Draft version April 3, 2026
Typeset using LaTeX **default** style in AASTeX7.0.1

Documentation for SIMLA First Delivery

1. INTRODUCTION

The *Spitzer* IRS Mapping Legacy Archive (SIMLA) contains spectral cube data products for nearly[1] every low resolution (short-low or SL, and long-low or LL) mapping mode program of the InfraRed Spectrograph (IRS) onboard *Spitzer*. Each data cube is background-subtracted and bad pixels have been masked with an automated algorithm; a complete description of the data reduction process can be found in the main paper, G. P. Donnelly et al. (2025). In this document, we describe basic aspects of delivered SIMLA data products, detail the provided metadata relating to these products and their quality, and give some caveats and recommendations related to using these data. SIMLA products are available at the NASA/IPAC Infrared Science Archive (DOI: 10.26131/IRSA655). The code for the SIMLA pipeline is available for view or use at https://github.com/simlacube/simla. See the IRS Instrument Handbook (IRS Instrument Team & Science User Support Team 2021) for more details on IRS data. A brief summary of the SIMLA pipeline, which introduces concepts and terms referred to in this document, is provided below.

1.1. *Need-to-know SIMLA Pipeline Summary*

The SIMLA pipeline for IRS mapping-mode observations identifies and builds cubes with accurate pixel-based background removal. The background frames, which are key to achieving good cube quality, were built by considering:

1. The zodiacal foreground emission that arises from dust in the Solar System. We use the zodiacal emission model (ZEM) from T. Kelsall et al. (1998), with a time-varying correction to account for the variable zodiacal emission captured by IRS calibration data.

2. A relatively stable artifact pattern that varies spatially across the detector but changes shape at different levels of modeled zodiacal intensity. To remove this, we produce "baseline frames" from the median combinations of large numbers of ZEM-subtracted sky observations within different bins of zodiacal intensity. We then interpolate a custom baseline frame for each cube based on the ZEM value for that cube.

3. Noisy pixel "pedestals" that are coherent over some non-negligible but varying timescale. These pedestals are subtracted off using (foreground and artifact-subtracted) sky observations that are similar in observation time (up to $\Delta t_{\max}$) and zodiacal intensity (up to $\Delta z_{\max}$) to the cube observations (see Table 1). This is accomplished using "shards," which are subdivisions of IRS slits in the slit spatial direction, allowing for sections of a given basic calibrated data frame (BCD) to be used for backgrounds even if another section is on-source (see Figure 1). For each cube, where possible, the SIMLA pipeline creates an average of many dark, off-source shards to capture the pixel pedestals. The background depth q_{targ} serves as the benchmark for background quality, see §3.

We identify IRS observations, or shards of observations, as "dark" based on a system of cuts to: the locally background-subtracted WISE W3 photometry at 12 μm (C_{W}), the median flux of the IRS spectrum after subtracting components 1. and 2. above (C_{s}), and the modeled intensity from the Galactic interstellar medium from the T. Kelsall et al. (1998) model (C_{ISM}). The full background image for each cube then consists of the sum of the ZEM, baseline frames, and pixel pedestals measured in dark frames. Once the background has been completed, the assembly of IRS basic calibrated data (BCDs) into cubes is facilitated by `CUBISM`[2] (J. D. T. Smith et al. 2007), which produces FITS files for each cube.

[1] Except for programs marked SIRTF Calibration, observations of the continuous viewing zone, and moving targets.
[2] https://irsa.ipac.caltech.edu/data/SPITZER/docs/dataanalysistools/tools/cubism/

