window extracted from an L2, L3 R or L3 C FITS file. Based on the names of these files, the L3 QL files will have the following format:

```
solo_L3-spice[-concat]-ql-[line][-product][-slit][-type][-db][-int]_[time]
_V[version]_SPIOBSID_RASTERNO.[ext]
```

- `[-concat]` is `-concat` for files stemming from L3 C files, otherwise empty.
- `[-line]` contains a string identifying the emission line.
- `[-product]` contains the name of the derived data product for narrow-slit observations, otherwise empty. The list of possible names is TBD.
- `[-slit]` is either `w` or `n`, for “wide” (30") or “narrow” (2"/4"/6") respectively.
- `[-type]` is either `-ras` (for “raster”), `-sit` (for "sit-and-stare") or `-exp` (for “single exposure”).
- `[-db]` is `-db` for dumbbell images/movies, otherwise empty
- `[-int]` is `-int` for intensity-windows images/movies, otherwise empty.
- `[time]` is the UTC time at the beginning of the study. For files stemming from L3 C files it is the start time of the first observation in the concatenated file.
- `[version]` is an incremental number padded with ‘0’ to two characters.
- `[ext]` is either `mpeg` or `jpeg`.

4.3 The `PCi_j` transformation matrix of FITS files

In general, the WCS coordinate c_i of a pixel with pixel indices $p_j=(p_1,p_2,p_3,\ldots,p_N)$ is expressed by:

$$c_i(p_1, p_2, p_3, \ldots, p_N) = \mathrm{CRVAL}i + \mathrm{CDELT}i \sum_{j=1}^{N} \mathrm{PCi\_j}(p_j - \mathrm{CRPIX}j) \tag{1}$$

Thus, for a four dimensional data cube, the WCS coordinate c_i of a pixel with indices $p_j=(p_1,p_2,p_3,p_4)$ is expressed by:

$$\begin{aligned} c_i(p_1,p_2,p_3,p_4) = \mathrm{CRVAL}i + \mathrm{CDELT}i*(&\mathrm{PC}i\_1*(p_1-\mathrm{CRPIX1}) + \\ &\mathrm{PC}i\_2*(p_2-\mathrm{CRPIX2}) + \\ &\mathrm{PC}i\_3*(p_3-\mathrm{CRPIX3}) + \\ &\mathrm{PC}i\_4*(p_4-\mathrm{CRPIX4})) \end{aligned}$$

For HDUs where each coordinate is coupled only to "its own" dimension (i.e. coordinate `i=1` is coupled to data cube dimension `j=1` only, coordinate `i=2` is coupled to dimension `j=2` only, etc), the `PCi_j` matrix has only diagonal entries and all off-diagonal entries have the default value of `0`. This represents cases with no shear, rotation, mirroring, or transposition.

In L1+ (L1, L2 and L3) FITS files the rotation of the field-of-view is described by four `PCi_j` matrix elements: for a non-zero roll angle `PC1_1` and `PC2_2` do not have the default value of `1`, and `PC1_2` and `PC2_1` do not have the default value of `0`. The spacecraft counter-clockwise roll angle relative to Solar north is given in the CROTA keyword. In L2 FITS files all geometrical distortions other than the rotation of the FOV are corrected for by linear interpolation of the data, and these distortions are therefore not described by FITS keywords. However, the L1 to L2 calibration routine (spice_prep.pro) will be made publicly available and the users will be able to turn off the data interpolation if desirable. In such cases the geometrical distortions will be described by FITS keywords, see [DISTORTIONS] and [DISPERSION].

For both Low Latency FITS files and L1+ FITS files, the jth dimension of the data cube may contribute to the ith coordinate for raster scans and wide-slit (or separately downlinked dumbbell) observations. In these cases, the off-diagonal entries of the PCi_j matrix are not 0:

- For raster observations, time increases for each new slit position. Since the scan direction of the scan mirror is from Solar West to Solar East this means that the time coordinate t *decreases* in the X dimension (j=1) of the data cube. Therefore, the PC4_1 element of the transformation matrix is less than zero, coupling the t coordinate to the X dimension of the data cube.
- For all wide-slit (and dumbbell) HDUs, the x coordinate increases in the dispersion dimension (j=3), and the headers have a non-zero PC1_3 matrix element. The value of PC1_3 is given by the spatial pixel size in the dispersion dimension divided by CDELT1.

L1, L2, and L3 R files do not contain multiple repetitions of raster scans; therefore, the time dimension of raster scan data cubes is degenerate and there is no natural value for CDELT4. On the other hand, CDELT4 cannot be zero, since the product CDELT4*PC4_1 must be equal to the time between two consecutive exposures within a raster scan. We therefore let CDELT4 have the default value of 1, and PC4_1 is thereby equal to the time between exposures.

For the HDUs of L3 R and L3 C files the dispersion coordinate is singular, and CDELT3 therefore has the default value of 1.

For sit-and-stare observations the t coordinate is only dependent on the fourth dimension (j=4), and PC4_1 in the equation above is 0. The time between two consecutive exposures is CDELT4 for sit-and-stare observations.

Note that the time coordinate t is the centre time of an exposure relative to the start time of the observation as a whole, i.e. the relative start time of the exposure + XPOSURE/2. Thus, the t coordinate of the first exposure of an observation is XPOSURE/2, the t coordinate of the exposure that corresponds to the reference pixel is CRVAL4, and t of the last exposure is the end time of the observation –XPOSURE/2. This is true for both rasters and sit-and-stare observations, even if the reference pixel for rasters corresponds to exposure number NAXIS1/2, while the reference pixel for sit-and-stare observations corresponds to exposure 1.

4.4 Data arrays and FITS headers

4.4.1 Data arrays and FITS headers: L1

Although we encourage users to use L2 files for their scientific analysis, we start by describing L1 files. These files contain additional metadata and binary table extensions that are not present in L2 files.

4.4.1.1 General format of L1 data arrays

The dimensions of the L1 data arrays are `(X,Y,D,t)` for all 6 window types in Table 4-1.

The `X` dimension of a data array always denotes the number of slit positions during the observation, and `Y` is the height along the slit, in pixels[8].

The third dimension, `D`, always represents a position along the dispersion direction of the detector. For type 1, this dimension unambiguously corresponds to wavelength (lambda). For wide-slit (and dumbbell) HDUs, however, this dimension corresponds to both spatial `X` and wavelength. Thus for any given `X` and `t`, the `(Y,D)` plane shows a transposed spatial image. However, the `X` position (relative to the centre of the solar disc) of every pixel in such an image is still given by the first WCS coordinate (`CTYPE1='HPLN-TAN'`). This means that there is an off-diagonal value in the `PCi_j` matrix to couple the `X` and `D` dimensions of the data cube, see Section 4.3

Dimension four, `t`, always represents the number of exposures per slit position. I.e. for sit-and-stare observations, `t` is the number of exposures and `X` is 1. For raster observations, where the slit is moved between each exposure, `t` is 1. For such observations, there is an off-diagonal element in the `PCi_j` matrix to couple time and the `X` position of the slit, see Section 4.3. In L3 FITS files, data cubes may represent repeated rasters, i.e. raster movies, and both `X` and `t` may be greater than 1.

Intensity-window observations, window type 4, have only been tested briefly and are not included in regular SPICE operations. These observations are binned in the dispersion dimension with a binning factor equal to the width of the window. Intensity-windows can only be defined for raster studies. The size of the resulting data cube is `(X,Y,1,1)`, i.e. it has no spectral or temporal resolution. The wavelength coordinate is such that the value reflects the central wavelength (i.e. the midpoint between the central two pixels) of the readout window. Intensity-windows are normally observed in pairs: one window covers an emission line (line-window), and a nearby window covers a region with no strong emission lines (background-window). The data from the background-window may be used as an estimate of the background level of the window covering the emission line. Such a background intensity-window may be used as a background estimate for multiple line-windows. The Science Data Pipeline uses line/background intensity-window pairs to create background-subtracted intensity data products, see Section 1.1.1.1. The link between an HDU storing the emission line data and the corresponding HDU storing the background data is established with FITS keywords, see Section 4.4.1.7.

For all study types, the maximum number of exposures is 480, i.e. `X` and `t` are always less than 480.

4.4.1.2 L1 FITS Header Example

Below is an example header of an L1 observational HDU including all keywords required by [METADATA].

The header describes a raster observation recorded obtained using the 4” slit. Sections 4.4.1.4 to 4.4.1.8 describe the data arrays resulting from all available window types and study types.

[8] When referring to pixels along detector Y (height along the slit) or pixels along the dispersion direction (lambda/X’), these might represent *binned* pixels, depending on the study.

A selection of the FITS keywords is further described in Section 4.4.1.3. The correspondence between the metadata parameter names used in [DATAICD] and the FITS keywords of the L1 HDUs is explained in Appendix B. The correspondence between the Study Generator fields and FITS keywords is outlined in Appendix C.

