1 INTRODUCTION

This guide describes the data produced by the SOFIA/FIFI-LS data reduction pipeline (REDUX) for guest investigators. The FIFI-LS observing modes are described in the SOFIA Observer’s Handbooks, available from the Proposing and Observing* page on the SOFIA website.

2 SI OBSERVING MODES SUPPORTED

2.1 FIFI-LS instrument information

FIFI-LS has two separate and independent grating spectrometers with common fore-optics feeding two large Ge:Ga detector arrays (16 x 25 pixels each). The wavelength ranges of the two spectral channels are 42 – 110 microns and 110 – 210 microns, referred to as the BLUE and RED channels, respectively.

* https://sofia.usra.edu/researchers/proposing-and-observing-0
Multiplexing takes place both spectrally and spatially. An image slicer redistributes 5 x 5 pixel spatial fields-of-view (approximately diffraction-limited in each wave band) along the 1 x 25 pixel entrance slits of the spectrometers. Anamorphic collimator mirrors help keep the spectrometer compact in the cross-dispersion direction. The spectrally dispersed images of the slits are anamorphically projected onto the detector arrays, to independently match spectral and spatial resolution to detector size, thus enabling instantaneous coverage over a velocity range of ~ 1500 to 3000 km/s around selected FIR spectral lines, for each of the 25 spatial pixels (“spaxels”).

The detectors are read out with integrating amplifiers: at each pixel a current proportional to the incident flux charges a capacitor. The resulting voltage is sampled at about 256Hz. After a certain number of read-outs (the ramp length), the capacitors are reset to prevent saturation. Thus, the data consist of linearly rising ramps for which the slope is proportional to the flux. See Figure 2 for an illustration of the read-out sequence.
Figure 2: FIFI-LS readout sequence

Frequencies of pattern generator output lines:

- CLK: 8192 Hz
- SYNC: 256 Hz
- RCRES: 256 Hz / (ramps per complete chop cycle)
- CHOP: 256 Hz / (ramps per fundamental chop cycle)
- MOVE_GRAT: 256 Hz / ([ramps per complete chop cycle] x [chop cycles per grating step])

Pattern generator clock frequency: 4 x 8192 Hz (CLK and everything else in pattern generator derived from this)

Digital multiplexer clock frequency: 64 x 8192 Hz (PLL-generated inside WRE from CLK)

Chopper signal sampling frequency (both x nd y): 2048 Hz = CLK/4, will produce 2 x 8 = 16 samples per 1/256 s ramp sampling interval which will be mapped onto 16 out of 18 "pixels" of one fake detector column
2.2 FIFI-LS observing modes

Symmetric chop mode, also known as nod-match-chop mode, is the most efficient observing mode. In this mode, the telescope chops symmetrically to its optical axis, with a matched telescope nod to remove background. A typical observation sequence will cycle through the A nod position and the B nod position in an ABBA pattern.

Most observations will be taken using symmetric chop mode. However, if the object is very bright, the efficiency is improved by observing in an asymmetric chopping mode. This mode typically consists of two map positions and one off-position per nod-cycle (an AAB pattern, where the B position contains only empty sky). Asymmetric chopping may also be used if an object’s size requires a larger chop throw than is possible with symmetric chopping.

![Figure 3: The geometry of chopping and nodding in the symmetric chop mode (left) and the asymmetric mode (right).](image)

Occasionally, for very bright targets, it may be advantageous to take data with no chopping at all. This mode, called total power mode, may be taken with either symmetric or asymmetric nodding, or with no nods at all.

At each chop and nod position, it is common to step the grating through a number of positions before each telescope move. These additional grating scans effectively increase the wavelength coverage of the observation.
3 ALGORITHM DESCRIPTION

3.1 Overview of data reduction steps

This section will describe, in general terms, the major algorithms used to reduce a FIFI-LS observation. See the figure below for a flow chart showing how these algorithms fit together.

![Flow chart showing data reduction steps](image)

**Figure 4:** Processing steps for FIFI-LS data. The blue box describes an overview of steps and the white box contains the actual steps carried out.

3.2 Reduction Algorithms

The following subsections detail each of the data reduction pipeline steps outlined in the flowchart above.

3.2.1 Split Grating and Chop

A single FIFI-LS raw FITS file contains the data from both chop positions (on and off) for each grating step used in a single nod position. FIFI-LS records its grating positions in “inductosyn” units. These are long integer values that are used to convert the data to a micron scale in the wavelength calibrate step.

The raw FIFI-LS data consist of a header (metadata describing the observation) and a table of voltage readings from the detector pixels. Each data section contains one frame, i.e. simultaneous readouts of all detector pixels and chopper values.

The data header is sent before each frame. The following 8 unsigned 16-bit words contain the header information.

- **Word 0:** The word #8000 marks the start of the header.
- **Word 1:** The low word of the 32-bit frame counter.
- **Word 2:** The high word of the 32-bit frame counter.
- **Word 3:** The flag word. Bit 0 is the chopper signal. Bit 1 is the detector (0=red, 1=blue). Unused bits are fixed to 1 to recognize this flag word.
- **Word 4:** The sample count as defined in the timing diagrams (see Figure 2 above). This count gets advanced at every sync pulse and reset at every long sync pulse.
• Word 5: The ramp count as defined in the timing diagrams. This counter gets advanced with a long sync pulse and reset by RCRES.
• Word 6: The scan index.
• Word 7: A spare word (for now used as “end of header”: #7FFF).