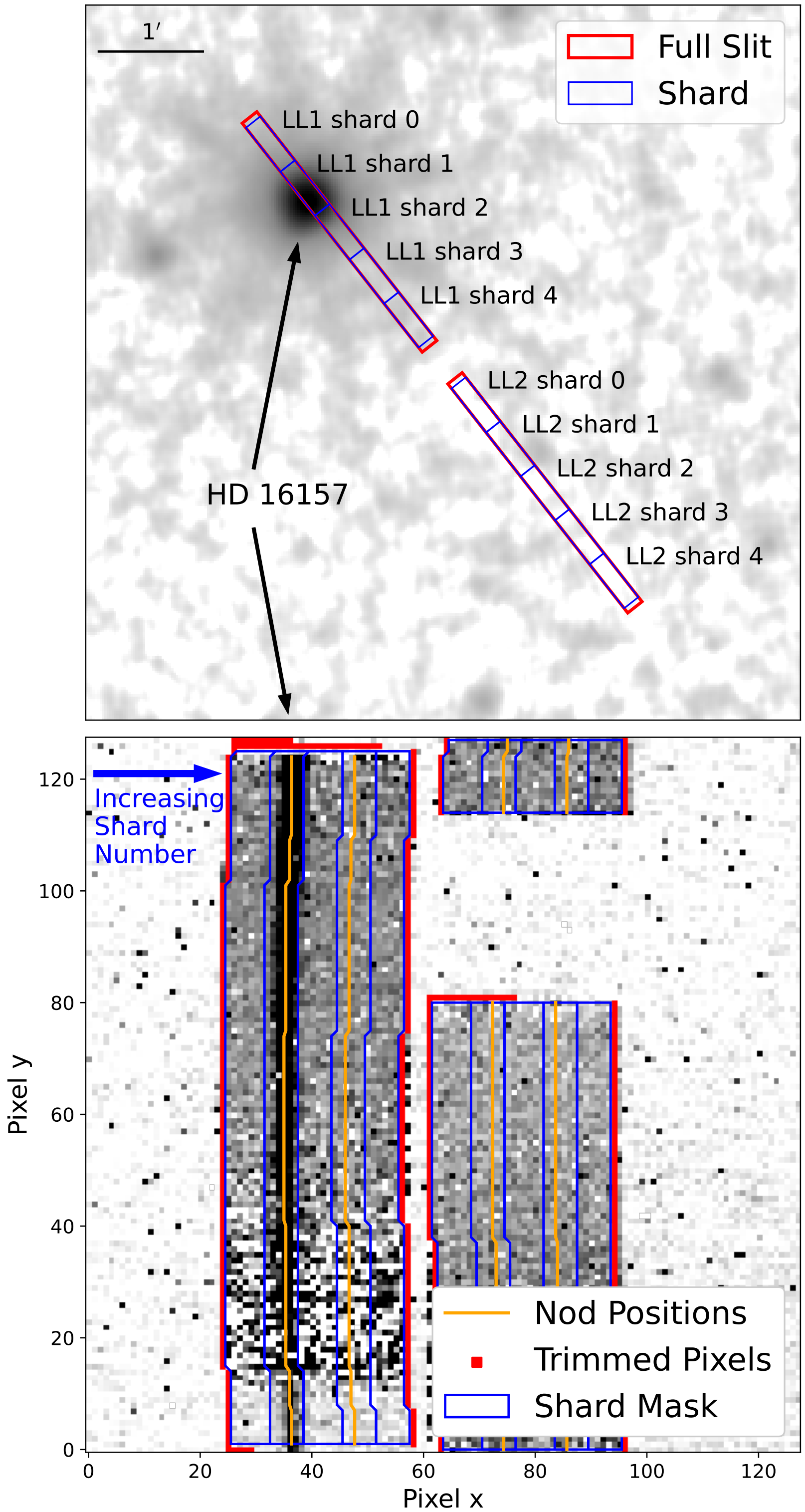

Figure 1. Adapted from G. P. Donnelly et al. (2025). Subdivisions of IRS slits, or shards, are used to increase the number of available background observations. This diagram shows how shards appear as sky apertures in the top panel, and as areas on the detector in the bottom panel. Note that even though this pointing integrates on a bright star, much of the image is suitable to be used as backgrounds.

2. DESCRIPTION OF DATA PRODUCTS

SIMLA data products are organized according to their program ID (PROGID) and area observation request key (AORKEY) assigned by the *Spitzer* pipeline and archive, following this directory structure: PROGID-###/AORKEY-###/, where "###" would be replaced with the one-to-five digit PROGID or the seven-to-eight digit AORKEY, respectively. Each data product has an associated AORKEY and a channel (denoted CHNLNUM, where SL=0, LL=2) in the file name. There is also an identifier (TMULTNUM, see §3) after the AORKEY to denote the cube number within this AOR; cubes with OBJTYPE = TargetMulti or TargetFixedCluster consist of multiple discontiguous pointings, so we have divided these AORs into multiple sub-cubes. For single-target AORs, TMULTNUM = 0. If the AOR has been split into multiple cubes, the smallest TMULTNUM for that AOR is 1. Thus, the general formula for file names follows: [AORKEY]-[TMULTNUM]_[subslit]_[product_type]. Each individual product type is described in the following subsections.

Note: Because all low-res IRS data include information from two spatially separated sub-slits, the cubes provided for each sub-slit (e.g SL1 and SL2 or LL1 and LL2) cover different fields of view; see top panel of Figure 1. These cubes may or may not have any overlapping area. However, the field of view for suborders 2 and 3 (which share a sub-slit) is always the same.

2.1. *Spectral Cubes*

SIMLA spectral cubes are similar to products from an integral field unit[3], in that they contain two spatial dimensions and one spectral dimension ($\sim 5.2 - 38.4\,\mu$m), however IRS spectral cubes are created by stepping the slit in parallel and perpendicular steps. This first SIMLA release corresponds to the low resolution $R \sim 60 - 130$ cubes, which comprise the majority of mapping mode observations.