Note that although the example is for a primary HDU and thus contain `SIMPLE=T` and `EXTEND=T`, it might just as well have been an image extension HDU, with modifications as required by [FITSpaper].

```
SIMPLE  =                    T / Written by IDL:  Sun Oct  8 22:50:48 2023
BITPIX  =                   16 / Integer*2 (short integer)
NAXIS   =                    4 / Number of dimensions
NAXIS1  =                   64 / Number of slit positions (x)
NAXIS2  =                  768 / Number of pixels along slit (y)
NAXIS3  =                   32 / Number of pixels in dispersion dimension
NAXIS4  =                    1 / Number of exposures per slit position (time)
EXTEND  =                    T /  FITS dataset may contain extensions
DATE    = '2023-10-08T23:03:04' / Date and time of FITS file creation

EXTNAME = 'Mg IX 706 - Peak'   / Extension name
FILENAME= 'solo_L1_spice-n-ras_20230927T050534_V02_218103821-002.fits' / Filenam


           --------------------------------------------------------
           | Study parameters valid for all Obs-HDUs in this file |
           --------------------------------------------------------
STUDYTYP= 'Raster  '           / Sit-and-stare, Raster or Single Exposure
STUDYDES= 'AR case, East Limb, FOV=4.3", 1" below, 3.4" above the limb' / Study
STUDY   = 'SCI_LIMB-FAST_SC_SL04_19.7S_FF' / SPICE Study name
OBS_MODE= 'SCI_LIMB-FAST_SC_SL04_19.7S_FF' / = STUDY
OBS_TYPE= 'EX8t    '           / Unique code for OBS_MODE
AUTHOR  = 'Tim Grundy'         / Author of study
OBS_ID  = 'SSPI_130A_LF6_111_EX8t_112' / SOC Observation ID
SPIOBSID=            218103821 / SPICE Observation ID (hex d00000d)
OBS_DESC= 'Limb dynamics raster - from SOOP template' / Observation description
PURPOSE = 'Science '           / Purpose of study (Science/Calibration/Checkout)
READMODE= 'Destructive'        / Destructive or non-destructive
TRIGGERD= 'none    '           / Event that triggered observation
TARGET  = 'none    '           / Planned type of target
SOOPNAME= 'L_FULL_LRES_MCAD_Probe-Quadrature' / SOOP Campaign name(s)
SOOPTYPE= 'LF6     '           / SOOP Campaign name code(s)
STP     =                  277 / SoLO Short-Term Plan number
SETFILE = 'flt16-study-set.xml' / Study Set (study definitions) filename
SETVER  =                   25 / Study Set version
APID    =                 1404 / Application Process ID
NRASTERS=                    3 / Number of planned rasters for this SPIOBSID
RASTERNO=                    2 / Raster number (starting at 0)
STUDY_ID=                    1 / On-board Study ID slot (0-63)
MISOSTUD=                 2220 / Ground study ID used in MISO planning tool
XSTART  =                  150 / [arcsec] Slit instr. x rel. to instr. boresight
XPOSURE =              19.7000 / [s] Total effective exposure time
```

```
FOCUSPOS=                12568 / Focus position
NSEGMENT=                    2 / Number of segments per window

POINT_ID= '00001756'           / SVO pointing ID
MOSAICID= '        '           / Blank when study is not a mosaic
SVO_GRP = 'L_FULL_LRES_MCAD_Probe-Quadrature' / SVO file group ID, =SOOPNAME

NWIN    =                   11 / Total number of windows in this file
NWIN_PRF=                   11 / Number of windows not Dumbbell or Intensity
NWIN_DUM=                    0 / Number of Dumbbell windows
NWIN_INT=                    0 / Number of Intensity-windows

PXCOV3  = '82-113, 750-781, 854-885, 920-951, 952-967, 1135-1166, 1701-1732, 1&'
CONTINUE  '733-1764, 1793-1824, 1843-1874, 1875-1906' / [pixel] All PXBEG3-PXEND


           ------------------------------------------------------
           | Other keywords valid for all Obs-HDUs in this file |
           ------------------------------------------------------
TIMESYS = 'UTC     '           / TIMESYS included for readability
DATEREF = '2023-09-27T05:05:34.083' / [UTC] Zero point of time coordinate
DATE-BEG= '2023-09-27T05:05:34.083' / [UTC] Beginning of data acquisition
DATE-OBS= '2023-09-27T05:05:34.083' / [UTC] Equals DATE-BEG
DATE-AVG= '2023-09-27T05:16:14.079' / [UTC] Data acquisition midpoint
DATE-END= '2023-09-27T05:26:53.785' / [UTC] End of data acquisition
SEQ_BEG = '2023-09-27T04:22:32' / [UTC] Approximated start of observation series
TELAPSE =       1279.70200014 / [s] Elapsed time between beg. and end of acqu.
OBT_BEG =       749106138.751 / Start acquisition time in OBT
LEVEL   = 'L1      '           / Data processing level
CREATOR = 'SDP-SPICE'          / Name of pipeline
ORIGIN  = 'University of Oslo' / Name of institution
VERS_SW = '4010    '           / UiO SVN revision number of L1 pipeline
VERSION = '02      '           / Incremental file version number
OBSRVTRY= 'Solar Orbiter'      / Observatory Name
INSTRUME= 'SPICE   '           / Instrument name
CROTA   =       7.34295325766 / [deg] S/C counter-clockwise roll rel to Solar N

COMPLETE= 'C       '           / Complete data set, all windows combined
PCT_CMPL=             100.000 / Completeness of data set, all windows combined


           -------------------------------------
           | STUDYFLG and its derived keywords |
           -------------------------------------
STUDYFLG=                    0 / Study flags
NOSPECTR=                    0 / Applies only to dumbbells
CALMODE =                    0 / Applies only to full frame readouts
DBLEXP  =                    0 / If set, double exposure is enabled
DBLEXPNO=                    0 / Applies only when DBLEXP=1
DARKMAP =                    0 / If set, a dark map was subtracted on-board
BLACKLEV=                    0 / If set, a bias frame was subtracted on-board


           -----------------------------------------------
```

```
           | Keywords valid for this HDU (WINDOW0_70.61) |
           ----------------------------------------------
WIN_TYPE= 'Narrow-slit Spectral' / Description of window type
DATAPROD= 'Narrow-slit Spectral Raster' / WIN_TYPE+STUDYTYP
TELESCOP= 'SOLO/SPICE/SW'       / Telescope/Sensor name/Detector array name
DETECTOR= 'SW      '            / Detector array name
WINNO   =                     0 / Window number (starting at 0) within this study
WINTABID=                    90 / Index in on-board window data table (0-255)
MISOWIN =                  2907 / Ground window ID used in MISO planning tool
WINSHIFT=                    -7 / [pixel] Win redshift rel to win 2907 base pos.
SLIT_ID =                     2 / Slit ID (0-3)
SLIT_WID=                     4 / [arcsec] Slit width
DUMBBELL=                     0 / 0/1/2: not a dumbbell/lower dumbbel/upper dumbb

WAVEUNIT=                    -9 / Power of 10 by which the metre is multiplied
WAVEREF = 'vacuum  '            / Wavelengths are given in vacuum
WAVEMIN =               70.4577 / [nm] Left edge of first read detector pixel
WAVEMAX =               70.7636 / [nm] Right edge of last read detector pixel
WINWIDTH=              0.305816 / [nm] Window width

BTYPE   = 'Intensity'           / Type of data
UCD     = 'phot.count;em.line'  / Unified Content Descriptors v1.23
BUNIT   = 'adu     '            / Units of uncalibrated data

BSCALE  =               1.00000 / Data value = BZERO + BSCALE*FITS array value
BZERO   =                     0 / Default value for unsigned integers
BLANK   =                 65535 / Value of undefined and lost pixels
NTOTPIX =               1572864 / Number of potentially usable pixels =all for L1
NDATAPIX=               1572864 / Number of usable pixels excl. saturated & BLANK
NSATPIX =                     0 / Number of fully saturated pixels
NLOSTPIX=                     0 / Number of lost pixels
NAPRXPIX=                     0 / Number of approximated pix. b/c telemetry loss

PCT_DATA=               100.000 / NDATAPIX/NTOTPIX*100
PCT_SATP=               0.00000 / NSATPIX/ NTOTPIX*100
PCT_LOST=               0.00000 / NLOSTPIX/NTOTPIX*100
PCT_APRX=               0.00000 / NAPRXPIX/NTOTPIX*100

DATAMIN =               563.000 / [adu] Minimum data value
DATAMAX =               5814.00 / [adu] Maximum data value
DATAMEAN=               486.780 / [adu] Mean    data value
DATAMEDN=               468.000 / [adu] Median  data value
DATAP01 =               422.000 / [adu] 1st  percentile of data values
DATAP10 =               441.000 / [adu] 10th percentile of data values
DATAP25 =               452.000 / [adu] 25th percentile of data values
DATAP75 =               494.000 / [adu] 75th percentile of data values
DATAP90 =               540.000 / [adu] 90th percentile of data values
DATAP95 =               588.000 / [adu] 95th percentile of data values
DATAP98 =               681.000 / [adu] 98th percentile of data values
DATAP99 =               788.000 / [adu] 99th percentile of data values
DATARMS =         96.3365126028 / [adu] RMS dev: sqrt(sum((data-DATAMEAN)^2)/N)
DATANRMS=        0.197905722704 / Normalised RMS dev: DATARMS/DATAMEAN
DATAMAD =         40.5195704912 / [adu] Mean abs dev: sum(abs(data-DATAMEAN))/N
DATASKEW=         19.7030628669 / Data skewness
```

```
DATAKURT=        756.410395647 / Data kurtosis

PXBEG1  =                   64 / [pixel] First read-out pixel in X dimension
PXEND1  =                    1 / [pixel] Last  read-out pixel in X dimension
PXBEG2  =                  101 / [pixel] First read-out pixel in Y dimension
PXEND2  =                  868 / [pixel] Last  read-out pixel in Y dimension
PXBEG3  =                   82 / [pixel] First read-out pixel in dispersion dim.
PXEND3  =                  113 / [pixel] Last  read-out pixel in dispersion dim.
PXBEG4  =                    1 / [pixel] First read-out pixel in time dimension
PXEND4  =                    1 / [pixel] Last  read-out pixel in time dimension

NBIN1   =                    1 / Binning factor in X dimension
NBIN2   =                    1 / Binning factor in Y dimension
NBIN3   =                    1 / Binning factor in dispersion dimension
NBIN4   =                    1 / Binning factor in time dimension
NBIN    =                    1 / Total binning factor

COMPRESS= 'Spatial Lossy'      / JPEG Compression description
COMP_RAT=              4.00000 / Compression ratio decompressed/compressed
COMPQUAL=             0.250000 / Compression ratio compressed/decompressed
COMPTYPE=                    6 / Compression type (0-7)
COMPPARA=                   64 / Compression amount parameter (0-255)
SHCFFTID=                    0 / Applies only to SHC-compressed data
COMP_ALG= 'Lossy/Spatial Lossy/4/6/64' / Lossy/COMP(RESS/_RAT/TYPE/PARA)


           ---------------------------------------------------------
           | Keywords describing telemetry of non-missing segments |
           ---------------------------------------------------------
NPACKETS=                  192 / Number of packets with observational data

LOSTPKTS=                    0 / Number of lost packets w/data, variable keyword
LOSTBINS=                    0 / Applies only to SHC-compressed data

NLOSTCHK=                    0 / Number of lost checksum packets
NFAILCHK=                    0 / Number of checksums failed
NAPRXPLN=                    0 / Number of approximated X-Y plane sections
NLOSTPLN=                    0 / Number of lost X-Y plane sections



           ------------------------------------------
           | World Coordinate System (WCS) keywords |
           ------------------------------------------
WCSNAME = 'Helioprojective-cartesian' / Name of coordinate system
LONPOLE =              180.000 / [deg] Native longitude of celestial pole
SPECSYS = 'TOPOCENT'           / Spectral coord. not corrected for S/C velocity
VELOSYS =              0.00000 / [m/s] Default for SPECSYS='TOPOCENT'

CTYPE1  = 'HPLN-TAN'           / Type of 1st coordinate
CNAME1  = 'Helioprojective longitude (Solar X), increases towards Solar West' /
CUNIT1  = 'arcsec  '           / Units for 1st coordinate (for CRVAL1, CDELT1)
CRVAL1  =     -2617.06821710 / [arcsec] 1st coordinate of reference point
CDELT1  =              4.00000 / [arcsec] Increment of 1st coord at ref point
```

```
CRPIX1  =              32.5000 / [pixel] 1st pixel index of reference point
PC1_1   =       0.991798906611 / Non-default value due to CROTA degrees S/C roll
PC1_2   =     -0.0350926093219 / Contribution of dim 2 to coord 1 due to roll
CRDER1  =       0.216088224675 / [arcsec] Mean stddev of Solar X
CWERR1  =       0.705631065045 / [arcsec] Max absolute distortion, Solar X
CWDIS1  = 'Lookup  '           / Type of WCS distortion correction
DW1     = 'EXTVER: 1'          / Extension version
DW1     = 'NAXES: 1'           / Axes in the distorion array
DW1     = 'AXIS.1: 1'          / Correction of Solar X
DW1     = 'ASSOCIATE: 1'       / Association stage is pixel coordinates
DW1     = 'APPLY: 6'           / Application stage is world coordinates

CTYPE2  = 'HPLT-TAN'           / Type of 2nd coordinate
CNAME2  = 'Helioprojective latitude (Solar Y), increases towards Solar North' /
CUNIT2  = 'arcsec  '           / Units for 2nd coordinate (for CRVAL2, CDELT2)
CRVAL2  =       -410.292032123 / [arcsec] 2nd coordinate of reference point
CDELT2  =              1.09829 / [arcsec] Increment of 2nd coord at ref point
CRPIX2  =              384.500 / [pixel] 2nd pixel index of reference point
PC2_1   =       0.465480600047 / Contribution of dim 1 to coord 2 due to roll
PC2_2   =       0.991798906611 / Non-default value due to CROTA degrees S/C roll
CRDER2  =       0.205898356675 / [arcsec] Mean stddev of Solar Y
CWERR2  =       0.631374091770 / [arcsec] Max absolute distortion, Solar Y
CWDIS2  = 'Lookup  '           / Type of WCS distortion correction
DW2     = 'EXTVER: 2'          / Extension version
DW2     = 'NAXES: 1'           / Axes in the distorion array
DW2     = 'AXIS.1: 1'          / Correction of Solar Y
DW2     = 'ASSOCIATE: 1'       / Association stage is pixel coordinates
DW2     = 'APPLY: 6'           / Application stage is world coordinates

CTYPE3  = 'WAVE    '           / Type of 3rd coordinate
CNAME3  = 'Wavelength'         / Description of 3rd coordinate
CUNIT3  = 'nm      '           / Units for 3rd coordinate (for CRVAL3, CDELT3)
CRVAL3  =              70.6106 / [nm] 3rd coordinate of reference point
CDELT3  =           0.00955676 / [nm] Increment of 3rd coord at ref point
CRPIX3  =              16.5000 / [pixel] 3rd pixel index of reference point
PC3_3   =              1.00000 / Default value, no rotation

CTYPE4  = 'UTC     '           / Type of 4th coordinate
CNAME4  = 'Time (Degenerate Dimension)' / Description of 4th coordinate
CUNIT4  = 's       '           / Units for 4th coordinate (for CRVAL4, CDELT4)
CRVAL4  =        639.850000000 / [s] 4th coord of ref point, relative to DATEREF
CDELT4  =        1.00000000000 / [s] Degenerate dimension; default value
CRPIX4  =              1.00000 / [pixel] 4th pixel index of reference point
PC4_4   =              1.00000 / Default value, no rotation
PC4_1   =       -20.0000000000 / Contribution of dimension 1 to coordinate 4


           -----------------------------------------------------------------
           | Auxiliary data and reference to bintab with variable keywords |
           -----------------------------------------------------------------
VAR_KEYS= 'VARIABLE_KEYWORDS;TIMAQOBT,MIRRPOS,TN_FOCUS,TN_GRAT,TN_SW,TN_LW,T_F&'
CONTINUE  'OCUS,T_GRAT,T_SW,T_LW,VN_MCPSW,VN_MCPLW,VN_GAPSW,VN_GAPLW,V_MCPSW,V&'
CONTINUE  '_MCPLW,V_GAPSW,V_GAPLW,TIMAQUTC,CRDER1,CRDER2                     &'
CONTINUE  '' / Variable keywords
```

```
TIMAQOBT=        749106768.751 / [OBT] Average Start time of data acquisition
MIRRPOS =              39051.4 / [adu] Average Scan mirror position
TN_FOCUS=              1727.00 / [adu]     Average SFM focus   temperature
TN_GRAT =              1585.09 / [adu]     Average SFM grating temperature
TN_SW   =              2778.48 / [adu]     Average HAS SW      temperature
TN_LW   =              2779.48 / [adu]     Average HAS LW      temperature
T_FOCUS =              20.4902 / [Celsius] Average SFM focus   temperature
T_GRAT  =              24.8640 / [Celsius] Average SFM grating temperature
T_SW    =             -19.9935 / [Celsius] Average HAS SW      temperature
T_LW    =             -20.0196 / [Celsius] Average HAS LW      temperature
VN_MCPSW=                 2709 / [adu] Average MCP SW voltage
VN_MCPLW=                 2584 / [adu] Average MCP LW voltage
VN_GAPSW=                 2932 / [adu] Average GAP SW voltage
VN_GAPLW=                 2933 / [adu] Average GAP LW voltage
V_MCPSW =                  899 / [V]   Average MCP SW voltage
V_MCPLW =                  848 / [V]   Average MCP LW voltage
V_GAPSW =                 2789 / [V]   Average GAP SW voltage
V_GAPLW =                 2790 / [V]   Average GAP LW voltage
TIMAQUTC= '2023-09-27T05:16:04.084' / [UTC] Average Start t. of data acquisition


           ----------------------------------------
           | Keywords describing processing steps |
           ----------------------------------------
PRSTEP1 = 'COMPRESSION'        / Type of processing, step 1
PRPROC1 = 'JPEG Compression (On-board)' / Name of procedure, step 1

PRSTEP2 = 'TELEMETRY-PARSING'  / XML decoding, decompression if applicable, etc
PRPROC2 = 'spice_process_telemetry.pro' / Name of procedure, step 2
PRPVER2 = '03.06.01'           / Version of procedure, step 2
PRLIB2A = 'uio-spice-pipeline' / Software library containing PRPROC2

PRSTEP3 = 'PIXEL-LEVEL-OFFSET-SUBTRACTION' / Type of processing, step 3
PRPROC3 = 'spice_l0_to_l1.pro' / Name of procedure, step 3
PRPVER3 = '03.06.01'           / Version of procedure, step 3
PRPARA3 =                  200 / Parameters for PRPROC3
PRLIB3A = 'uio-spice-pipeline' / Software library containing PRPROC3

PRSTEP4 = 'SPATIAL-COORDINATE-CORRECTION-X' / Type of processing, step 4
PRPROC4 = 'correct_spice_offset_relative_to_spacecraft' / Name of procedure, ste
PRPVER4 = '1.0     '           / Version of procedure, step 4
PRPARA4 = 'delta_instrument_x=-73.56256' / =T_GRAT*0.46-85.0, param for PRPROC4
PRLIB4A = 'uio-spice-pipeline' / Software library containing PRPROC4

PRSTEP5 = 'SPATIAL-COORDINATE-CORRECTION-Y' / Type of processing, step 5
PRPROC5 = 'correct_spice_offset_relative_to_spacecraft' / Name of procedure, ste
PRPVER5 = '1.0     '           / Version of procedure, step 5
PRPARA5 = 'delta_instrument_y=-72.4' / =constant, parameter for PRPROC5
PRLIB5A = 'uio-spice-pipeline' / Software library containing PRPROC5


           -------------------------------------------------
           | Hardware and software processing environment  |
           -------------------------------------------------
```

```
LONGSTRN= 'OGIP 1.0'           / The OGIP long string convention may be used.
COMMENT This FITS file may contain long string keyword values that are
COMMENT continued over multiple keywords.  This convention uses the  '&'
COMMENT character at the end of a string which is then continued
COMMENT on subsequent keywords whose name = 'CONTINUE'.
PRENV2  = '  Kernel: Linux                                                    &'
CONTINUE  '  Kernel release number: 3.10.0-1160.90.1.el7.x86_64               &'
CONTINUE  '  Architecture: x86_64                                             &'
CONTINUE  '  Host name: astro-sdc-fs.uio.no                                   &'
CONTINUE  '  OS: Red Hat Enterprise Linux Server release 7.9 (Maipo)          &'
CONTINUE  '  CPU: Intel(R) Xeon(R) CPU E5-2630L v4 @ 1.80GHz                  &'
CONTINUE  '  IDL 8.8.2 (Apr 06 2022 (430118)), memory bits: 64, file offset bi&'
CONTINUE  'ts: 64       ' / Hardware and software


           ----------------------------------------------
           | SOLARNET keywords, and additional keywords |
           ----------------------------------------------
SOLARNET=              1.00000 / Fully/Partially/No SOLARNET compliant (1/0.5/-1
OBS_HDU =                    1 / HDU contains observational data (1) or not (0)
PARENT  = 'solo_L0_spice-n-ras_0749106139_V202310082250C_218103821-002.fits' / L
FILE_RAW= 'sc_2023_09_26.xml;sc_2023_09_27.xml' / Telemetry file
INFO_URL= 'https://spice.ias.u-psud.fr/' / URL to additional information


           ----------------------------
           | Solar Ephemeris Keywords |
           ----------------------------
DSUN_OBS=        55110443473.1 / [m]  S/C distance from Sun
DSUN_AU =       0.368390577399 / [AU] S/C distance from Sun
RSUN_ARC=        2603.69512960 / [arcsec] Apparent photospheric Solar radius
RSUN_REF=          6.95700E+08 / [m]      Assumed  photospheric Solar radius
SOLAR_B0=       -7.98698708837 / [deg] Tilt angle of Solar North toward S/C
SOLAR_P0=       -26.4031244523 / [deg] S/C Celestial North to Solar North angle
SOLAR_EP=       -4.01777025122 / [deg] S/C Ecliptic  North to Solar North angle
CAR_ROT =                 2276 / Carrington rotation number
HGLT_OBS=       -7.98698708837 / [deg] S/C Heliographic latitude (B0 angle)
HGLN_OBS=       -164.204283381 / [deg] S/C Heliographic longitude
CRLT_OBS=       -7.98698708837 / [deg] S/C Carrington   latitude (B0 angle)
CRLN_OBS=        245.836092183 / [deg] S/C Carrington   longitude (L0 angle)
HEEX_OBS=       -53054663461.3 / [m]   S/C Heliocentric Earth Ecliptic X
HEEY_OBS=       -14794693589.0 / [m]   S/C Heliocentric Earth Ecliptic Y
HEEZ_OBS=       -1865665140.34 / [m]   S/C Heliocentric Earth Ecliptic Z
HCIX_OBS=       -30276080788.4 / [m]   S/C Heliocentric Inertial X
HCIY_OBS=        45407958131.0 / [m]   S/C Heliocentric Inertial Y
HCIZ_OBS=       -7657496344.10 / [m]   S/C Heliocentric Inertial Z
HCIX_VOB=       -33185.5821441 / [m/s] S/C Heliocentric Inertial X Velocity
HCIY_VOB=       -47756.6412276 / [m/s] S/C Heliocentric Inertial Y Velocity
HCIZ_VOB=        3319.63532324 / [m/s] S/C Heliocentric Inertial Z Velocity
HAEX_OBS=        38166.4958182 / [m]   S/C Heliocentric Aries Ecliptic X
HAEY_OBS=       -43918.5266620 / [m]   S/C Heliocentric Aries Ecliptic Y
HAEZ_OBS=       -2735.18464885 / [m]   S/C Heliocentric Aries Ecliptic Z
HEQX_OBS=       -52514978526.7 / [m]   S/C Heliocentric Earth Equatorial X
HEQY_OBS=       -14856000803.8 / [m]   S/C Heliocentric Earth Equatorial Y
```

```
HEQZ_OBS=       -7657496344.10 / [m]   S/C Heliocentric Earth Equatorial Z
GSEX_OBS=        203022566111. / [m]   S/C Geocentric   Solar Ecliptic X
GSEY_OBS=        14794693589.0 / [m]   S/C Geocentric   Solar Ecliptic Y
GSEZ_OBS=       -1865665140.34 / [m]   S/C Geocentric   Solar Ecliptic Z
OBS_VR  =       -21578.8889377 / [m/s] Radial velocity of S/C away from the Sun
EAR_TDEL=        316.410425431 / [s] Time(Sun to Earth)-Time(Sun to S/C)
SUN_TIME=        183.828652131 / [s] Time(Sun centre to S/C)
DATE_EAR= '2023-09-27T05:10:50.493' / [UTC] DATE-BEG + EAR_TDEL
DATE_SUN= '2023-09-27T05:02:30.254' / [UTC] DATE-BEG - SUN_TIME


           -------------------------
           | HISTORY and checksums |
           -------------------------
HISTORY process_request.pro
HISTORY spice_process_telemetry.pro
HISTORY spice_l0_to_l1.pro
DATASUM = '1861580276'         / data unit checksum created 2023-10-08T21:03:15
CHECKSUM= 'lGBjm9BglEBgl9Bg'   / HDU checksum created 2023-10-08T21:03:15
O_BLANK =                32767 / Original BLANK value
O_BZERO =              32768.0 / Original BZERO Value
END
```