Only columns 3, 4, and 5 are used in the split grating/chop step. The following shows an example header for a raw RED FIFI-LS file, as read in IDL:

```
IDL> data = mrdfits('fifi_raw.fits', 1, /unsigned)
IDL> print, data.header
32768  28160    15    1    0    0   89   32767
32768  28161    15    1    1    0   89   32767
32768  28162    15    1    2    0   89   32767
32768  28163    15    1    3    0   89   32767
32768  28164    15    1    4    0   89   32767
32768  28165    15    1    5    0   89   32767
32768  28166    15    1    6    0   89   32767
32768  28167    15    1    7    0   89   32767
32768  28168    15    1    8    0   89   32767
32768  28169    15    1    9    0   89   32767
32768  28170    15    1   10    0   89   32767
32768  28171    15    1    11   0   89   32767
32768  28172    15    1    12   0   89   32767
32768  28173    15    1    13   0   89   32767
32768  28174    15    1    14   0   89   32767
32768  28175    15    1    15   0   89   32767
32768  28176    15    1    16   0   89   32767
32768  28177    15    1    17   0   89   32767
32768  28178    15    1    18   0   89   32767
32768  28179    15    1    19   0   89   32767
32768  28180    15    1    20   0   89   32767
32768  28181    15    1    21   0   89   32767
32768  28182    15    1    22   0   89   32767
32768  28183    15    1    23   0   89   32767
32768  28184    15    1    24   0   89   32767
32768  28185    15    1    25   0   89   32767
32768  28186    15    1    26   0   89   32767
32768  28187    15    1    27   0   89   32767
32768  28188    15    1    28   0   89   32767
32768  28189    15    1    29   0   89   32767
32768  28190    15    1    30   0   89   32767
32768  28191    15    1    31   0   89   32767
32768  28192    15    1    0    1   89   32767
32768  28193    15    1    1    1   89   32767
32768  28194    15    1    2    1   89   32767
32768  28195    15    1    3    1   89   32767
...
```

The third column is all ones, indicating that the data is for the RED channel, the fourth column counts the readouts from 0 to 31, and the fifth column indicates the ramp number.
Where each chop and grating position starts and stops in the raw data table is determined using the header keywords RAMPLN_[B,R], C_CYC_[B,R], C_CHOPLN, G_PSUP_[B,R], G_PSDN_[B,R] keywords. A RED data example is as follows:

\[
\begin{align*}
\text{RAMPLN}_R &= 32 / \text{number of readouts per red ramp} \\
\text{C_CHOPLN} &= 64 / \text{number of readouts per chop position} \\
\text{G_PSUP}_R &= 5 / \text{number of grating position up in one cycle} \\
\text{G_PSDN}_R &= 0 / \text{number of grating position down in one cycle}
\end{align*}
\]

Here, \( \text{C_CHOPLN} / \text{RAMPLN}_R \) is \( 64 / 32 = 2 \); therefore, there are 2 ramps per chop.

Each chop switch index is determined using the 5th column in the header. It is chop 0 if the value is odd and chop 1 if the value is even. Grating scan information determines how that chop phase is split up into separate extensions using the following formula:

\[
\text{binsize} = \left( n\text{readout} \times \text{ramplength} \right) / \left( n\text{step} \times \text{choplength} \right)
\]

where \( n\text{readout} \) is the total number of readouts (frames) in the file, \( \text{ramplength} \) is determined by the appropriate RAMPLN keyword, \( n\text{step} \) is the number of grating steps (\( \text{G_PSDN} + \text{G_PSUP} \)), and \( \text{choplength} \) is the number of readouts per chop position (C_CHOPLN).

The binary data section is comprised of 468 signed 16-bit words: one each for 25 spaxels, plus one control value, times 18 spectral channels (“spexels”). The spaxels are read out one spectral channel at a time. Spectral channel zero of all 25 pixels are read out, and then a chopper value (analog readout from the secondary mirror) is recorded; then the next channel of all the pixels is read out, and then the next chopper value, and so on, through all the spectral channels. The chopper values are discarded during pipeline processing. Of the 18 spectral channels, channel 0 is the CRE resistor row and row 17 is the blind CRE row (“dummy channels”). These two channels are discarded; the other 16 channels are considered valid spexels.

In IDL, when the raw FITS file is read in using the \texttt{mrdfits} task, the binary table is represented as an array of structures consisting of an unsigned, one-dimensional eight-element array (element name HEADER), and a \( 26 \times 18 \times n\text{frame} \) signed long-integer array (element name DATA). The size of the third dimension of the array equals the number of frames acquired by the instrument.

The first step in the data reduction pipeline is to split out the data from each grating scan position into separate FITS tables, and to save all grating positions from a single chop position into a common file. For example, if there are five grating scans per chop, and two chop positions, then a single one-extension input raw FITS file will be reorganized into two files (chop 0 and chop 1) with 5 extensions each. For total power mode, there is only one chop position, so there will be one output file, with one extension for each grating step. Each extension in the output FITS files contains the data table corresponding to the grating position recorded in its header (keyword INDPOS).
Hereafter, in the pipeline, until the Combine Grating Scans step, each grating scan extension is handled separately.

3.2.2 Fit Ramps
The flux measured in each spatial and spectral pixel is reconstructed from the readout frames by fitting a line to the voltage ramps. The slope of the line corresponds to the flux value.