- *_cube.fits: Background-subtracted and bad pixel-masked IRS spectral cube, in units of MJy sr^{-1}.
- *_cube_unc.fits: The uncertainty cube associated with the above, with the same data shape, field of view, and units (MJy sr^{-1}). See also §5.5.

2.2. *Inter-order Signal Corrected Cubes*

The inter-order (IO) signal is a spatially (in detector space) and temporally-varying artifact in IRS data that is most apparent in pixels between spectral orders on the detector, but affects the pixels on-order as well. This signal is especially significant in SL data, where it can appear as spectral offsets at the $\lesssim 1\,\mathrm{MJy\,sr^{-1}}$ level, see Figure 2. The standard SIMLA pipeline removes much of this signal via the background subtraction, but it can still be present in some cubes. The sl_io_correct code from C. A. Starkey (2016) mitigates the remaining IO artifact by sampling the pixels in the IO region of BCDs. SIMLA cubes for the short-low (SL) channel (SL1, SL2, and SL3) each come with separate versions that have had the inter-order signal artifact removed via this code. See §5.4 for more information.

- *_cube-iocorr.fits: SL SIMLA cube that has been corrected for the IO artifact.
- *_cube-iocorr_unc.fits: Uncertainty cubes for the IO-corrected SL cubes.

Note: The sl_io_correct code can not work properly for cubes consisting of only one BCD. For such cases, most commonly for some cubes with OBJTYPE = TargetMulti or TargetFixedCluster, we do not provide IO-corrected products.

2.3. *Extracted Spectra*

- *_spec.tbl: a 1D spectrum extracted over the entire field of view for the corresponding cube, in IPAC table format. Columns are lam: the wavelength in μm, Inu: the mean surface brightness in MJy sr^{-1}, and Inu_unc:, the propagated surface brightness uncertainty in MJy sr^{-1}.
- *_spec-iocorr.tbl: the 1D spectrum for an IO-corrected version of an SL cube, if present, in the same format as above.

[3] Observations of each slit step are *not* simultaneous.

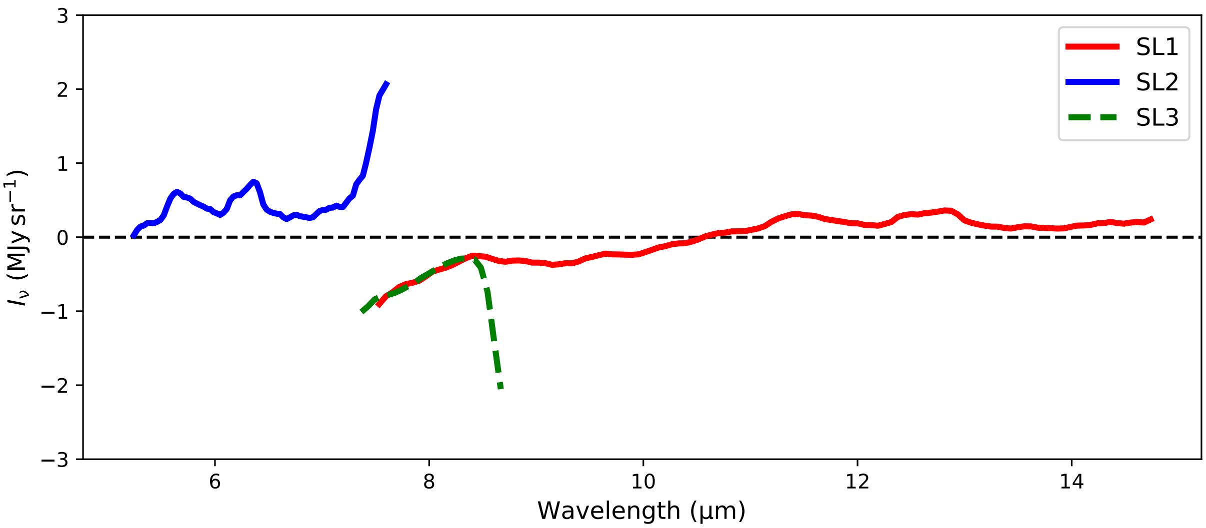

Figure 2. The typical shape of the inter-order signal artifact in SL spectra. Note the large discontinuity at the SL1-SL2 interface, which can lead to difficulties in spectral stitching. Consider using the IO-corrected version of an SL cube if a similar shape appears in the spectra of dark regions in the uncorrected cube, and/or there is difficulty in stitching spectra. The spectra presented here are the mean of the difference between uncorrected and IO-corrected SL cubes that have low or moderate flux in the "peak-up" fields.