4.4.1.3 Brief description of some selected FITS keywords

4.4.1.3.1 Keywords identifying the study

The keyword `OBS_TYPE` is a 3-character alphanumeric string that uniquely identifies the `OBS_MODE`, which contains the name of the SPICE study. The value of the Solar Obiter-wide `OBS_MODE` is repeated in the SPICE-specific keyword `STUDY`.

The keyword `OBS_ID` contains the SOC Observation ID. The `OBS_ID` is built up of parameters that together uniquely identify the observation in a Solar Orbiter-wide context. The format of the SOC Observation ID is:

    AAAA_PPvV_SSS_III_OOOO_JJJ

The components of this string are:

AAAA: 4-character instrument ID, is always "SSPI" for SPICE
PPvV: 4-character string, Long-Term Planning plan ID
SSS: 3-digit alphanumeric SOOP Type, equals FITS keyword `SOOPTYPE`
III: 3-digit (base-58) SOOP Instance
OOOO: 4-character (base-58) Observation Type, equals FITS keyword `OBS_TYPE`
JJJ: 3-digit (base-58) Observation Instance

An observation may belong to multiple SOOPs. In such cases, the SOC Observation ID contains multiple strings as described above, each sub-string being separated by a semicolon. See [IOR-ICD] for details.

Note that the SPICE Observation ID is given in the keyword `SPIOBSID`. This keyword uniquely identifies an observation series in a SPICE context (the combination of `SPIOBSID` and `RASTERNO` uniquely identifies a single observation). The `SPIOBSID` is a 32-bit on-board monotonically increasing observation counter.

4.4.1.3.2 Keywords derived from the value of STUDYFLG

The 8-bit value of the `STUDYFLG` keyword stems from the Science Header Packets, and the 6 keywords following `STUDYFLG` in the header example are set based on this value, in accordance with Section 4.2.6.1. of [DATAICD].

4.4.1.3.3 Keywords describing the size of the readout windows

Note that in L2 files any adjacent windows are merged, see Section 4.4.2. In such cases the keywords described below apply to the merged window.

The window start column on the detector and the window width are both given in pixel coordinates in the telemetry, see Table 4-10. The values of `WAVEMIN` and `WINWIDTH` are derived from these pixel values, using the conversion from pixel indices to nm given in [DISPERSION] and [MAN]. `WAVEMIN` is the wavelength of the leftmost edge of the first detector pixel column of the readout window, ignoring any binning, and `WINWIDTH` is the edge-to-edge window width. `WAVEMAX` is the wavelength corresponding to the rightmost edge of the window's last pixel column. For narrow-slit observations the pixel value of the window start column (starting at 1 for the leftmost pixel column on the SW detector, 1025 for the leftmost column on the LW detector) can be found in the Solar Orbiter mandatory keyword `PXBEG3`, and the pixel value of the last column is stored in `PXEND3`. The edge-to-edge pixel width of the window is therefore `PXEND3 – PXBEG3 + 1`. `PXCOV3` contains a comma separated list of the detector coverage in the dispersion dimension of all HDUs in the file, i.e. a list of all `PXBEG3`-`PXEND3` ranges.

The window start *row*, starting at 1 for the lowermost[9] pixels on the detector, is also given in the telemetry and is stored in the FITS files as `PXBEG2`. The window end row is stored in `PXEND2`.

The scan direction of the scan mirror is from Solar West to Solar East. This is indicated by always having `PXEND1 = 1`. For raster observations `PXBEG1 > 1`, for other study types `PXBEG1 = PXEND1 = 1`.

`PXBEG4 = 1` for all observation types. For sit-and-stare observations `PXEND4 > 1`, for other study types `PXEND4 = PXBEG4 = 1`.

4.4.1.3.4 The BLANK keyword

The data type of a L1 HDU's data array is normally 16-bit *unsigned* integer. However, when reconstructing SHC-compressed data some pixel values may end up being negative, and the data type for SHC-compressed data is therefore always 16-bit *signed* integer. Due to this difference the value of undefined pixels (`BLANK`) is set to $2^{16}-1$ for unsigned integer array HDUs and $2^{15}-1$ for signed integer array HDUs.

Note that in L2 files the data arrays are of type floating point and the `BLANK` keyword is not present. Undefined pixels in L2 files are set to `NaN`.

[9] i.e. the southernmost detector pixel row in the case of no spacecraft roll

4.4.1.3.5 Keywords describing telemetry and telemetry loss and consequences thereof

L1 and L2 files contain a collection of FITS keywords that briefly describe the completeness of the telemetry of the file as a whole, and of each of the observational HDUs.

4.4.1.3.5.1 L1 and L2 keywords having the same values for all observational HDUs in the file

`COMPLETE`: Completeness of all windows in the file combined. '`C`' if complete,'`I`' if incomplete.
`PCT_CMPL`: Completeness of all windows in the file combined, in percent.