Before fitting a line to a ramp, some likely bad frames are removed from the data. The chopper values (in the 26th spaxel position), the first ramp from each spaxel, and the first and last readout per ramp are all removed before fitting. Also, a ramp may be marked as saturated if it does not have its highest peak in the last readout of the ramp. If this occurs, the readout before the highest peak is removed before fitting, along with any readouts after it. This ensures that the slope is not contaminated by any non-linearity near the saturation point.

Typically, multiple ramps are taken at each chop position. After the slope of each ramp is derived, the slopes are combined with a robust weighted mean. This final averaged value is recorded as the flux for the pixel and the error on the mean is recorded as the error on the flux for that pixel. After this pipeline step, there is a flux value for each spatial and spectral pixel, recorded in a 5 x 5 x 18 data array in a separate FITS extension for each grating scan and chop position. The error values are recorded in a separate 5 x 5 x 18 table in the same FITS extension. The extension elements are named DATA and STDDEV, respectively.

Some pixels in the data array may be set to not-a-number (NaN). These are either known bad detector pixels, or pixels for which the ramp fits did not have sufficient signal-to-noise. These pixels will be ignored in all further reduction steps.
3.2.3 Subtract Chops

To remove instrument and sky background emission, the data from the two chop positions must be subtracted. For A nods, chop 1 is subtracted from chop 0. For B nods, chop 0 is subtracted from chop 1. All resulting source flux in the chop-subtracted data should therefore be positive, so that the nods are combined only by addition. This pipeline step produces one output file for each pair of input chop files. In total power mode, no chop subtraction is performed.
3.2.4 Combine Nods
After the chops are subtracted, the nods must be combined to remove residual background.

In symmetric chopping mode, the A nods are paired to adjacent B nods. In order to match a given A nod, a B nod must have been taken at the same dither position (FITS header keywords DLAM_MAP and DBET_MAP), and with the same grating position (INDPOS). The B nod meeting these conditions and taken nearest in time to the A nod (keyword DATE-OBS) is added to the A nod.

In asymmetric mode, a single B nod may be added to multiple A nods. For example, in an AAB pattern the B nod is combined with each preceding A nod. The B nods in this mode need not have been taken at the same dither position as the A nods, but the grating position must still match. The matching B nod taken nearest in time is added to each A nod.
This pipeline step produces an output file for each input A nod file, containing the chop- and nod-combined flux and error values for each grating scan.

![Image](image.png)

**Figure 7:** The chop-subtracted nod A flux array, with the corresponding nod B added.

### 3.2.5 Wavelength Calibrate

The wavelength calibrate step calculates wavelength values in microns for each of the spectral pixels in each grating scan, based on the known grating position, a model of the optical geometry of the instrument, and measurements of the positions of known spectral lines. The optics within FIFI-LS tend to drift with time and therefore the FIFI-LS team updates the wavelength solution every year. The wavelength equation (below) is stored in a script, while all relevant constants are stored in a reference table, with an associated date of applicability.

The wavelength ($\lambda_i$) for the pixel at spatial position $i$ and spectral position $j$ is calculated from the equation:

$$\phi_i = 2\pi ISF \frac{ind + ISOFF_i}{2^4}$$
\[ \delta_j = [j - 8.5] \cdot PS + sign[j - QOFF] \cdot [j - QOFF]^2 \cdot QS \]
\[ g_i = g_0 \cdot \cos \left( \tan^{-1} \left( \frac{\text{SlitPos}_i - NP}{a} \right) \right) \]
\[ \lambda_{ij} = 1000 \frac{g_i}{m} \left[ \sin(\phi_i - \gamma) + \sin(\phi_i + \gamma + \delta_j) \right] \]

where:

- \textit{ind}: the input inductosyn position
- \textit{m}: the spectral order of the observation (1 or 2)

are inputs that depend on the observation settings, and

- \textit{ISF}: inductosyn scaling factor
- \textit{PS}: pixel scale, in radians
- \textit{QOFF}: offset of quadratic pixel scale part from the “zero” pixel, in pixels
- \textit{QS}: quadratic pixel scale correcting factor, in radians/pixel^2
- \textit{g_0}: grating constant
- \textit{NP}: slit position offset
- \textit{a}: slit position scale factor
- \textit{\gamma}: offset from Littrow angle

are constants determined by the FIFI-LS team, and

- \textit{ISOFF}: offset of the home position from grating normal in inductosyn units for the \textit{i}th spaxel
- \textit{SlitPos_i}: slit position of the \textit{i}th spaxel

are values that depend on the spatial position of the pixel, also determined by the FIFI-LS team. The spaxels are ordered from 1 to 25 spatially as follows:

<table>
<thead>
<tr>
<th>Spaxel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td>4</td>
</tr>
<tr>
<td>5</td>
</tr>
<tr>
<td>6</td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
<tr>
<td>17</td>
</tr>
<tr>
<td>18</td>
</tr>
<tr>
<td>19</td>
</tr>
<tr>
<td>20</td>
</tr>
<tr>
<td>21</td>
</tr>
<tr>
<td>22</td>
</tr>
<tr>
<td>23</td>
</tr>
<tr>
<td>24</td>
</tr>
<tr>
<td>25</td>
</tr>
</tbody>
</table>

with corresponding slit position:

<table>
<thead>
<tr>
<th>Slit position</th>
<th>Spaxel</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>23</td>
</tr>
<tr>
<td>4</td>
<td>24</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
</tr>
<tr>
<td>6</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>19</td>
</tr>
<tr>
<td>11</td>
<td>20</td>
</tr>
<tr>
<td>12</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>11</td>
</tr>
<tr>
<td>14</td>
<td>12</td>
</tr>
<tr>
<td>15</td>
<td>13</td>
</tr>
<tr>
<td>16</td>
<td>14</td>
</tr>
<tr>
<td>17</td>
<td>15</td>
</tr>
<tr>
<td>18</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>6</td>
</tr>
<tr>
<td>20</td>
<td>7</td>
</tr>
<tr>
<td>21</td>
<td>8</td>
</tr>
<tr>
<td>22</td>
<td>9</td>
</tr>
<tr>
<td>23</td>
<td>10</td>
</tr>
<tr>
<td>24</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>1</td>
</tr>
</tbody>
</table>
Note that each spectral pixel has a different associated wavelength, but it also has a different effective spectral width. This width \( (d\lambda/dp) \) is calculated from the following equation:

\[
d\lambda_i/dp = 1000 \frac{q_i}{m} [PS + 2 * \text{sign}(i - QOFF) * (j - QOFF) * QS] \cos(\phi_i + \gamma + \delta_j)
\]

where all variables and constants are defined above.

At this stage, the first and last spectral pixels (the dummy channels, top and bottom rows in Figure 7) are removed from the flux and error arrays, so that their sizes become 5 x 5 x 16. The wavelength values and spectral widths calculated by the pipeline for each pixel are stored in a new 5 x 5 x 16 table in each grating scan extension (elements LAMBDA and DLAMDPIX).

### 3.2.6 Spatial Calibrate

The locations of the spaxels are not uniform across the detector due to the optics not being perfectly aligned. See Figure 8 for a plot of the average of the center of each spaxel location as measured in the lab. This location is slightly different at each wavelength. These spaxel positions are determined by the FIFI-LS team and recorded in a look-up table.
Figure 8: Average fitted spaxel positions in telescope simulator coordinates (red and blue channel) in the focal plane of the telescope. In these coordinates, blue spaxels are about 1.5 mm along an edge; red spaxels are about 3 mm. The plate scale is about 4 arcseconds per mm.

For a particular observation, the recorded dither offsets in arcseconds are used to calculate the x and y coordinates for the pixel in the ith spatial position and the jth spectral position using the following formulae:

\[
\begin{align*}
x_{ij} &= ps \cdot (xpos_{ij} + dx) + d\lambda \cos(\theta) - d\beta \sin(\theta) \\
y_{ij} &= ps \cdot (ypos_{ij} + dy) + d\lambda \sin(\theta) + d\beta \cos(\theta)
\end{align*}
\]

where \(ps\) is the plate scale in arcseconds/mm (FITS header keyword PLATSCAL), \(d\lambda\) is the right ascension dither offset in arcseconds (keyword DLAM_MAP), \(d\beta\) is the declination dither offset in arcseconds (keyword DBET_MAP), \(\theta\) is the detector angle, \(xpos_{ij}\) and \(ypos_{ij}\) are the fitted spaxel positions in mm for pixel \(i, j\), and \(dx\) and \(dy\) are the spatial offsets between the primary array (usually BLUE), used for telescope pointing, and the secondary array (usually RED). The \(dx\) and \(dy\) offsets are zero for the data taken with primary camera. By default, the coordinates are then rotated by the detector angle.
(minus 180 degrees), and the y-coordinates are inverted in order to set North up and East left in the final coordinate system:

\[
\begin{align*}
    x'_{ij} &= -x_{ij} \cos(\theta) + y_{ij} \sin(\theta) \\
    y'_{ij} &= x_{ij} \sin(\theta) + y_{ij} \cos(\theta)
\end{align*}
\]

The pipeline stores these calculated x and y coordinates in two 5 x 5 x 16 data tables in each FITS extension (elements XS and YS).

### 3.2.7 Apply Flat

In order to correct for variations in response among the individual pixels, the FIFI-LS team has generated flat field arrays that correct for the average difference in spatial and spectral response. There are normalized flat fields available for a range of standard central wavelengths: 52, 57, 63, 70, 72, 79, 88, 96, 118, 145, 153, 158, 163, 185, and 200 microns. The pipeline selects the flat that is nearest the central wavelength for the observation (FITS keyword G_WAVE_R for the RED channel, G_WAVE_B for the BLUE channel), and divides the flux array by the flat. The output file format does not change for this step.

![Figure 9: The flat-corrected flux array.](image)

### 3.2.8 Combine Grating Scans

Up until this point, all processing has been done on each grating scan extension separately. The pipeline now combines the data from all grating scans, in order to fill in the wavelength coverage of the observation.
Since each pixel has a different spectral width, the pipeline first applies a scaling factor to each pixel to correct it to a common width. The scale factor for the pixel in the $i$th spatial position and the $j$th spectral position is the mean spectral width of the first extension in the input file, divided by $d\lambda/dp_{ij}$, as calculated above, in the wavelength calibration step. The integrated flux for each pixel in each grating scan should then be directly comparable. The common spectral width is recorded in the output FITS header under the keyword SPEXLWID.

Due to slight variations in the readout electronics, however, there may also be additive offsets in the overall flux level recorded in each grating scan. To correct for this bias offset, the pipeline calculates the mean value of all pixels in the overlapping wavelength regions for each grating scan. This value, minus the mean over all scans, is subtracted from each grating scan, in order to set all extensions to a common bias level.