2.4. *Moment Zero Maps (Spectrum-averaged Images)*

- `*_mom0.fits`: the corresponding cube is averaged (mean) over the spectral axis, making a moment 0 map, in units of MJy sr^{-1}. This is useful to see contents of the cube field of view at a glance.

2.5. *Dark Pixel Masks*

It is often desirable to extract dark spectra from a cube to help correct for remaining spectral offsets (see §5.3), evaluate noise within a cube, or check for the IO signal, among other things. To these ends, we provide masks that indicate pixels that were *only* contained within shards selected as dark by the conditions within the SIMLA pipeline.

- `*_darkmask.fits`: mask for cube pixels that are contained within dark shards, i.e., background sky regions used in cube backgrounds. Key: 1 if considered dark, 0 otherwise.

There are a few things to note about these files:

- Pixels in a dark mask are evaluated as "dark" or "not dark" based upon the same criteria that are used to select shards for backgrounds, so dark masks are subject to those same assumptions (see §1.1 and G. P. Donnelly et al. 2025).

- As not all cube fields of view have SIMLA-selected dark regions, many dark mask files will not have any dark pixels.

- The inverse of a dark mask should not be interpreted as a "bright mask," as noise within a shard spectrum or offsets in the WISE photometry may have caused an off-source shard to be rejected for use in backgrounds.

- Dark pixels indicate areas used as part of the background for the corresponding cube. If searching for signals within these dark regions, be aware that these pixels *contributed to their own background subtraction.*

2.6. *Cube Stats Table*

Finally, we also provide `cube_stats.csv`, which contains a table where each row corresponds to one built SIMLA cube, and columns are some basic cube properties. The columns are described in §3.

3. DESCRIPTION OF SIMLA-SPECIFIC KEYWORDS

The following are descriptions of the SIMLA-specific keywords that appear as columns of `cube_stats.csv` and/or as header keywords in `*_cube.fits` files.

- `SIMLAVER`: Version number of this SIMLA run (FITS headers only).
- `SAMPSPEC` and `SAMPSPAT`: Classifications that broadly describe the pixel sampling from slit steps in the spectral (i.e., perpendicular to the long side of the slit) or spatial (i.e., parallel to the long side of the slit) directions, respectively. The sampling for both directions can be classified as `ROBUST`, where step spacing results in no less than 10% overlap between apertures, `MARGINAL`, where the overlap is less than 10% but greater than 0%, and `SPARSE`, where there is no overlap between steps. If the number of parallel mapping steps is one (`STEPSPAR` = 1 in `*_cube.fits` headers), `SAMPSPAT` is set to `SPARSE`. For `SAMPSPEC`, we include a fourth category, `SLIT`, where there is no additional step in the perpendicular direction (`STEPSPER` = 1); see Figure 3. This fourth category calls attention to the fact that the resulting cube is similar to an IRS staring mode observation. Otherwise, `SLIT` and `SPARSE` describe the same sampling depth, and the calibration conditions assumed by `CUBISM` may not apply; see §5.6.
- `N_BCDS`: The number of BCD images that constitute the built cube. For cubes with `OBJTYPE = TargetMulti` or `TargetFixedCluster`, this can be different than the number of BCDs of this `CHNLNUM` in the corresponding `AORKEY`.
- `MEAN_MJD`: The mean modified Julian date of observations within the cube AOR.
- `ZODI12UM` and `ISM12UM`: Model intensities at 12 μm of the zodiacal and interstellar foregrounds/backgrounds from the T. Kelsall et al. (1998) model, respectively. In the SIMLA pipeline, we use the 12 μm intensity as a benchmark for these emission components for convenient comparisons to the WISE W3 12 μm band. Note that the zodiacal emission has been removed from each cube, but the Galactic ISM emission has not; see §5.1.
- `BG_RANK`: The mean quality rank across all 10 shard areas of the cube background, with a rank of one being the best, and four being the worst. The descriptions of background ranks are:
  1. Background shards come entirely from within the cube AOR and/or from a separate "dedicated background" AOR[4], and there are at least $q_{\rm targ}$ (see §1.1 and 4) shards.
  2. The shard background depth is equal to the target depth $q_{\rm targ}$, which is reached using shards from outside the cube AOR.
  3. The shard background depth is greater than zero but less than $q_{\rm targ}$.
  4. No usable dark shards were found within $\Delta z_{\rm max}$ and $\Delta t_{\rm max}$ (see §4).
- `BG_DZODI`: The mean of the differences between the 12 μm zodiacal intensity estimate from the ZEM for each shard used in the background, and the model-estimated 12 μm zodiacal intensity of the cube.
- `BG_DTIME`: The mean difference between the modified Julian date (MJD) of each shard used in the background and the mean MJD of cube observations.
- `BG_IN` and `BG_OUT`: The total number of background shards used in the cube background *across all shard areas* that come from the cube AOR, or come from outside the cube AOR[5], respectively.
- `TMULTNUM`: The index number for cubes from an `AORKEY`. If the AOR has a single target, `TMULTNUM` = 0. If the AOR has `OBJTYPE` = `TargetMulti` or `TargetFixedCluster`, the SIMLA pipeline splits the pointings into separate cubes, and `TMULTNUM` is a number 1 or greater.