4.4.1.3.5.2 L1 and L2 keywords describing the telemetry of an individual HDU

`NTOTPIX`: The number of potentially usable pixels
`NDATAPIX`: The number of usable pixels excluding saturated & BLANK
`NSATPIX`: The number of fully saturated pixels
`NLOSTPIX`: The number of lost pixels due to telemetry loss
`NAPRXPIX`: The number of pixels with approximated values due to loss of compressed telemetry
`PCT_DATA: NDATAPIX/NTOTPIX*100`
`PCT_SATP: NSATPIX/ NTOTPIX*100`
`PCT_LOST: NLOSTPIX/NTOTPIX*100`
`PCT_APRX: NAPRXPIX/NTOTPIX*100`

4.4.1.3.5.3 L1 only keywords describing the telemetry of an individual HDU in detail

Note that neither the keywords mentioned in this section nor the binary table extensions they are referring to are present in L2 files.

`NPACKETS`: the total number of packets *with observational data* that were prepared on-board for downlink. This number includes the total number of Data Packets of all the window's *n* segments, plus *n* if the *n* Final Packets contained observational data.
`LOSTPKTS`:the number of lost packets with observational data
`LOSTBINS`:the number of lost FFT bins (in the case of SHC data)
`NLOSTCHK`: the number of lost Final Packets (with or without observational data)
`NFAILCHK`: the number of lost data checksum test that have failed
`NAPRXPLN`: the number of approximated image plane ranges due to lost compressed telemetry
`NLOSTPIX`: the number of lost pixels due to lost compressed or uncompressed telemetry
`NLOSTPLN`: the number of lost image plane ranges due to lost compressed telemetry (or lost FFT coefficient planes for SHC data)

Sections 4.4.4.4 and 4.4.5 describe how the indices of lost telemetry packets, lost FFT coefficient planes, approximated image plane ranges, lost image plane ranges, and lost FFT bins are stored in binary table extensions in L1 files. Lost pixels (and lost image planes) can easily be identified by selecting pixels having the `BLANK` value and therefore the indices of such pixels and image planes are not stored in binary table extensions.

4.4.1.3.6 Keywords describing coordinate distortions

After a significant re-pointing or a wheel off-loading it may take tens of minutes, even hours, until the pointing of Solar Orbiter is stable. If SPICE is observing during a period of unstable spacecraft pointing it may be necessary to take the pointing instability into account when calculating the Solar X and Solar Y coordinates of the observation.

For single exposure observations `CRDER1` and `CRDER2` give the standard deviation of the S/C pointing in arc seconds during the exposure. For multi-exposure observations these keywords give the mean of the standard deviations of the S/C pointing. The individual standard deviations are stored in a binary table extension, see Section 4.4.4. for details on *variable keywords*.

The `CWERR1` and `CWERR2` keywords give the maximum absolute deviation in arc seconds from the Solar X and Solar Y coordinates calculated by Equation (1). Following [WCSDISTORTIONS], these deviations, or coordinate distortions, are stored in two separate image extension both having `EXTNAME` = '`WCSDVARR`'. The distortions of Solar X are stored in the `WCSDVARR` extension having `EXTVER=1`, the Solar Y distortions are stored in the `WCSDVARR` extension having `EXTVER=2`.

The distortion mechanisms described in [WCSDISTORTIONS] support distortions of *pixel coordinates* that are used to calculate the world coordinates. In SPICE files, however, we use the SOLARNET mechanisms mentioned in [SSTRED] to store the distortions of the calculated Solar X and Solar Y world coordinates.

We represent the distortions according to the `Lookup` mechanism outlined in Section 3.4 of [WCSDISTORTIONS]. The spacecraft roll is assumed to be constant during the observation, making each of the two coordinate distortion arrays 1-dimensional, with one coordinate distortion value per exposure. See Figure 4-2 for an example of Solar X and Solar Y coordinate distortion arrays. Figure 4-3 shows the corresponding CIII raster image without and with corrected coordinates.

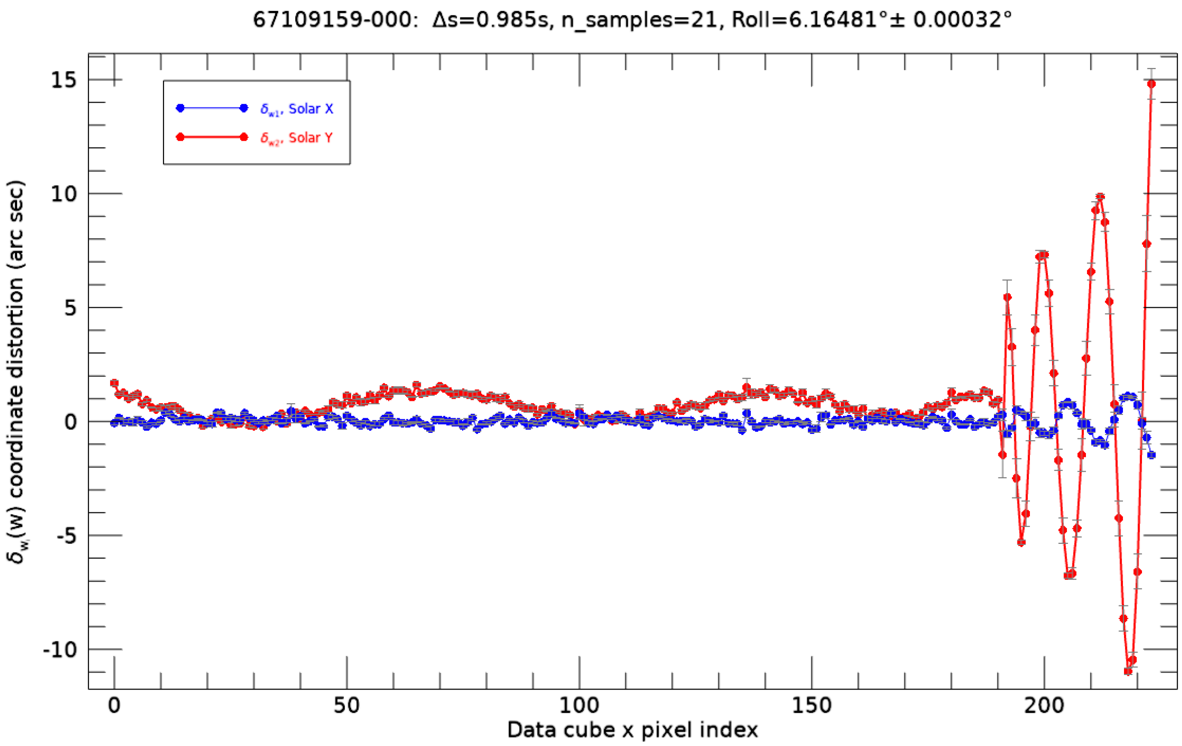

Figure 4-2: 2021-09-14, `SPIOBSID` 67109159: An example of Solar X and Solar Y coordinate distortions due to unusually pronounced spacecraft pointing instability. The error bars correspond to the standard deviation of the coordinate distortions during a 20 sec exposure. These standard deviations are given in the variable keywords `CRDERi`. The maximum amplitudes of the distortions are given by `CWERRi`. The duration of the raster is 1 hour

15 minutes, 224 exposures were taken. Note that the distortion correction values for Solar Y are not fluctuating around 0, but are shifted towards positive distortion. The reason for this shift is that the distortion correction at the reference pixel `CRPIX1=112.5` **is by definition 0.**

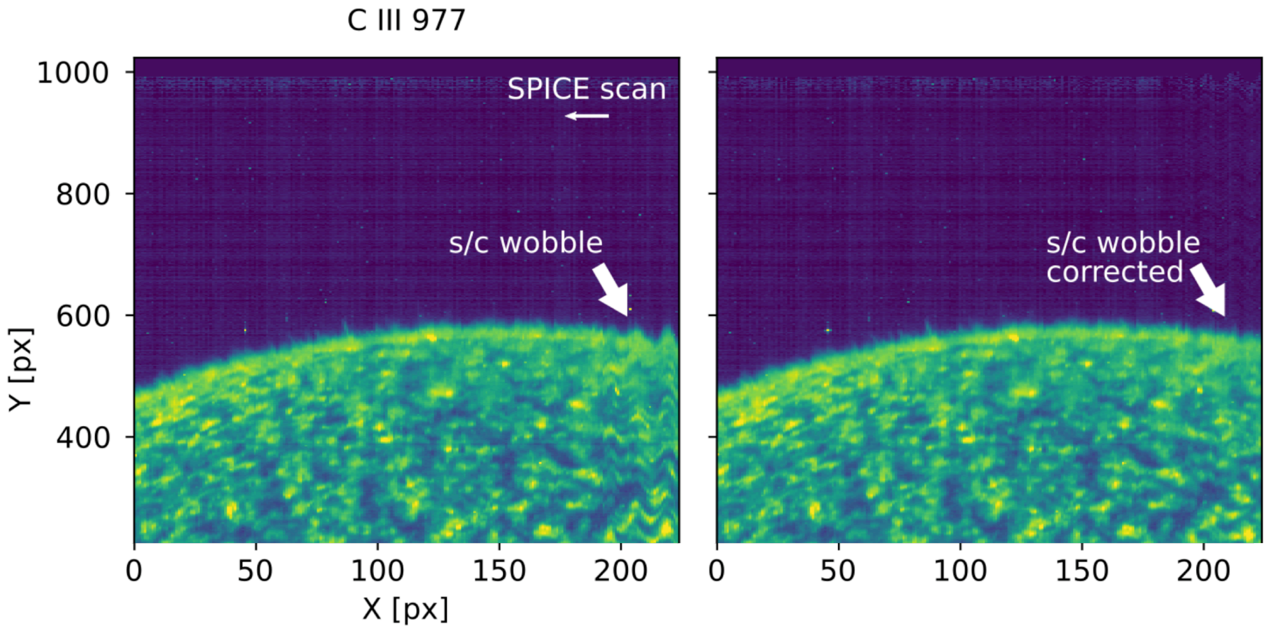

Figure 4-3: Example of a SPICE CIII raster before and after coordinate distortion correction. (G. Pelouze)

The IDL SolarSoft function `wcs_get_coord.pro` supports the SOLARNET interpretation of the `Lookup` mechanism that is used in SPICE FITS files[10]. When calculating coordinates, the distortions are automatically taken into account:

```
IDL> data = readfits(file, header, ext=0)
IDL> wcs = fitshead2wcs(header, filename=file)
IDL> corrected_coordinates = wcs_get_coord(wcs)
```

The SPICE IDL quicklook and analysis tools [IDLANA] use `wcs_get_coord.pro` to calculate corrected coordinates.

As of January 2023, the existing WCS tools in astropy/python do not support the `Lookup` method for calculating coordinate distortions, nor the SOLARNET implementation of the distortion mechanisms. Python users will therefore have to calculate the Solar X and Solar Y coordinates using the existing tools, before manually adding the distortions to obtain the corrected coordinates. I.e., for slit position `n`:

- calculate the Solar X coordinates of all pixels along the slit and add element `n` of the distortion array stored in the `WCSDVARR` extension having `EXTVER=1`
- calculate the Solar Y coordinates of all pixels along the slit and add element `n` of the distortion array stored in the `WCSDVARR` extension having `EXTVER=2`

[10] `wcs_get_coord.pro` version 11 (11th May 2023) and later

4.4.1.3.7 Keywords describing the processing steps

The individual processing steps that have been applied to the data are described by the `PRxxxxn` keywords. The type of processing performed and the name of the procedure performing the processing are given by `PRSTEPn` and `PRPROCn` respectively. For on-board data processing, i.e. binning, dark subtraction and/or compression, `PRSTEPn` and `PRPROCn` are the only `PRxxxxn` keywords that are defined.

Processing performed on-ground are described with additional `PRxxxxn` keywords giving the version number of the processing routine (`PRPVERn`) and the name of the library that contains the processing routine (`PRLIBnA`). Finally, the keyword `PRENVn` describes the hardware and software environment of the on-ground processing. `PRENVn` is valid for all processing steps starting at step `n`, until a new `PRENVn` keyword is defined. However, for SPICE FITS files created by the pipeline, the processing environment is normally the same for all processing steps, and therefore only a single `PRENVn` keyword that is valid for all processing steps performed on ground will normally be defined.

For L0 files the only on-ground processing that is described is the XML telemetry parsing itself, which includes XML decoding and decompression. If a fixed value was added to the data on-board, this value is subtracted when creating a L1 file, and this *pixel level offset subtraction* is described as a separate on-ground processing step.

4.4.1.3.8 Keyword giving the name(s) of the telemetry file(s)

The name of the source telemetry file is given in the keyword `FILE_RAW`. If also a telemetry file from the preceding day is given as input to the pipeline, the value of `FILE_RAW` is the names of the telemetry files of the preceding day and the current day, separated by a semicolon.