Finally, the pipeline sorts the data from all grating scans by their associated wavelength values in microns, and stores the result in a single data array with dimensions $5 \times 5 \times (16 \times n_{scan})$, where $n_{scan}$ is the total number of grating scans in the input file. Note that the wavelengths are still irregularly sampled at this point, due to the differing wavelength solutions for each grating scan and spatial pixel. All arrays in the output fits file (DATA, STDDEV, LAMBDA, XS, and YS) now have dimensions $5 \times 5 \times (16 \times n_{scan})$. The DLAMDP IX array is removed, since the fluxes have been corrected to a common spixel width.
Figure 10: Example spectral flux from the center spaxel for a single dither position. Each different symbol and color represents a different grating scan, after correction to a common spectral width. The gray line indicates the combined flux array, after bias correction.
Figure 11: The full 5 x 5 x 80 flux array (flattened to 25 x 80), after combining 5 grating scans.
3.2.9 Telluric Correct

Telluric correction is the process of attempting to correct an observed spectrum for absorption by molecules in the earth’s atmosphere, in order to recover the intrinsic (“exo-atmospheric”) spectrum of the source. The atmospheric molecular components (primarily water, ozone, CO2) can produce both broad absorption features that are well resolved by FIFI-LS and narrow, unresolved features. The strongest absorption features are expected to be due to water. Because SOFIA travels quite large distances during even short observing legs, the water vapor content along the line of sight through the atmosphere can vary substantially on fairly short timescales during an observation. Therefore, observing a “telluric standard,” as is usually done for ground-based observations, will not necessarily provide an accurate absorption correction spectrum. For this reason, telluric corrections of FIFI-LS data relies on models of the atmospheric absorption, as provided by codes such as ATRAN, in combination with the estimated line-of-sight water vapor content (precipitable water vapor, PWV) provided by the water vapor monitor (WVM) aboard SOFIA. Currently, the WVM does not generate PWV values that are inserted into the FITS headers of the FIFI-LS data files. It is expected that these values will become available in the near future and at that point the PWV values will be used to generate telluric correction spectra. Until then, correction spectra are generated using the expected PWV value for the flight altitude and airmass appropriate for the observations. Experience with both FIFI-LS and the FORCAST instrument has shown that this method can produce reasonably accurate corrections in the vicinity of broad, shallow telluric absorption features. However, corrections in regions with deep, sharp features (e.g., near 63 microns) are extremely sensitive to the actual PWV along the line of sight. Without accurate atmospheric models and measurements of the PWV, it cannot be expected that standard models will provide reliable telluric corrections. Furthermore, accurate correction of spectral lines in the vicinity of narrow telluric absorption features is problematic even with the use of good atmospheric models and knowledge of the PWV. This is due to the fact that the observed spectrum is the result of a multiplication of the intrinsic spectrum by the telluric absorption spectrum, and then a convolution of the product with the instrumental profile, whereas the correction derived from a model is the result of the convolution of the theoretical telluric absorption spectrum with the instrumental profile. The division of the former by the latter does not necessarily yield the correct results, and the output spectrum may retain telluric artifacts after telluric correction.

A set of ATRAN models appropriate for a range of altitudes and zenith angles has been generated for pipeline use. In this step, the pipeline selects the model closest to the observed altitude and zenith angle, smooths the transmission model to the resolution of the observed spectrum, bins the transmission data to the observed wavelength and spectral width of each spexel, and then divides the data by the transmission model. Very low transmission values result in poor corrections, so any pixel for which the transmission is less than 60% (by default) is set to NaN. For reference, the smoothed,
binned transmission data is attached to the FITS file as a 5 x 5 x (16 * nscan) data table in the first FITS extension (element ATRAN).

Since the telluric correction may introduce artifacts, or may, at some wavelength settings, produce flux cubes for which all pixels are set to NaN, the pipeline also propagates the uncorrected flux cube through the remaining reduction steps. The telluric-corrected cube and its associated error are stored in the DATA and STDDEV fields of the binary table. The uncorrected cube and its associated error are stored in the UNCORRECTED_DATA and UNCORRECTED_STDDEV fields.
Figure 12: The telluric-corrected flux array. The lower pixels are set to NaN due to poor atmospheric transmission at those wavelengths.
### 3.2.10 Flux Calibrate

Flux calibration of FIFI-LS data is carried out via the division of the instrumental response, as recorded in response files appropriate for each grating setting, wavelength range, and dichroic. The response values have units of V/s/Jy and are derived from observations of “flux standards.” At the wavelengths at which FIFI-LS operates, there are very few stars bright enough to yield high signal-to-noise data useful for flux calibration purposes. Therefore, observations of asteroids, planets, and planetary moons are used, along with models of such objects, to derive the response curves. Since the observed fluxes of such solar system objects vary with time, the models must be generated for the time of each specific observation. To date, observations of Mars have been used as the primary flux calibration source. Predicted total fluxes for Mars across the FIFI-LS passband at the specific UT dates of the observations have been generated using the model of Lellouch and Amri. Predicted fluxes at several frequencies have been computed and these have then been fit with blackbody curves to derive values at a large number of wavelength points. The deviations of the fits from the input predictions are much less than 1%. After the models have been generated, the telluric-corrected spectra of the standards, in units of V/s, are divided by the theoretical spectra, in Jy. The results are smoothed and then fit with a polynomial to derive response functions (V/s/Jy) that can then used to flux calibrate the telluric-corrected spectra of other astronomical sources.