[4] In many cases, observers took background observations using a different AOR from the cube AOR. The SIMLA pipeline attempts to identify these cases by finding background shards with the same `PROGID` and `RAMPTIME` as the cube, taken within one day of the cube, and that are within one degree separation of the cube. Such backgrounds are treated as if they are within the cube AOR for the purposes of background construction.

[5] Separate-AOR "dedicated backgrounds" are still counted as outside of the cube AOR here, so it is possible for a background to have `BG_RANK`=1 but `BG_IN`=0.

- `IOCORR`: T if this cube has been corrected for the IO artifact, F otherwise (FITS headers only).
- `PTYPE`: Identifier for the type of SIMLA product file (FITS headers only). `CUBE` for spectral cubes, `CUBEUNC` for spectral cube uncertainties, `DARKMASK` for the dark mask files, or `MOMENT0` for the moment 0 maps.

In addition to certain keywords taken directly from standard BCD file headers (e.g., `OBJNAME`, `AORKEY`, etc.) the `cube_stats.csv` file also contains these columns:

- `SUBORDER`: Built order number of the cube. For example, the suborders of an SL1 and a LL3 cube are 1 and 3, respectively.
- `BG_SHARD_RANKS`: Each integer is the rank of an individual shard area, with the first five belonging to sub-slit 1, the second five to sub-slit 2, and the shard indices for each sub-slit go from left to right as 0–4. These data are averaged to produce `BG_RANK`. For example, if `BG_SHARD_RANKS` = 1121113111 for an SL AOR, the SL1_2 shard has a rank of 2, and the SL2_1 shard has a rank of 3 (the first shard is index 0).

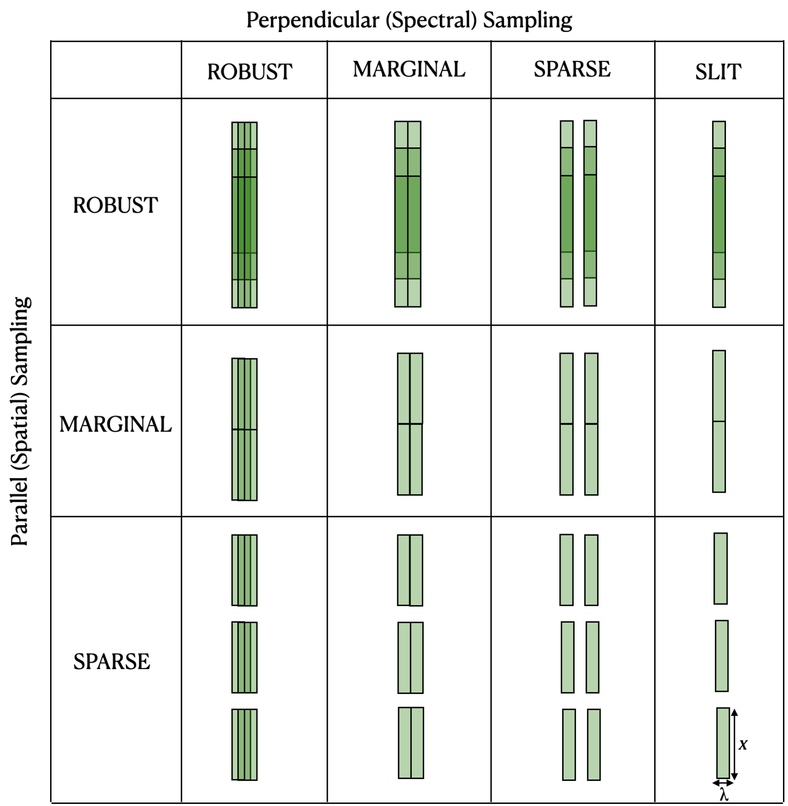

Figure 3. Diagram of the `SAMPSPEC` and `SAMPSPAT` classifications.