4.4.1.3.9 Keywords used to determine the study type: sit-and-stare or raster

The pipeline uses the IORs and study definition files to determine whether a multi-exposure observation is a raster or a sit-and-stare observation. These files are only available for cruise phase and nominal mission phase observations. In order to determine the study type for commissioning phase observations, we must therefore instead use the scan mirror positions reported in the telemetry. Due to readout noise the scan mirror position values are neither constant for sit-and-stare observations nor monotonically increasing for raster observations. Therefore, the study type is determined from a linear fit of the scan mirror positions. For pipeline debugging purposes the parameters of the fit are included in the headers of commissioning observation files: `MIRRDELT` gives the slope of fit and `SMIRRDEL` gives the 1-sigma uncertainty estimate of the fit.

4.4.1.4 L1 HDUs with narrow-slit spectral-profile data (window type 1)

Data arrays:
SPICE narrow-slit (2”, 4” or 6”) spectral-profile data arrays have dimensions according to Table 4-3 below. Readout windows have a width in the dispersion direction of `D = 4, 8, 16` or `32` pixels.

Study type	Dimensions
Sit-and-stare	`(1,Y,D,t)`
Raster	`(X,Y,D,1)`
Single exposure	`(1,Y,D,1)`

Table 4-3: Dimensions of narrow-slit spectral-profile data arrays

4.4.1.5 L1 HDUs with dumbbell stack data (window type 2)

Note that due to technical reasons stand-alone dumbbell observations are rarely recorded, see page 26 of Section 4.

At each end of the slit in the `Y` direction there is an area of nominal size 30" x 30" used for making small context images, so called dumbbell or alignment windows. It is possible to downlink one or both dumbbells, but only for a single window per study. The dumbbells may be downloaded in addition to, or instead of, the spectral-profile data cube, and are included in the same FITS file as the spectral-profile data, in separate HDUs.

Data arrays:
SPICE dumbbell stack data arrays have dimensions according to Table 4-4 below.

Study type	Dimensions
Sit-and-stare	`(1,32,64,t)`
Raster	`(X,32,64,1)`
Single exposure	*Not applicable*

Table 4-4: Dimensions of dumbbell stack data arrays

4.4.1.6 L1 HDUs with wide-slit data (window type 3)

Data array:
SPICE wide-slit (30”) data arrays have dimensions according to Table 4-5 below. Readout windows have in most cases a width in the dispersion direction of `D = 32` pixels.

Study type	Dimensions
Sit-and-stare	`(1,Y,D,t)`
Raster	`(X,Y,D,1)`
Single exposure	`(1,Y,D,1)`

Table 4-5: Dimensions of wide-slit data arrays

4.4.1.7 L1 HDUs with intensity-window data (window type 4)

Note that intensity-window observations are rarely recorded, see page 26 of Section 4.

An intensity-window is binned in the dispersion dimension with a binning factor equal to the width of the window. Intensity-windows are normally observed in pairs, with one window covering a spectral line, and a nearby window covering a part of the spectrum without any strong emission lines. The intensity-window data cubes are saved in separate HDUs. The value of the wavelength coordinate is the central wavelength (i.e. the midpoint between the central two pixels before binning) of the readout window.

Data array:
SPICE intensity-window data arrays have dimensions according to Table 4-6 below.

Study type	Dimensions
Sit-and-stare	*Not applicable*
Raster	(X,Y,1,1)
Single exposure	*Not applicable*

Table 4-6: Dimensions of intensity-window data arrays

If the HDU stems from an intensity-window that was defined as either a line window or a background window in the MISO planning tool, the HDU has two of the following three keywords that are not present in the HDUs of any other window type: `IWINTYPE`, and either `IWINBKG` or `IWINLINE`. HDUs storing stand-alone intensity-windows do not include any of these three keywords.

If the data of the HDU stems from a window covering an emission line the value of `IWINTYPE` is `'LINE'`. In that case an additional keyword `IWINBKG` gives the HDU number of the HDU storing the data that is to be regarded as the background level.

If the data stems from a background window the value of `IWINTYPE` is `'BACKGROUND'`. A background window may be defined as background for multiple `'LINE'` windows. `IWINLINE` gives a comma separated list of the HDU numbers of all `'LINE'` window HDUs that has this window defined as the background.

As an example, the following keywords may be present in an HDU stemming from a window covering the emission line Ne VIII 770 A:

```
EXTNAME  = 'Ne VIII 770 A'     / Extension name
…
WINNO    =                   1 / Window number (starting at 0) within this study
…
IWINTYPE = 'LINE'              / This intensity-window covers an emission line
IWINBKG  =                   4 / HDU number (primary=0) storing background
```

The following keywords may be present in the corresponding HDU with data to be used for background-subtraction from the Ne VIII 770 A line:

```
EXTNAME  = 'Red wing of Ne VIII 770 A' / Extension name
…
WINNO    =
…
IWINTYPE = 'BACKGROUND'       / This intensity-window considered as background
IWINLINE = '0,1,2    '        / HDU numbers (primary=0) storing emission line
```

4.4.1.8 L1 HDUs with full-detector read-out data (window types 5 and 6)

It is possible to make a full-frame read-out where the full areas of both detector arrays are recorded. A full-frame read-out always consists of a single exposure[11]. Data from each detector array are stored in separate HDUs[12]. Both the three narrow slits and the wide slit may be used, in the latter case the `PC1_3` transformation matrix element has a non-zero value.

Data array:
SPICE full-detector read-out data arrays have dimensions according to Table 4-7 below.

Study type	Dimensions
Sit-and-stare	*Not applicable*
Raster	*Not applicable*
Single exposure	(1,1024,1024,1)

Table 4-7: Dimensions of full-detector read-out data arrays

4.4.2 Data arrays and FITS headers: L2

In order to include data from all pixels present in the L1 data array, the geometrically corrected L2 data arrays are by default a few pixels wider in the dispersion dimension compared to the L1 data arrays described in Section 4.4.1. The values of padded pixels are set to `NaN`.

In order to ensure that an emission line is completely covered, or to catch multiple closely spaced lines, two or more adjacent readout windows may be defined. In L2 files the data arrays of such adjacent windows are merged (concatenated) if the windows have the same binning and compression. Neither Intensity-windows nor dumbbells are merged (but neither of these window types have been used other than for a very few test runs).

If adjacent windows have been merged the L2 file will contain fewer image HDUs than the L1 file.

The metadata that describe the data array are updated to reflect the wider dispersion dimension of the merged window, and the reduced number of HDUs in the file. This window concatenation is described by a set of `PRxxxxn` keywords, see the `'WINDOW-CONCATENATION'` processing step in Section 4.4.2.1

Two windows may be very close without being adjacent. Due to the geometrical correction of L2 files, pixels close to the adjoining edge of two closely spaced L2 windows may originally stem from the nearby window. A future update of the calibration routines will ensure that pixels are only present in their original window.

In L2 files the spatial plate scale in the direction of the slit is the same for the two detector arrays, and a given pixel along the slit on the SW or LW detector correspond the same location on the

[11] However, it is possible to repeat full detector observations, with or without moving the scan mirror. Rasters and sit-and-stare observations may therefore be constructed also for full detector observations by combining the individual exposures that are stored in separate FITS files.
[12] SPICE is capable of transmitting full-frame data in both compressed and uncompressed format. In this case the L1 file contains *four* HDUs, *two* for each detector array, one with decompressed data and one with the uncompressed data.

Sun. The plate scale in the dispersion dimension is adjusted in order to get the same spatial plate scale in the dispersion and y dimensions for wide-slit (and dumbbell) observations.

Note that if telemetry is lost for compressed observations, entire image planes or even full data cubes may be missing (i.e. are set to `NaN`) in L2 files, see Sections 4.4.4.4 and 4.4.5. The keywords describing the telemetry and any loss of telemetry are listed in Sections 4.4.1.3.5.1 and 4.4.1.3.5.2.

Keywords that are present in L1 files are updated in order to account for e.g. calibrated data array values and merged windows.

Some keywords are modified as required by [FITSpaper]. As an example, the `BLANK` keyword only applies to integer data HDUs, as is the case for L1. The floating point data HDUs in L2 files do not have a `BLANK` keyword, and undefined pixels are set to `NaN`.

L2 files contain additional keywords not present in L1 files:

- Keywords that describe the calibration. This includes:
  - `VERS_CAL`, a Solar Orbiter-wide keyword giving the version of the calibration software
  - Additional `PRxxxxn` keywords describing the processing steps, see Section 4.4.2.1.
  - The variable keyword `RADCAL`, giving the radiometric calibration factor as a function of dispersion pixel, see Section 4.4.4.3
  - `WAVECOV` is a comma separated list of the wavelength coverage (i.e. '`WAVEMIN-WAVEMAX`') of all image HDUs in the file.
- Keywords that describe random and systematic errors (`CRDERi` and `CSYERi` respectively, see Table 3-8 of [METADATA])
- The keyword `UCD` giving the Unified Content Descriptors, e.g.:
  ```
  UCD = 'phot.radiance; em.line' ; for narrow-slit observations
  UCD = 'phot.radiance; em.UV'   ; for wide-slit/dumbbell observations
  ```
  The intensity unit of SPICE L2 files, given by `BUNIT`, is `W/m2/sr/nm` (narrow-slit) or `W/m2/sr` (wide-slit/dumbbell). The intensity unit is described by `BTYPE`, with the value '`Spectral Radiance`' or '`Radiance`' respectively.

For files with two or more merged windows the following keywords will have a different value in L2 than in L1:

- `EXTNAME` is based on the `EXTNAMEs` of the individual windows, ending with the string "`(Merged)`".
- `NWIN` is the total number of observational HDUs in the L2 file, i.e. `NWIN` is no longer the number of windows defined for the study in the MISO planning tool
- `WINNO` is the window number in the L2 file (starting at 0)
- `MISOWIN` and `WINTAB` will have the value of the first window in the group of merged windows.
- All keywords that depend on the width of the windows are updated to reflect the width of the concatenated window.

4.4.2.1 Keywords describing the individual L1 to L2 calibration steps

The individual L1 to L2 processing steps are described by the `PRxxxxn` keywords:

```
PRSTEP6 = 'DARK-SUBTRACTION'   / Type of processing, step 6
PRPROC6 = 'spice_prep_dark_offset_correction.pro' / Name of procedure, step 6
PRPVER6 = '1.10    '           / Version of procedure, step 6
PRPARA6 = 'dark="solo_L1_spice-n-exp_20210912T021124_V07_67109105-009.fits"' / P
PRLIB6A = 'uio-spice-pipeline' / Software library containing PRPROC6

PRSTEP7 = 'FLATFIELDING'       / Type of processing, step 7
PRPROC7 = 'spice_prep_flat_field_correction.pro' / Name of procedure, step 7
PRPVER7 = '1.4     '           / Version of procedure, step 7
PRPARA7 = 'flat="ground-calibration flat field"' / Parameters for PRPROC7
PRLIB7A = 'uio-spice-pipeline' / Software library containing PRPROC7

PRSTEP8 = 'BURN-IN-CORRECTION' / Type of processing, step 8
PRPROC8 = 'spice_prep_burnin_correction.pro' / Name of procedure, step 8
PRPVER8 = '1.2     '           / Version of procedure, step 8
PRPARA8 = 'burn_in_correction_data_version="2023-08-03",                       &'
CONTINUE  'burn_in_correction_defined_for_this_line=0                          &'
CONTINUE  '' / Parameters for PRPROC8
PRLIB8A = 'uio-spice-pipeline' / Software library containing PRPROC8

PRSTEP9 = 'SPATIAL-SPECTRAL-DISTORTION-CORRECTION' / Type of processing, step 9
PRPROC9 = 'spice_prep_distortion_correction.pro' / Name of procedure, step 9
PRPVER9 = '2.3     '           / Version of procedure, step 9
PRPARA9 = 'distortion_correction_matrix_version="2021-09-08"' / Parameters for P
PRLIB9A = 'uio-spice-pipeline' / Software library containing PRPROC9

PRSTEP10= 'RADIOMETRIC-CALIBRATION' / Type of processing, step 10
PRPROC10= 'spice_prep_radiometric_calibration.pro' / Name of procedure, step 10
PRPVER10= '2.0     '           / Version of procedure, step 10
PRPARA10= 'method="based on comparison to QS SUMER spectrum",                  &'
CONTINUE  'response_file_version="2023-10-11",                                 &'
CONTINUE  'interpolated_sw_response=0.48677482' / Parameters for PRPROC10
PRLIB10A= 'uio-spice-pipeline' / Software library containing PRPROC10

PRSTEP11= 'WINDOW-CONCATENATION' / Type of processing, step 11
PRPROC11= 'spice_prep.pro'     / Name of procedure, step 11
PPRPVER11= '4019    '           / Version of procedure, step 11
PRPARA11= '   WINNOs_of_concatenated_windows=[0,1],                            &'
CONTINUE  'WINTABIDs_of_concatenated_windows=[170,173],                        &'
CONTINUE  ' MISOWINs_of_concatenated_windows=[3057,2907],                      &'
CONTINUE  ' EXTNAMEs_of_concatenated_windows="O III 703 Peak,Mg IX 706 - Peak"&'
CONTINUE  '' / Parameters for PRPROC11
PRLIB11A= 'uio-spice-pipeline' / Software library containing PRPROC11
```

Note that all observational HDUs of L2 files will include a description of the burn-in correction processing step. However, the burn-in data file contains burn-in corrections only for sections of the detectors where strong lines have significantly reduced the detector's sensitivity. Whether the burn-in data file contains information that allowed for a correction of the HDU in question is given by the `PRPARA` parameter `burn_in_correction_defined_for_this_line`. In the header example above the value of this parameter is `0`, i.e. this line has *not* been corrected for burn-in. Finally, if one or more windows that make up a merged window are corrected for burn-in, the merged window is also marked as corrected for burn-in.