The pipeline stores a set of response functions for each channel and dichroic value. To perform flux calibration, it selects the correct response function for each input observation, interpolates the response onto the wavelengths of each spexel, and divides the flux by the response value to convert it to physical units (Jy). From this point on, the data products are considered to be Level 3 (FITS keyword PROCSTAT=LEVEL_3). For reference, the resampled response data is attached to the FITS file as a 5 x 5 x (16 * nscan) data table in the first FITS extension (element RESPONSE). Flux calibration is applied to both the telluric-corrected cube and the uncorrected cube. The estimated systematic error in the flux calibration is recorded in the CALERR keyword in the FITS header, as an average fractional error. At this time, flux calibration is expected to be good to within about 20% (CALERR = 0.2).

### 3.2.11 Correct Wave Shift

Due to the motion of the earth with respect to the local standard of rest, the wavelengths of features in the spectra of astronomical sources will appear to be slightly shifted, by different amounts on different observation dates. In order to avoid introducing a broadening of spectral features when multiple observations obtained over different nights are combined, the wavelength calibration of FIFI-LS observations must be adjusted to remove the barycentric wavelength shift. This shift is calculated as an expected wavelength shift \( \frac{d \lambda}{\lambda} \), from the radial velocity of the earth with respect to the sun and the sun with respect to the local standard of rest, on the observation date, toward the RA

---

† See http://www.lesia.obspm.fr/perso/emmanuel-lellouch/mars/index.php
and Dec of the observed target. This shift is recorded in the header keyword BARYSHIFT, and applied to the wavelength calibration in the LAMBDA field as:

\[ \lambda' = \lambda + \lambda \left(\frac{d\lambda}{\lambda}\right) \]

Since the telluric absorption lines do not change with the motion of the earth, the barycentric wavelength shift cannot be applied to non-telluric-corrected data. Doing so would result in a spectrum in which both the intrinsic features and the telluric lines are shifted. Therefore, the unshifted wavelength calibration is also propagated in the output file, in the field UNCORRECTED_LAMBDA.

### 3.2.12 Wavelength Resample

Next, the pipeline resamples the spectral flux for each spatial pixel onto a regular grid of wavelengths. It first determines the maximum and minimum wavelengths present in all input files, from all dither positions for the observation. This sets the range of the output grid. The spacing of the output grid is set by the desired oversampling. By default, the pipeline samples the average expected spectral FWHM for the observation with 8 output pixels. For example, for a BLUE order 2 observation, as in the Figures above, the average spectral resolution is 900. If the input files span a wavelength range of 51.65 to 51.99 microns, the average wavelength is 51.82 microns, and the expected spectral FWHM is 51.82 / 900 = 0.0576 microns. Sampling this FWHM with 8 pixels creates a grid with a spectral width of 0.0072 microns. This will sample the full range of wavelengths with 48 spectral pixels (see Figure 14). Since the pixel width changes in the resampling from the intrinsic spectral width, as recorded in the combine grating scans step (keyword SPEXLWID), the output flux will be corrected by a factor of the new width, divided by the old spectral width, in order to conserve integrated flux.

For each spatial pixel, the resampling is performed by looping over the output wavelength grid, and finding all flux values with assigned wavelengths within a fitting window (typically ±0.25 times the spectral FWHM). Outlier flux values (typically those greater than 5 sigma from the mean value) may be rejected. Then, a low-order one-dimensional polynomial is fit to all the good data points, with weighting by the error on the flux, as calculated by the pipeline, and the distance of the input data value from the grid location. The output flux value is the value of the polynomial, evaluated at the grid location, and the associated error is the error on the fit.

Both the telluric-corrected and the uncorrected flux cubes are resampled in this step, onto the same wavelength grid. However, the telluric-corrected cube is resampled using the wavelengths corrected for barycentric motion, and the uncorrected cube is resampled using the original wavelength calibration. The spectra from the uncorrected cube will appear slightly shifted with respect to the spectra from the telluric-corrected cube.

The pipeline stores the resampled data as a 2D FITS image extension (nwave x 25) with extension name FLUX. The associated error (also nwave x 25) is stored in a separate extension, with the name ERROR. The non-telluric-corrected cubes are stored in UNCORRECTED_FLUX and UNCORRECTED_ERROR extensions, respectively, and
will also have size \( n_{\text{wave}} \times 25 \). The output wavelength values for the grid are stored in a 1D extension called WAVELENGTH, with size \( n_{\text{wave}} \). The \( x \) and \( y \) coordinates of the spatial pixels are recorded in 1D \( X \) and \( Y \) extensions, each with size 25. For reference, a model of the atmospheric transmission spectrum, smoothed to the resolution of the observation, and the instrumental response curve used in flux calibration are also attached to the FITS file in 1D extensions called TRANSMISSION and RESPONSE, with size \( n_{\text{wave}} \). In the example above, \( n_{\text{wave}} \) is 48.