4. PIPELINE PARAMETERS

In Table 1, we give the values of various parameters that were used for the SIMLA cube building pipeline. `SIMLAVER` is the version number of the built cubes, C_{W} is the cut on the shard-aperture background-subtracted WISE W3 photometry, C_{s} is the cut on the median of the zodiacal emission and baseline-subtracted IRS spectra extracted from shards, C_{ISM} is the cut on the modeled intensity in a shard originating from the Galactic ISM at 12 μm, Δt_{max} is the maximum allowed time between a shard observation and the mean MJD of a given cube AOR, similarly, Δz_{max} is the maximum allowed difference in the model estimate of the zodiacal intensity at 12 μm between a shard and that of the given cube AOR, q_{targ} is the target depth of shards in the background stack, i.e., the depth necessary to achieve a `BG_RANK` of 1 or 2 (see §3), and σ_{trim} is the cut threshold during the pixel trimming of the background stack. See G. P. Donnelly et al. (2025) for more details on how each parameter is used.

Table 1. Used Pipeline Parameters

SIMLAVER	C_{W}	C_{s}	C_{ISM}	Δt_{max}	Δz_{max}	q_{targ}	σ_{trim}
1.0	$\pm 0.1\,\mathrm{MJy\,sr^{-1}}$	$\pm 2\,\mathrm{MJy\,sr^{-1}}$	$0.5\,\mathrm{MJy\,sr^{-1}}$	10 days	$10\,\mathrm{MJy\,sr^{-1}}$	50	1.5

5. NUANCES, RECOMMENDATIONS, AND CAVEATS

5.1. *What counts as background emission?*

Background subtraction is, in part, intended to remove real signal that is not related to the astrophysical target of the cube. This forces a choice about which components of emission should be considered. For SIMLA cubes, the zodiacal light arising from dust in the Solar System is the only astrophysical emission that is removed. Importantly, we do *not* include emission from the Milky Way interstellar medium in the background model, so as to preserve this signal for interested users. The consequence of this is that Galactic targets and extragalactic targets near the Galactic plane may have a contribution from Milky Way "foreground." If the science case requires removing such foregrounds, we recommend that users extract a spectrum from an off-target part of the cube, and then subtract this spectrum in 1D from cube extractions.

5.2. *Surface brightness limit for background selection*

The SIMLA pipeline selects IRS observations to be used as backgrounds based in part on an aperture-matched and background-subtracted WISE photometry cut at $\pm C_{\mathrm{W}} = 0.1\,\mathrm{MJy\,sr^{-1}}$; any apertures with fluxes that fall below this cut may be used as backgrounds if they pass the other cuts. Although this cut is near the WISE sensitivity limit, there may be cases where corresponding IRS observations are sufficiently deep that there is real signal below the level of the cut, and this signal becomes a component of the background. Such cases would appear as an inverted (negative) spectrum in parts of a cube that are emission-free, and less obviously as a slightly altered spectral shape elsewhere. The practical implication of this is that the WISE cut sets a limit below which care should be exercised when interpreting fluxes and features in SIMLA cubes.

5.3. *Correcting for small surface brightness offsets*

Some cubes contain spectrophotometric offsets on the order of $\lesssim 0.5\,\mathrm{MJy\ sr^{-1}}$. These offsets can arise from a mixture of background differences, zodiacal model errors, and uncertainty in the baseline calibration. Such offsets may be present in any cube, and can be identified in cubes that have dark regions within them by extracting a spectrum from these regions and checking for a positive or negative flux offset. This offset may not be constant with wavelength, but it should be consistent spatially within the cube. In the case where a cube has a dark region, i.e., it has pixels with a value of 1 in its associated `*_darkmask.fits` file, the offset can be corrected by subtracting a spectrum extracted from the dark mask region.

5.4. *Recommendations Regarding the SL Inter-Order Artifact*

While the IO correction can often improve the shape of extracted SL spectra, in particular where orders meet, it may also add noise. We recommend that the decision to use IO-corrected cubes is made on a case-by-case basis after inspecting both versions. Two tests that can be done to determine whether significant IO signal is present are: 1) extracting a spectrum from a dark region within a cube, if possible (see §2.5), and looking for the IO artifact in the spectrum (e.g. Figure 2) and 2) looking at the spectral agreement at the SL1/SL2 interface, as the IO artifact causes stitching problems. We have also found that `sl_io_correct` often positively impacts the spectral shape of cubes for which the peak-up (PU) fields were pointed at very bright sources, by removing stray/scattered PU light. In such exceptional cases, the removed signal can be as bright as a few hundred $\mathrm{MJy\,sr^{-1}}$, scaling with the intensity of the PU-observed emission; see Figure 4.