The sensitivity of the SPICE detectors decreases with time. The radiometric calibration includes a correction for this time-dependent responsivity. From a table of estimated response values we interpolate a single response correction value for each detector for the `DATE-AVG` of the observation. This factor is given by the `PRPARA` parameter `interpolated_sw_response`.

4.4.3 Data arrays and FITS headers: L3 R and L3 C

The pipeline is as of November 2023 creating simplified L3 R files containing gaussian fit parameters (see [IDLANA]). Each observational HDU in the input L2 file leads to multiple L3 HDUs. The number of L3 HDUs may be significantly reduced in the future when upcoming mechanisms for e.g. linking data arrays in external FITS files are implemented. For the time being, the L3 file consists of 7 HDUs for each L2 HDU:

1. The Gaussian fit parameters
2. L2 data
3. Wavelength for each pixel
4. Residuals of the fit compared to the original data for each pixel
5. Weights of each pixel
6. "Include" flags of each pixel and fit component
7. "Constant" flags of each pixel and fit parameter

These HDUs contain all the information that is needed to make the HDUs in L3 files fully self-consistent. It is therefore easy to manually modify the automatically determined initial values of the fit and re-run the line fitting with the code of the user's choice. The line fitting routine must use the parameterisation used by `cfit.pro` (see [CFIT]), but may of course be written in other programming languages than IDL. However, for IDL users manipulating L3 files is especially convenient:

– use `spice_xfiles` to select the file you want to analyse
– click the "copy window to user file" button in the "LEVEL 3 – official" column
– click the "View/Edit window" button in the "LEVEL 3 – user" column to start `xcfit_block`

To be implemented: based on the fitted Gaussian line parameters, secondary derived parameters can be estimated:

- Abundances
- FIP bias
- Density
- Temperature

The format of L3 files including secondary derived parameters is not yet finalised.

4.4.3.1 Keywords describing individual L2 to L3 processing steps

As outlined in Section 4.4.2.1 the individual processing steps are described using the `PRxxxxn` keywords. When processing L2 files to L3 additional processing steps are described: finding the peak(s) in the spectrum to which Gaussian fits were made, and the actual line fitting:

```
PRSTEP11= 'PEAK-FINDING'        /  Processing step type
PRPROC11= 'spice_gt_peaks'      /  Name of procedure performing PRSTEP11
PRPVER11=                     1 /  Version of procedure PRPROC11
PRLIB11A= 'solarsoft/so/spice/idl/quicklook' /  Software library containing PRPR
```

```
PRSTEP12= 'LINE-FITTING'       /  Processing step type
PRPROC12= 'spice_data::create_l3_file' /  Name of procedure performing PRSTEP12
PRPVER12=                    5 /  Version of procedure PRPROC12
PRPARA12= 'LINE_LIST = 0,                                                  &'
CONTINUE  'MASKING   = 1,                                                  &'
CONTINUE  'FITTING   = 1,                                                  &'
CONTINUE  'POSITION  = 0,                                                  &'
CONTINUE  'VELOCITY  = -999,                                               &'
CONTINUE  'POSSIBLE_MANUAL_EDITING = 0' /  Parameters for PRPROC12
PRLIB12A= 'solarsoft/so/spice/idl/quicklook' /  Software library containing PRPR
```

In future L3 files, the calculation of secondary derived data products will be described by e.g.:

```
PRSTEP9 = 'ABUNDANCE-ESTIMATION'  / Processing step type 9
PRPVER9 =                    0.32 / Version of procedure PRPROC9
PRPROC9 = 'spice_abundances.pro'  / Name of procedure performing PRSTEP9

PRSTEP10= 'FIP-BIAS-ESTIMATION'   / Processing step type 10
PRPVER10=                     1.0 / Version of procedure PRPROC10
PRPROC10= 'spice_fip_bias.pro'    / Name of procedure performing PRSTEP10

PRSTEP11= 'DENSITY-ESTIMATION'    / Processing step type 11
PRPVER11=                       2 / Version of procedure PRPROC11
PRPROC11= 'spice_densities.pro'   / Name of procedure performing PRSTEP11

PRSTEP12= 'TEMPERATURE-ESTIMATION'/ Processing step type 12
PRPVER12=                     0.1 / Version of procedure PRPROC12
PRPROC12= 'spice_temperatures.pro'/ Name of procedure performing PRSTEP12
```

4.4.4 Storing variable keyword values in binary table extensions

Some SPICE FITS keywords may have multiple values. Such keywords have either:

- one value per exposure: acquisition times, mirror positions, and temperatures, see Section 4.4.4.1, or
- one value per segment: voltages, see Section 4.4.4.2, or
- one value per dispersion pixel: the radiometric calibration factor that converts the intensities in Level 2 files to counts, see Section 4.4.4.3, or
- L1 only: values that are not directly linked to individual exposures or other dimensions of the data cube: lost telemetry packet indices and lost FFT Bin indices, see Section 4.4.4.4.

In all cases we use the *variable-keyword mechanism* outlined in Appendix I of [S-META] (see http://sdc.uio.no/open/solarnet/ for the latest version) to store the individual values of variable keywords in binary table extensions.

4.4.4.1 Variable keywords with one value per exposure: times, mirror positions, and temperatures

The acquisition time (OBT), scan mirror position and 4 instrument temperatures are recorded for each exposure of a SPICE observation. For single exposure observations, the values of these measurements are stored in the FITS keywords `TIMAQOBT`, `MIRRPOS`, `TN_FOCUS`, `TN_MIRR`, `TN_SW`, and `TN_LW`, and the temperatures converted from data numbers to degrees Celsius in `T_FOCUS`, `T_MIRR`, `T_SW`, and `T_LW`. In L1+ FITS files, `TIMAQUTC` stores the acquisition time converted to UTC. For multi exposure observations, these keywords hold the *average* values, and the *individual* *values* of each keyword, i.e. one value per exposure, are stored in a binary table extension using the variable-keyword mechanism.

In the header of a SPICE observational HDU that uses the variable-keyword mechanism, the `VAR_KEYS` keyword always have the following value:

```
VAR_KEYS= 'VARIABLE_KEYWORDS;TIMAQOBT,MIRRPOS,TN_FOCUS,TN_MIRR,TN_SW,TN_LW',&'
CONTINUE  'T_FOCUS, T_MIRR, T_SW, T_LW, TIMAQUTC'/ Variable keywords
```

This means that in the binary table extension `VARIABLE_KEYWORDS`, the individual values of the keywords `TIMAQOBT, MIRRPOS, TN_xxxxx, T_xxxxx,` and `TIMAQUTC` are stored in columns with `TTYPEn` equal to the keyword names. Following Appendix I-d of [S-META], the binary table columns and the referring HDU are to be associated pixel-by-pixel, since the `WCSNn` keywords are set to '`PIXEL-TO-PIXEL`'. The dimensions of the value columns, given by `TDIMn`, are `(1,1,1,NAXIS4)`, the singular dimensions signalling that there is one value per exposure that is valid for all `(x,y,d)` pixels for that exposure.

Below is an excerpt of such a binary table extension header, including column specific keywords that define the columns storing the individual `TIMAQOBT` and `TN_FOCUS` values.

```
XTENSION= 'BINTABLE'           / Written by IDL:  Mon Sep 25 12:03:41 2017
:
EXTNAME = 'VARIABLE_KEYWORDS'  / Extension name
:
WCSN1   = 'PIXEL-TO-PIXEL'     / Value column/referring HDU association type
TFORM1  = '64D      '          / Real*8 (double precision)
TTYPE1  = 'TIMAQOBT'           / [OBT] Start time of data acquisition
TDIM1   = '(1,1,1,64)'         / Array dimensions for column 1
TUNIT1  = '         '          / Units of column 1
TDMIN1  =        481295089.350 / Minimum value in column 1
TDMAX1  =        481295146.051 / Maximum value in column 1
TDESC1  = 'Variable values for TIMAQOBT' / Axis labels for column 1
:
WCSN3   = 'PIXEL-TO-PIXEL'     / Value column/referring HDU association type
TFORM3  = '64I      '          / Unsigned Integer*2 (short integer)
TTYPE3  = 'TN_FOCUS'           / [adu] SFM focus adu temperature
TDIM3   = '(1,1,1,64)'         / Array dimensions for column 3
TUNIT1  = '         '          / Units of column 3
TSCAL3  =                    1 / Scale parameter for column 3
TZERO3  =                32768 / Zero offset for column 3
TDMIN3  =                  846 / Minimum value in column 3
TDMAX3  =                  871 / Maximum value in column 3
TDESC3  = 'Variable values for TN_FOCUS' / Axis labels for column 3
```

4.4.4.2 Variable keywords with one value per segment: voltages

4 instrument voltages per segment are downlinked in the science telemetry. For single-segment observations the raw values of these measurements are stored in the FITS keywords `VN_MCPSW`, `VN_MCPLW`, `VN_GAPSW`, and `VN_GAPLW`, and the voltages converted from data numbers to Volt in `V_MCPSW`, `V_MCPLW`, `V_GAPSW`, and `V_GAPLW`. For multi-segment observations these keywords hold the *average* values, and the *individual* values for each keyword, i.e. one value per segment, are stored in the same binary table extension described in the previous Section.

In the header of a SPICE observational HDU stemming from a multi-segment observation, the `VAR_KEYS` keyword have the following value:

```
VAR_KEYS= 'VARIABLE_KEYWORDS;TIMAQOBT,MIRRPOS,TN_FOCUS,TN_GRAT,TN_SW,TN_LW,T_F&'
CONTINUE  'OCUS,T_GRAT,T_SW,T_LW,VN_MCPSW,VN_MCPLW,VN_GAPSW,VN_GAPLW,V_MCPSW,V&'
CONTINUE  '_MCPLW,V_GAPSW,V_GAPLW,TIMAQUTC' / Variable keywords
```

Below is an excerpt of the header of the `'VARIABLE_KEYWORDS'` binary table extension, including column specific keywords that define the column storing the individual `VN_MCPSW` values, one value for each segment.

```
XTENSION= 'BINTABLE'            / Written by IDL:  Fri Oct 25 13:47:08 2019
:
EXTNAME = 'VARIABLE_KEYWORDS'   / Extension name
:
           -------------------------------
           | Column 12 specific keywords |
           -------------------------------
WCSN12  = 'PIXEL-TO-PIXEL'      / Value column/referring HDU association type
TFORM12 = '4I      '            / Unsigned Integer*2 (short integer)
TTYPE12 = 'VN_MCPLW'            / [adu] MCP LW voltage
TDIM12  = '(1,1,1,4)'           / Array dimensions for column 12
TUNIT12 = '        '            / Units of column 12
TSCAL12 =                    1 / Scale parameter for column 12
TZERO12 =                32768 / Zero offset for column 12
TDMIN12 =                  553 / Minimum value in column 12
TDMAX12 =                  553 / Maximum value in column 12
TDESC12 = 'Variable values for VN_MCPLW' / Axis labels for column 12
```

4.4.4.3 Variable keyword with one value per lambda pixel: radiometric calibration factor

In the Level 1 to Level 2 calibration the photon counts are converted to physical intensity units. In order to get back the counts from the Level 2 intensity the variable keyword `RADCAL` should be applied to the Level 2 data.

In the header of a Level 2 observational HDU the `RADCAL` keyword gives the mean radiometric conversion factor of the window. The individual values of the conversion factor are stored in the binary table extension `VARIABLE_KEYWORDS`. To retrieve the counts the Level 2 data array is multiplied with the conversion factor array values:

```
FOR lam=0,n_lam-1 DO counts[*,*,lam,*] = L2_int[*,*,lam,*]*radcal_array[lam]
```

The binary table column storing the individual `RADCAL` values for a given observational HDU is named `RADCAL` followed by a tag enclosed by square brackets. This tag equals the `EXTNAME` of the observational HDU, possibly shortened in order to make the length of the string `RADCAL[tag]` shorter than 68 characters.