![Figure 13: Example spectral flux from the center spaxel. Black is the input flux; red is the resampled flux.](image)
3.2.13 Spatial Resample

Finally, the pipeline resamples each wavelength plane onto a regular spatial grid. This step combines the spatial information from all input nod-combined dither positions into a single output map. Note that these input files are resampled separately in the wavelength resample step, but onto the same master wavelength grid, so that they may be compared directly. The maximum and minimum x and y offsets in the input data set the range of the spatial grid. The spacing ($dx$) is set by a reference spatial FWHM for the observed channel (5" for BLUE, 10" for RED), divided by the desired oversampling. For example, for a BLUE order 2 observation, the default oversampling of 5 pixels sets the output grid spacing to 1 arcsecond. If the range of x offsets is -14.0 to 15.8 arcseconds, and the range of y offsets is -15.9 to 15.7 arcseconds, then the output spatial grid will have dimensions 31 x 33. The pixel width again changes after resampling, so the output flux is multiplied by the area of the new pixel ($dx^2$), divided by the intrinsic area of the spaxel (approximately 36 arcseconds$^2$ for BLUE, 144 arcseconds$^2$ for RED).

The output files are specified to have a spatial sampling of 1” per pixel in the BLUE and 2” per pixel in the RED.

![Figure 14: Wavelength resampled flux array, now 25 x 48](image-url)
The spatial resampling algorithm is similar to the wavelength-resampling algorithm. For each wavelength plane, the algorithm loops over the output spatial grid, finding values within a local fitting window. For the spatial grid, a larger fit window is typically necessary than for the spectral grid, since the observation setup usually allows more oversampling in wavelength than in space. The default value for the fit window is 3.0 times the average FWHM for the channel observed. Outlier flux values (typically those greater than 10 sigma away from the mean within the window) are rejected, and remaining good values are fit with a low order polynomial surface fit. The fits are usually weighted by the error on the flux and a Gaussian function of the distance from the data point to the grid location. The output flux for each pixel is the value of the surface polynomial, evaluated at the grid location, and the associated error value is the error on the fit.

For some types of observations, especially undithered observations of point sources, for which the spatial FWHM is undersampled, the polynomial surface fits may not return good results. In these cases, it may be beneficial to use an alternate resampling algorithm. In this algorithm, the master spatial grid is determined as above, but each input file is interpolated onto it individually, for each wavelength plane, using a radial basis function interpolation. Areas of the spatial grid for which there is no data in the input file are set to NaN. The interpolated cubes are then mean-combined, ignoring any NaNs, to produce the final output cube.

For either algorithm, the pipeline also generates an exposure map cube indicating the number of observations of the source that were taken at each pixel (see Figure 16). The final resampled data cube is considered to be Level 4, regardless of whether it is generated from data from a single flight or multiple flights (FITS keyword PROCSTAT=LEVEL_4).

The final output from the pipeline is a FITS file with 10 image extensions:
• FLUX: The $nx \times ny \times nw$ cube of flux values.
• ERROR: The associated error values on the flux (also $nx \times ny \times nw$).
• UNCORRECTED_FLUX: The $nx \times ny \times nw$ cube of flux values that have not been corrected for atmospheric transmission.
• UNCORRECTED_ERROR: The associated error values on the uncorrected flux (also $nx \times ny \times nw$).
• WAVELENGTH: The wavelength values associated with each plane of the cube ($nw$).
• X: The $x$-coordinates of the data, in arcsecond offsets from the base position ($nx$).
• Y: The $y$-coordinates of the data, in arcsecond offsets from the base position ($ny$).
• TRANSMISSION: The atmospheric transmission model ($nw$).
• RESPONSE: The instrumental response curve ($nw$).
• EXPOSURE_MAP: The exposure map ($nx \times ny \times nw$).

Figure 16: Exposure map of input dither positions. Values range from 4 to 16.
Figure 17: The final output flux cube. The image on top is a spatial slice at wavelength 51.804 µm. The plot on the bottom is a spectral slice at pixel 13, 14 (offset 0.26, -0.47 arcseconds).
4 DATA PRODUCTS

4.1.1 Filenames
FIFI-LS output files from Redux are named according to the convention:

FILENAME = F####_FI_IFS_AOR-ID_CHANNEL_Type_FN1[-FN2].fits,

where #### is the four-digit SOFIA flight number, FI is the instrument identifier, IFS specifies that it is integral field spectroscopy data, AOR-ID is the 8 digit AOR identifier for the observation, CHANNEL is either BLU or RED, Type is three letters identifying the product type (listed in the table below), and FN1 is the file number corresponding to the input file. FN1-FN2 is used if there are multiple input files for a single output file, where FN1 is the file number of the first input file and FN2 is the file number of the last input file.