5.5. *Uncertainty Estimates*

The formal uncertainties in `*_unc.fits` cubes (and therefore in extracted spectra) are primarily per-pixel estimates created from measurements of noise in the ramp readouts made during IRS integrations. However, there are other

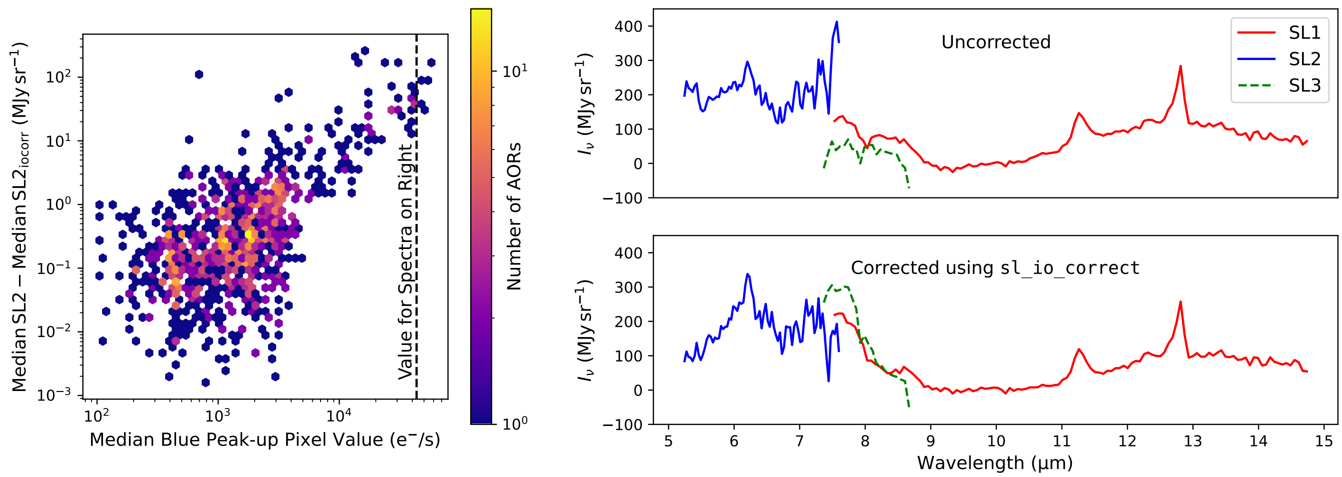

Figure 4. Left: The difference of full-spectrum medians for SL2 (which is near the PU fields in detector-space) spectra with and without a correction via `sl_io_correct` is larger for AORs with high peak-up flux, indicating that stray/scattered light from the peak-up fields is being corrected for. Right, we show spectra for an example with unusually high brightness and contamination from the peak-up region, before (top) and after (bottom) the correction. The agreement between the spectral orders is significantly improved after the correction, and emission features from polycyclic aromatic hydrocarbons can be seen.