As an example, the value of the `VAR_KEYS` keyword of an observational HDU may be:

```
VAR_KEYS='VARIABLE_KEYWORDS;TIMAQOBT,MIRRPOS,TN_FOCUS,TN_GRAT,TN_SW,TN_LW,T_F&'
CONTINUE'OCUS,T_GRAT,T_SW,T_LW,VN_MCPSW,VN_MCPLW,VN_GAPSW,VN_GAPLW,V_MCPSW,V&'
CONTINUE  '_MCPLW,V_GAPSW,V_GAPLW,TIMAQUTC,CRDER1,CRDER2,RADCAL[O III 703 / Mg&'
```

```
CONTINUE  ' IX 706 (Merged)]
```

The relevant part of the header of the binary table extension 'VARIABLE_KEYWORDS' may then be:

```
WCSN22  = 'PIXEL-TO-PIXEL'      / Value column/referring HDU association type
TFORM22 = '80D     '            / Real*8 (double precision)
TTYPE22 = 'RADCAL[O III 703 / Mg IX 706 (Merged)]' / [DN/(W/m2/sr/nm)] Calibrati
TDIM22  = '(1,1,80,1)'          / Array dimensions for column 22
TUNIT22 = 'DN/(W/m2/sr/nm)]'    / Units of column 22
TDMIN22 =         335.757886613 / Minimum value in column 22
TDMAX22 =         367.519855635 / Maximum value in column 22
TDESC22 = 'Variable values for RADCAL' / Axis labels for column 2
```

4.4.4.4 Variable keywords describing lost telemetry packets and lost FFT Bins

In L1 files, if telemetry packets containing observational data are lost the packet indices (starting at 0 for the first data packet) are stored in the binary table extension with name 'LOST_TELEMETRY'. Note that this extension is not present in L2 and L3 FITS files. Instead, the user should refer to the keywords described in Sections 4.4.1.3.5.1 and 4.4.1.3.5.2 for an overview of the telemetry completeness.

The name of the binary table column storing the indices of the lost packets is 'LOSTPKTS', followed by a tag giving the EXTNAME of the observational HDU that uses the variable-keyword mechanism (the referring HDU). The lost packet indices are stored primarily for pipeline debugging purposes.

Below is an excerpt of the header of a 'LOST_TELEMETRY' binary table extension of a L1 file, including keywords that define the column storing the individual lost packet indices:

```
XTENSION= 'BINTABLE'            / Written by IDL:  Mon Sep  9 11:08:01 2019
:
EXTNAME = 'LOST_TELEMETRY'      / Extension name
:
TFORM1  = '3J      '            / Integer*4 (long integer)
TTYPE1  = 'LOSTPKTS[WINDOW0_724.05A]' / Lost packets w/data, variable keyword
TUNIT1  = '        '            / Units of column 1
TDMIN1  =                     1 / Minimum value in column 1
TDMAX1  =                    28 / Maximum value in column 1
TDESC1  = 'Indices of lost packets containing observational data' / Axis labels
```

The value of the VAR_KEYS keyword of the referring HDU contains the binary table extension name and the column name:

```
VAR_KEYS= 'VARIABLE_KEYWORDS;TIMAQOBT,MIRRPOS,TN_FOCUS,TN_GRAT,TN_SW,TN_LW,T_F&'
CONTINUE  'OCUS,T_GRAT,T_SW,T_LW,TIMAQUTC,                                    &'
CONTINUE  'LOST_TELEMETRY;LOSTPKTS[WINDOW0_724.05A]' / Variable keywords
```

As described in [SPICELOST] the implications of lost telemetry packets in the case of SHC data highly depends on how many and which FFT Bins are lost. This information is therefore stored in the FITS files in order to help the advanced user to determine the degree of degradation of the approximated data cube. An additional binary table column storing the lost FFT Bin indices is added to the 'LOST_TELEMETRY' binary table extension. The name of this column is 'LOSTBINS' plus a tag with the EXTNAME of the referring HDU.

L1 files may contain data cubes that are reconstructed from incomplete SHC telemetry, but in the L2 and L3 files produced by the Science Data Pipeline all pixels in such cubes are set to NaN. However, the user may choose to keep incomplete data cubes when running the L1 to L2 calibration routines manually and use the information in the LOSTBINS column when interpreting the data.

Below is an excerpt of the header of a 'LOST_TELEMETRY' binary table extension, including keywords that define the column storing the individual lost FFT Bin indices:

```
XTENSION= 'BINTABLE'            / Written by IDL:  Mon Sep  9 11:08:01 2019
:
EXTNAME = 'LOST_TELEMETRY'      / Extension name
:
TFORM2  = '8I      '            / Integer*2 (short integer)
TTYPE2  = 'LOSTBINS[WINDOW0_724.05A]' / Lost FFT bins, variable keyword
TUNIT2  = '        '            / Units of column 2
TDMIN2  =                    0 / Minimum value in column 2
TDMAX2  =                   31 / Maximum value in column 2
TDESC2  = 'Indices of lost FFT Bins' / Axis labels for column 2
```

The VAR_KEYS keyword of the referring HDU contains the binary table extension name and the column name:

```
VAR_KEYS= 'VARIABLE_KEYWORDS;TIMAQOBT,MIRRPOS,TN_FOCUS,TN_GRAT,TN_SW,TN_LW,T_F&'
CONTINUE  'OCUS,T_GRAT,T_SW,T_LW,TIMAQUTC,                                    &'
CONTINUE  'LOST_TELEMETRY;LOSTPKTS[WINDOW0_724.05A],LOSTBINS[WINDOW0_724.05A]&'
/
CONTINUE  '' / Variable keywords
```

The referring HDU contains representative scalar values of variable keyword (see Appendix I of [S-META]). The representative scalar values of LOSTPKTS and LOSTBINS are the number of lost telemetry packets and the number of lost FFT Bins respectively:

```
LOSTPKTS =                     3 / Number of lost packets w/data, variable
keyword
LOSTBINS =                     8 / Number of lost FFT bins, variable keyword
```

4.4.5 Flagging image planes with approximated or missing values due to lost compressed telemetry using binary table pixel list

As described in [SPICELOST] it is possible to reconstruct a data cube in the case of lost telemetry, even if the data is compressed. However, one or more X-Y or dispersion-Y image planes of the data cube may in such cases be lost or have approximated values, depending on the file level, which kind of compression was applied, and which telemetry packets were lost.

If a lost telemetry packet contained JPEG header information it is not possible to decompress the JPEG image, and the entire image plane is lost. Decompression is possible if a missing telemetry packet did *not* contain JPEG header information, but the resulting values will be approximated. In such cases L1 files contain the approximated values, but in L2 and L3 files all pixels affected by telemetry loss are set to `NaN`.

If compressed telemetry packets are missing from a multi-segment observation, then *ranges* of image planes may be set to `NaN`, or have approximated values (L1 only). Each image plane range corresponds to an image plane of a single segment. Flagging of image plane ranges may also be applied for compressed full-detector readouts. For such observations each detector array is split into 16 Lambda-Y regions of 64x1024 pixels that are JPEG compressed on-board separately.

Note that for technical reasons the flagging of lost image plane ranges in L1 files described below currently does *not* work if *all* packets of a segment are missing. However, the L1+ FITS keywords mentioned in Sections 4.4.1.3.5.1 and 4.4.1.3.5.2 are correctly set even when entire segments are missing.

In L1 files, we use the pixel list mechanism described in Appendix II-a of [S-META] to flag image planes ranges that have approximated values due to missing compressed telemetry.

For every observational HDU that uses the pixel list mechanism there is a corresponding binary table extension containing a single pixel list where the vertices of approximated or lost image plane ranges are stored. The name of these binary table extensions are `'APRXPLNPIXLIST'` and `'LOSTPLNPIXLIST'` respectively, plus a tag with the `EXTNAME` of the referring HDU. In L2 and L3 files, pixels corresponding to L1 Image plane ranges that are defined in an `'APRXPLNPIXLIST'` extension are set to `NaN`. In the future the advanced user will be able to run the Level 1 to Level 2 calibration routines manually with an option to retain approximated pixel values in L2 files.

In the header of the L1 observational HDU the names of the any pixel list binary table extensions are given by `PIXLISTS`. `NLOSTPLN` and `NAPRXPLN` give the total number of lost or approximated image plane ranges in the L1 data cube.

As an example, we consider an actual full-detector JPEG compressed observation obtained in March 2021. A telemetry packet belonging to the third Lambda-Y JPEG of the Long Wavelength detector never made it to the ground, leading to approximated pixel values. The corresponding range of the full LW Lambda-Y image plane is flagged by the pixel list binary table by the following two rows:

	DIMENSION1	DIMENSION2	DIMENSION3	DIMENSION4	PIXTYPE
Row 1	1	1	65	1	1
Row 2	1	1024	128	1	2

An excerpt of the header of the L1 pixel list binary table extension looks like this:

```
XTENSION= 'BINTABLE'           / Written by IDL:  Wed Sep 29 13:20:51 2021
…
TFIELDS =                    5 / Number of columns
EXTNAME = 'APRXPLNPIXLIST[Full LW 4:1 Focal Lossy]' / Extension name
```

```
                -------------------------------
                | Column 1 specific keywords |
                -------------------------------
TFORM1  = '1I      '           / Integer*2 (short integer)
TTYPE1  = 'DIMENSION1'         / Pixel indices dimension 1
TCTYP1  = 'PIXEL   '           / Indicates that column 1 is a pixel index
TPC1_1  =                    1 / Indicates that column 1 is a pixel index
TUNIT1  = '        '           / Units of column 1
TDMIN1  =                    1 / Minimum value in column 1
TDMAX1  =                    1 / Maximum value in column 1
TDESC1  = 'Lower Left/Upper Right pixel indices of 1 approximated Lambda-Y ima&'
CONTINUE  'ge plane ranges due to loss of compressed telemetry packets&' / Axis
CONTINUE  '' / labels for column 1


                -------------------------------
                | Column 2 specific keywords |
                -------------------------------
TFORM2  = '1I      '           / Integer*2 (short integer)
TTYPE2  = 'DIMENSION2'         / Pixel indices dimension 2
TCTYP2  = 'PIXEL   '           / Indicates that column 2 is a pixel index
TPC2_2  =                    1 / Indicates that column 2 is a pixel index
TUNIT2  = '        '           / Units of column 2
TDMIN2  =                    1 / Minimum value in column 2
TDMAX2  =                 1024 / Maximum value in column 2
TDESC2  = 'Lower Left/Upper Right pixel indices of 1 approximated Lambda-Y ima&'
CONTINUE  'ge plane ranges due to loss of compressed telemetry packets&' / Axis
CONTINUE  '' / labels for column 2


                -------------------------------
                | Column 3 specific keywords |
                -------------------------------
TFORM3  = '1I      '           / Integer*2 (short integer)
TTYPE3  = 'DIMENSION3'         / Pixel indices dimension 3
TCTYP3  = 'PIXEL   '           / Indicates that column 3 is a pixel index
TPC3_3  =                    1 / Indicates that column 3 is a pixel index
TUNIT3  = '        '           / Units of column 3
TDMIN3  =                   65 / Minimum value in column 3
TDMAX3  =                  128 / Maximum value in column 3
TDESC3  = 'Lower Left/Upper Right pixel indices of 1 approximated Lambda-Y ima&'
CONTINUE  'ge plane ranges due to loss of compressed telemetry packets&' / Axis
CONTINUE  '' / labels for column 3


                -------------------------------
                | Column 4 specific keywords |
                -------------------------------
TFORM4  = '1I      '           / Integer*2 (short integer)
TTYPE4  = 'DIMENSION4'         / Pixel indices dimension 4
TCTYP4  = 'PIXEL   '           / Indicates that column 4 is a pixel index
TPC4_4  =                    1 / Indicates that column 4 is a pixel index
TUNIT4  = '        '           / Units of column 4
TDMIN4  =                    1 / Minimum value in column 4
```

```
TDMAX4  =                    1 / Maximum value in column 4
TDESC4  = 'Lower Left/Upper Right pixel indices of 1 approximated Lambda-Y ima&'
CONTINUE  'ge plane ranges due to loss of compressed telemetry packets&' / Axis
CONTINUE  '' / labels for column 4


           ------------------------------
           | Column 5 specific keywords |
           ------------------------------
TFORM5  = '1I      '           / Integer*2 (short integer)
TTYPE5  = 'PIXTYPE '           / Pixel type
TUNIT5  = '        '           / Units of column 5
TDMIN5  =                    1 / Minimum value in column 5
TDMAX5  =                    2 / Maximum value in column 5
TDESC5  = 'Pixel indices types: 1 = lower left corner indices, 2 = upper right&'
CONTINUE  ' corner indices' / Axis labels for column 5


           ------------------------------
           | File and study identifiers |
           ------------------------------
FILENAME= 'solo_L2_spice-n-exp_20210322T033925_V06_50331861-018.fits' / Filename
…
```

Note that for SHC data *all* image planes stemming from a single segment are approximated if telemetry packets with observational data are lost. The amount of approximation may be estimated using the variable L1 keyword `LOSTBINS` described in Section 4.4.4.4.

APPENDIX A SPICE DATA PRODUCTS MATRIX

A FITS file may contain different types of HDUs/data products. Each HDU type may occur in files with a number of different combinations of descriptor elements.

L1 and L2:
In Table 4-8 we have listed all possible L1 and L2 data products, and the file descriptors that a FITS file containing such a product may have. The overwhelming majority of Science observations recorded the first 3 years of the mission are `n-ras`, with a few `n-sit` and fewer still `w-sit`. For all observation types including calibration observations, `n-exp` is the most common descriptor. The usage of `-int` and `-db` is negligible. See Table 4-2 for actual usage of the different data products.