4.1.2 Pipeline Products
The following table lists all intermediate products generated by Redux for FIFI-LS, in the order in which they are produced. The product type is stored in the primary FITS header of the file, under the keyword PRODTYPE. By default, the scan_combined, flux_calibrated, and wavelength_resampled, and wxy_resampled products are saved. Specifying the appropriate option in either the automatic or interactive modes will save all products. Note that as of pipeline version 1.3.2, the STDDEV and ERROR extensions do not include the systematic error on the flux calibration; this is recorded separately in the FITS header keyword CALERR in the primary extension.
<table>
<thead>
<tr>
<th>Step</th>
<th>Description</th>
<th>Product Type</th>
<th>Proc. status</th>
<th>File code</th>
<th>Saved</th>
<th>Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Split Grating / Chop</td>
<td>Data split into separate files by chop number and extensions by grating scans</td>
<td>grating_chop_split LEVEL_2</td>
<td></td>
<td>CP0, CP1</td>
<td>N</td>
<td>N\textsubscript{scan} image extensions, each 26x18x1xN\textsubscript{real}, where N\textsubscript{scan} is the number of grating scans and N\textsubscript{real} is the number of readout frames</td>
</tr>
<tr>
<td>Fit Ramps</td>
<td>Readout ramps fit and averaged</td>
<td>ramps_fit LEVEL_2</td>
<td></td>
<td>RP0, RP1</td>
<td>N</td>
<td>N\textsubscript{scan} binary table extensions, each containing DATA and STDDEV data cubes, each 5x5x18</td>
</tr>
<tr>
<td>Subtract Chops</td>
<td>Chop positions subtracted</td>
<td>chop_subtracted LEVEL_2</td>
<td></td>
<td>CSB</td>
<td>N</td>
<td>N\textsubscript{scan} binary table extensions, each containing DATA and STDDEV data cubes, each 5x5x18</td>
</tr>
<tr>
<td>Combine Nods</td>
<td>Nod positions combined</td>
<td>nod_combined LEVEL_2</td>
<td></td>
<td>NCM</td>
<td>N</td>
<td>N\textsubscript{scan} binary table extensions, each containing DATA and STDDEV data cubes, each 5x5x18</td>
</tr>
<tr>
<td>Lambda Calibrate</td>
<td>Wavelength values calculated and attached to data cube</td>
<td>wavelength_calibrated LEVEL_2</td>
<td>WAV</td>
<td></td>
<td>N</td>
<td>N\textsubscript{scan} binary table extensions, each containing DATA, STDDEV, LAMBDA, and DLAMDPPIX data cubes, each 5x5x16</td>
</tr>
<tr>
<td>Spatial Calibrate</td>
<td>Spatial values calculated and attached to data cube</td>
<td>spatial_calibrated LEVEL_2</td>
<td>XYC</td>
<td></td>
<td>N</td>
<td>N\textsubscript{scan} binary table extensions, each containing DATA, STDDEV, LAMBDA, DLAMDPPIX, XS, and YS data cubes, each 5x5x16</td>
</tr>
<tr>
<td>Apply Flat</td>
<td>Flat field correction applied to flux</td>
<td>flat_fielded LEVEL_2</td>
<td>FLF</td>
<td></td>
<td>N</td>
<td>N\textsubscript{scan} binary table extensions, each containing DATA, STDDEV, LAMBDA, DLAMDPPIX, XS, and YS data cubes, each 5x5x16</td>
</tr>
<tr>
<td>Combine Scans</td>
<td>Combine grating scans into an unevenly spaced grid in a single extension</td>
<td>scan_combined LEVEL_2</td>
<td>SCM</td>
<td>Y</td>
<td></td>
<td>1 binary table extension, containing DATA, STDDEV, LAMBDA, XS, and YS data cubes, each 5x5x(16*N\textsubscript{scan})</td>
</tr>
<tr>
<td>Telluric Correct</td>
<td>Flux corrected for atmospheric absorption</td>
<td>telluric_corrected LEVEL_2</td>
<td>TEL</td>
<td></td>
<td>N</td>
<td>1 binary table extensions, containing DATA, STDDEV, UNCORRECTED_DATA, UNCORRECTED_STDDEV, LAMBDA, XS, YS, and ATRAN data cubes, each 5x5x(16*N\textsubscript{scan})</td>
</tr>
<tr>
<td>Flux Calibrate</td>
<td>Flux values converted to physical units (Jy)</td>
<td>flux_calibrated LEVEL_3</td>
<td>CAL</td>
<td>Y</td>
<td></td>
<td>1 binary table extension, containing DATA, STDDEV, UNCORRECTED_DATA, UNCORRECTED_STDDEV, LAMBDA, XS, YS, ATRAN, and RESPONSE data cubes, each 5x5x(16*N\textsubscript{scan})</td>
</tr>
<tr>
<td>Correct Wave Shift</td>
<td>Wavelength calibration corrected for barycentric</td>
<td>wavelength_shifted LEVEL_3</td>
<td>WSH</td>
<td></td>
<td>N</td>
<td>1 binary table extension, containing DATA, STDDEV, UNCORRECTED_DATA, UNCORRECTED_STDDEV,</td>
</tr>
</tbody>
</table>

GI Handbook for FIFI-LS Data 30
| Wave Resample | Resample data onto a regular wavelength grid | wavelength_resampled | LEVEL_3 | WGR | Y | LAMBDA, UNCORRECTED_LAMBDA, XS, YS, ATRAN, and RESPONSE data cubes, each 5x5x(16*Nscan) |
| Spatial Resample | Resample data onto a regular spatial grid | wxy_resampled | LEVEL_4 | WXY | Y | 9 image extensions: FLUX (Nw x 25), ERROR (Nw x 25), UNCORRECTED_FLUX (Nw x 25), UNCORRECTED_ERROR (Nw x 25), WAVELENGTH (Nw), X (25), Y (25), TRANSMISSION (Nw), and RESPONSE (Nw), where Nw is the number of resampled wavelength pixels |

### Table 1: Final and intermediate data products

#### 4.1.3 Data Format

All files produced by the pipeline are multi-extension FITS files, for which the primary HDU contains only the primary header, and all data is contained in separate extensions. Before the combine scans step, all outputs will have a separate binary table extension for each grating position; the output of the combine scans and flux calibrate steps will have a single binary table extension. The binary tables contain the elements listed in the table above, along with a HEADER element. After the resampling steps, the data is contained in image extensions, as listed above.