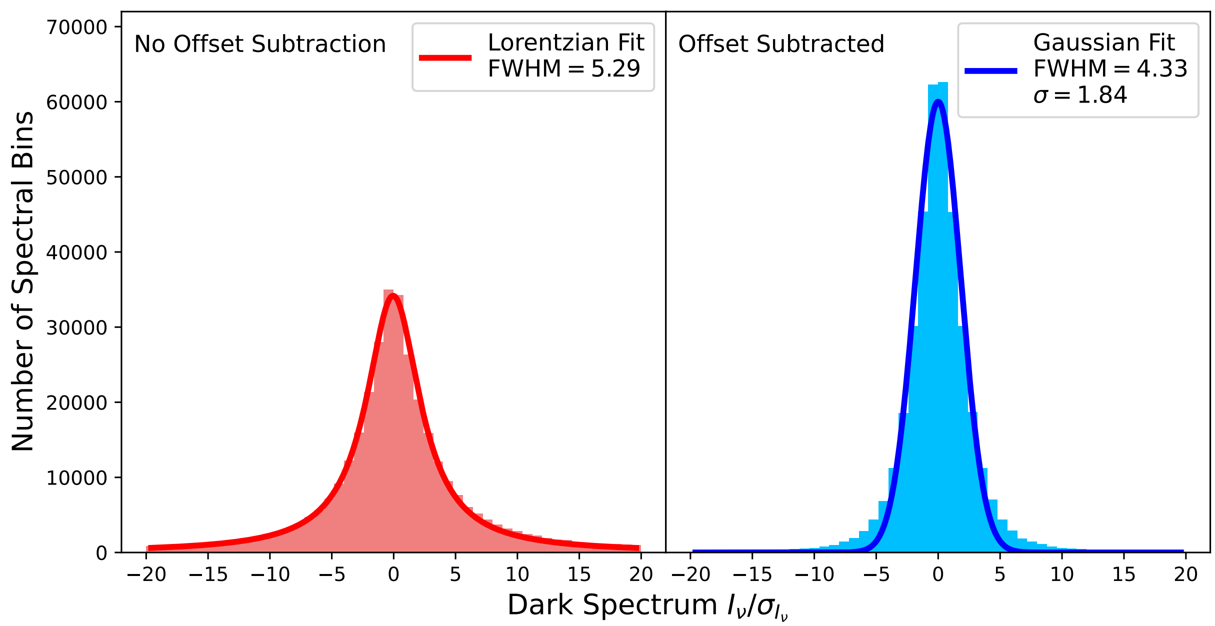

Figure 5. Distributions of surface brightness relative to the formal uncertainty for spectra extracted using `*_darkmask.fits` files (see §2.5). On the left, there is no accounting for spectrophotometric offsets (see §5.3), indicating that these offsets can be much larger than the formal uncertainties. On the right, we have corrected for smooth offsets by subtracting a second-order polynomial fit from the spectra before dividing by the uncertainty. The Gaussian shape that emerges indicates close correspondence between statistical pixel and cube-level uncertainties for offset-subtracted spectra. See §5.5.

sources of uncertainty from the initial BCD reduction and the SIMLA pipeline that would increase these estimates. In Figure 5, we show the surface brightness distribution of spectra from dark regions (via the dark masks) relative to the reported uncertainty. If the BCD ramp readout noise constituted the only noise present in cubes, this would resemble a normal distribution, i.e., a Gaussian with $\sigma = 1$ (FWHM = 2.35). Spectrophotometric offsets (which can often be corrected for at the cube level, see Section 5.3) produce larger wings in the distribution, visible in the left panel, showing that these offsets can be significant compared to per-pixel statistical uncertainties. If these offsets are removed as in the right panel, the expected Gaussian shape emerges, indicating that the formal uncertainties are well-correlated with the offset-subtracted dark spectrum noise.

The Gaussian in the right panel has $\sigma = 1.84$, so reported uncertainties are best interpreted as a lower limiting uncertainty estimate; across SIMLA, a more conservative uncertainty estimate would on average be $\sim 1.84\times$ larger. However, for an individual cube, this increase depends on several factors like `RAMPTIME`, source brightness, and module; see Figure 17 in G. P. Donnelly et al. (2025). If a more conservative estimate is desired for the uncertainty of a particular cube, and a dark spectrum can be extracted from the cube, a similar process as described above could be applied to derive an appropriate uncertainty increase on a cube-by-cube basis.

5.6. *Caveats*

- In rare cases, cubes may be affected by saturation or other artifacts from the original IRS observations. We did not flag or remove these cubes from SIMLA. It is always best to visually inspect cubes before analyzing spectra.
- SIMLA cubes are built using `CUBISM`, which assumes a spatially sampled slit map. Thus, for SIMLA maps that have a single sample in the spectral direction, i.e., covering effectively the width of one slit (`SAMPSPEC` = `SLIT` or `SPARSE`), cube rows will be highly correlated along the direction perpendicular to the slit. In these cases, absolute integrated flux values may be unreliable by up to a factor of 2.
- The pixel-level uncertainty estimates are propagated using standard methods that assume uncorrelated uncertainties. However, given that any pixel in a cube is a projection of multiple pixels on a BCD, uncertainties for adjacent cube pixels are not entirely uncorrelated.
- The SL detector features peak-up (PU) fields adjacent to the spectral orders. In some rare cases where a PU field covered a source bright enough to saturate there, there may be positional and wavelength-dependent discontinuities within the resulting cube (unrelated to order mismatch where order segments meet/end). As LL BCDs do not include peak-up regions, these issues does not effect those cubes.
- Despite the background subtraction and the automatic flagging done by `CUBISM`, residual bad pixels may effect some cubes. Spectral mapping was performed by stepping a slit across the sky. Therefore, persistent bad pixels can leave striping pattern artifacts in some cubes.

6. SIMLA REFERENCE

Published use of SIMLA data products should reference the main SIMLA paper: G. P. Donnelly et al. (2025).

REFERENCES

Donnelly, G. P., Whitcomb, C. M., Hands, L., et al. 2025, arXiv e-prints, arXiv:2512.12434, doi: 10.48550/arXiv.2512.12434

IRS Instrument Team, & Science User Support Team. 2021, IRS Instrument Handbook,, NASA IPAC DataSet, IRSA487 doi: 10.26131/IRSA487

Kelsall, T., Weiland, J. L., Franz, B. A., et al. 1998, ApJ, 508, 44, doi: 10.1086/306380

Smith, J. D. T., Armus, L., Dale, D. A., et al. 2007, PASP, 119, 1133, doi: 10.1086/522634

Starkey, C. A. 2016, PhD thesis, University of Toledo, Ohio