Data Product (HDU type)	Description	Descriptors	Avg cadence	Expected Daily Vol
Narrow-slit spectral-profile raster		n-ras[-db][-int]		
Narrow-slit intensity-window raster		n-ras[-db]-int		
Wide-slit raster		w-ras		
Dumbbell raster		n-ras-db[-int]		
Narrow-slit spectral-profile sit-and-stare		n-sit[-db]		
Wide-slit sit-and-stare		w-sit		
Dumbbell sit-and-stare		n-sit-db		
Narrow-slit full detector single exposure		n-exp		
Wide-slit full detector single exposure		w-exp		

Table 4-8: SPICE L1 and L2 Data Products Matrix

L3 R and L3 C:
In Table 4-9 we have listed all possible L3 R and L3 C data products, and the file descriptors that a FITS file containing such a product may have.

Data Product (HDU type)	Description	Descriptors	Avg cadence	Expected Daily Vol
Narrow-slit spectral-profile raster		n-ras[-db][-int]		
Narrow-slit spectral-profile sit-and-stare		n-sit[-db]		

Table 4-9: SPICE L3 R and L3 C Data Products Matrix

APPENDIX B RELEATIONSHIP BETWEEN SCIENCE HEADER PACKET PARAMETERS AND FITS KEYWORDS

Section 4.2.6.1 of [DATAICD] lists the Science Header Packet parameters that describe the collection of Science Data Packets that builds up a raster segment. These parameters may apply to all windows of an observation, a single window, a single raster segment, or the parameters may have one value per exposure. The values of these parameters are stored as FITS keywords, see Table 4-10. Note that the "StudyFlags" parameter is an 8-bit integer with the value of each bit indicating different instrument settings, see Table 4-11. Both the primary HDU and all image extensions store all the FITS keywords described in the table. The shaded table rows indicate parameters/keywords with one value for each exposure. The mean values of these keywords are stored in the primary HDU and all image extensions, and the individual values for each exposure are stored in a binary table extension, see Section 4.4.4.

Science Header Packet Parameters	FITS keywords
Total number of CCSDS packets for this window	NPACKETS -1 or -2 (NPACKETS is the number of packets with observational data, excluding the Header Packet and the Final Packet if the latter only contains a checksum)
Observation ID	SPIOBSID
Focus position	FOCUSPOS
Slit position	SLIT_ID
Exposure time	XPOSURE * 10
Study ID	STUDY_ID
StudyFlags	STUDYFLG
Total Raster Segments per Window	NSEGMENT
ObsRasterNumber	RASTERNO
Raster Segment Number	-
Window Total Number	NWIN
Window Number	WINNO
Window Data Table ID	WINTABID
Window Start Column	PXBEG3
Window Width	PXEND3 – PXBEG3 + 1
Wavelength Binning Factor	NBIN3
Window Starting Row	PXBEG2
Window Height	PXEND2 – PXBEG2 + 1
Y Binning Factor	NBIN2
Compression Type	COMPTYPE
Compression Amount Parameter	COMPPARA
SHC FFT ID	SHCFFTID
Pixel Level Offset	PIXELOFF
Alignment Window Status	DUMBBELL
MCP SW Monitor Voltage	VN_MCPSW (and V_MCPSW)
MCP LW Monitor Voltage	VN_MCPLW (and V_MCPLW)

Science Header Packet Parameters	FITS keywords
Gap SW Monitor Voltage	`VN_GAPSW` (and `V_GAPSW`)
Gap LW Monitor Voltage	`VN_GAPLW` (and `V_GAPLW`)
Segment X Size	–
Acquisition Time when initiating black level reset after arriving at X	`TIMAQOBT`
Scan Mirror Position	`MIRRPOS`
Temperature 1	`TN_FOCUS` (and `T_FOCUS`)
Temperature 2	`TN_GRAT` (and `T_GRAT`)
Temperature 3	`TN_SW` (and `T_SW`)
Temperature 4	`TN_LW` (and `T_LW`)

Table 4-10: Telemetry Science Header Packet Parameters and their keyword equivalents. Note that the two parameters that describe a single raster segment do not have FITS keyword equivalents. Keywords in red are Solar Orbiter-wide FITS keywords. The values of the individual bits of `STUDYFLG` determine the values of the 6 derived keywords in Table 4-11. The voltages and temperatures are given in engineering units in the `VN_xxxxx` and `TN_xxxxx` keywords, and are converted to Volt and Celsius in the `V_xxxxx` and `T_xxxxx` keywords. Orange shading indicates parameters with one value per segment, blue shading indicates parameters with one value per exposure. In L1+ FITS files the onboard time `TIMAQOBT` is converted to UTC and given in `TIMAQUTC` (this keyword is not present LL01 FITS files).

Bit	Study Flag Description	FITS Keywords
0 -1	*Spare*	–
2	AlignExcludeSpectral – 0 = Do not exclude spectral window data – 1 = Exclude spectral window data	`NOSPECTR`
3	Cal Mode Config – 0 = 1 Data Plane – 1 = 2 Data Planes	`CALMODE`
4	Double Exposure Number – 0 = First Exposure – 1 = Second Exposure	`DBLEXPNO`
5	Double Exposure Enabled – 0 = False – 1 = True	`DBLEXP`
6	Dark Map Subtraction Used? – 0 = False – 1 = True	`DARKMAP`
7	Black Level Subtraction Used – 0 = False – 1 = True	`BLACKLEV`

Table 4-11: FITS keywords derived from the value of `STUDYFLG`. All FITS keywords have values `0` or `1`.

APPENDIX C RELEATIONSHIP BETWEEN STUDY SETS, IORS AND FITS KEYWORDS

IAS provide Study Set files containing the definitions of all on-board studies, and IORs[13] containing the commanded parameters of each observation (i.e. each instance of a study). These files are used by the Science Data Pipeline in order to populate L1+ FITS keywords.

For each observation the SPICE Observation ID found in the telemetry (SPIOBSID, see Appendix B) is used to find the IOR containing the information about that specific observation. The IOR contains information about which Study Set was used by the planning tool in creating the IOR (the Study Set **<version>**, FITS keyword SETVER). Having found the correct Study Set file (FITS keyword SETFILE), the STUDY_ID of the telemetry is then used to find the Study Set's definition of the study in question.

If a parameter found in the IOR or Study Set is also found in the telemetry, the telemetry value is the one used as a FITS keyword value. This is to ensure that the metadata describe the actual contents of file, and not what was commanded.

The following FITS keywords are set based on XML tags, parameters values and XML comments in the IORs:

STP – Solar Orbiter Short-Term Plan number
WINSHIFT – The number of pixels the window is shifted towards the red on the detector relative to the base position of windows with the current MISOID
OBS_ID – SOC Observation ID (not to be confused with the SPICE Observation ID, SPIOBSID)
OBS_TYPE – Unique code for OBS_MODE/STUDY, derived from OBSID
SOOPTYPE – Unique code for SOOPNAME, derived from OBSID
SOOPNAME – The name of the SOOP (given in XML comments in the IORs)
NRASTERS – The number of planned rasters for a given SPIOBSID

Table 4-12 below summarises the Study Set XML tags describing a study and their translation into FITS keywords.

XML tags	XML tag data type and values	FITS keywords	FITS keyword data type and values	In TLM
<studySet>				
<version>	0 ≤ int ≤ ?	SETVER	*Same as XML tag*	No
<studyInfo>				
<groundId>	0 ≤ int ≤ ?	MISOSTUD	*Same as XML tag*	No
<name>	String, free text	STUDY and OBS_MODE	*Same as XML tag*	No
<purpose>	String, {"Calibration", "Science", "Engineering"}	PURPOSE	*Same as XML tag*	No

[13] Instrument Operation Requests

XML tags	XML tag data type and values	FITS keywords	FITS keyword data type and values	In TLM
<description>	String, free text	STUDYDES	*Same as XML tag*	No
<author>	String, free text	AUTHOR	*Same as XML tag*	No
</studyInfo>				
<onboardInfo>				
<onboardID>	0 ≤ int ≤ 63	STUDY_ID	*Same as XML tag*	Yes
</onboardInfo>				
<type>	String, {"Full Spectrum", "Spatial Scan", "Time Series", "Scanned Time Series"}	A mix of STUDYTYP and WIN_TYPE. These two keywords are populated using information in the telemetry.	STUDYTYP, {"Sit-and-stare", "Raster", "Single Exposure"} WIN_TYPE, {"Narrow-slit Spectral", "Dumbbell (lower)", "Dumbbell (upper)", "Wide-slit", "Intensity-window", "Full Detector Narrow-slit", "Full Detector Wide-Slit"}	Yes, implicitly
<slit>	1 ≤ int ≤ 4	SLIT_ID	*Same as XML tag*	Yes
<exposures>	1 ≤ int ≤ 480 [14]	CDELT1 or CDELT4 (when **<exposures>** eq actual number of exposures)	*Same as XML tag* (when **<exposures>** eq actual number of exposures)	Yes
<expTime>	0.1 ≤ float ≤ 1023.5 [15]	XPOSURE	*Same as XML tag*	Yes
<totalExpTime>		XPOSURE for double exposure	*Same as XML tag*	No
<xStart>	float, default 0	XSTART. Also used when calculating CRVAL1	*Same as XML tag*	No
<stepSize>	float, default ?	CDELT1	*Same as XML tag*	No

[14] Any number from 1 to 64, multiples of 2 between 66 and 128, multiples of 4 between 132 and 256, multiples of 8 between 264 and 480

[15] Overheads for resetting detector pixels are added to these numbers. The overhead scales with the number of pixels in the dispersion dimension. Resolution of expTime is 0.1 s between 0.1 s and 204.7 s, 0.5 s between 205 s and 1023.5 s.

XML tags	XML tag data type and values	FITS keywords	FITS keyword data type and values	In TLM
<delayTime>	float, default ?	*None*	–	No
<alignmentWindowConfig>				
<excludeSpectral>	boolean, default `false`	`NOSPECTR`	Int, {0,1}	Yes
<compressionUpper>				
<type>	String, {"`Spatial Lossy`", "`Focal Lossy`", "`Spatial Uncompressed`", "`Focal Uncompressed`", "`SHC Low`", "`SHC High`"}	`COMPRESS`	*Same as XML tag*	Yes, implicitly
<rawFactor>	`1` ≤ Int ≤ `256?`	`COMP_RAT` = raw data volume/compressed volume. Applies to all observations (not only JPEG compressed)	Int, {1,128}	Yes, implicitly
<shcCoef>	`0` ≤ Int ≤ `7`	`SHCFFTID`	*Same as XML tag*	Yes
</compressionUpper>				
<compressionLower>				
		See <compressionUpper>		
</compressionLower>				
</alignmentWindowConfig>				
<options>				
<readoutMode>	string, {"`Destructive`", "`Non Destructive`"}	`READMODE`	*Same as XML tag*	No
<calMode>	string, {"`Uncompressed`", "`Compressed`", "`Both`"}	`CALMODE`	Int, `{0,1}` (`0` = one data plane (compressed or uncompressed), `1` = two data planes (compressed ***and*** uncompressed)	Yes
<focusCal>		*None*	–	No
<macroSteps>		*None*	–	No
<macroDelay>		*None*	–	No
<darkMap>	boolean	`DARKMAP`	Int, {0,1}	Yes

XML tags	XML tag data type and values	FITS keywords	FITS keyword data type and values	In TLM
</options>				
<window>				
<windowInfo>				
<description>	string	EXTNAME. For merged windows the EXTNAME is based on the EXTANMEs of the constituent windows. For dumbbells, either "DUMBBELL_UPPER_" or "DUMBBELL_LOWER_" is prepended to the description string.	*Same as XML tag*	No
<wavelength>		*None* (keywords describing wavelength, e.g. WAVEMIN, WAVEMAX) are calculated from telemetry parameters)	–	Yes. implicitly
<onboardID>	0 ≤ int ≤ 255	WINTABID	*Same as XML tag*	Yes
<groundID>	0 ≤ int ≤ ?	MISOWIN	*Same as XML tag*	No
<iwin_type>	String, {"", "Background", "Line"}	IWIN_TYPE	*Same as XML tag*	
<background_number>	0 ≤ int ≤ 255	Used to determine IWINBKG	EXTNAME of extension storing background intensity-window	
<line_number>	0 ≤ int ≤ 255	Used to determine IWINLINE	(Comma-separated list of) EXTNAME(s) of extension(s) storing line intensity-window	
</windowInfo>				
<axis axisName="lambda">				
<start>		PXBEG3		Yes
<size>		PXEND3 – PXBEG3 + 1		Yes
<binFactor>		NBIN3		Yes
</axis>				
<axis axisName="Y">				
<start>		PXBEG2		Yes

XML tags	XML tag data type and values	FITS keywords	FITS keyword data type and values	In TLM
<size>		PXEND2 – PXBEG2 + 1		Yes
<binFactor>		NBIN2		Yes
</axis>				
<compression>				
		See <compressionUpper>		
</compression>				
<alignmentWindow>		DUMBBEXT *Not implemented yet*	(Comma-separated list of) EXTNAME(s) of the extension(s) containing dumb-bell data for this window	No
</window>				

Table 4-12: Relationship between Study Generator XML tags and FITS keywords.