Grizli Visual Checking README

Date: 02/23/2024

Installation:

- python
- ds9 and X11 (This is the one we found works. There could be other ways to install ds9)
    - install X11 from here
    - download X11 version of ds9 from here
    - Once you download ds9 package, you can put them in a folder. And add Path to that folder in the bash or zshrc script. For example, I put ds9 and ds9.zip in "/Users/lushen/.bin". Then, in my .zshrc file, I add a line: *export PATH="$PATH:/Users/lushen/.bin"*
- Pyds9 (here)

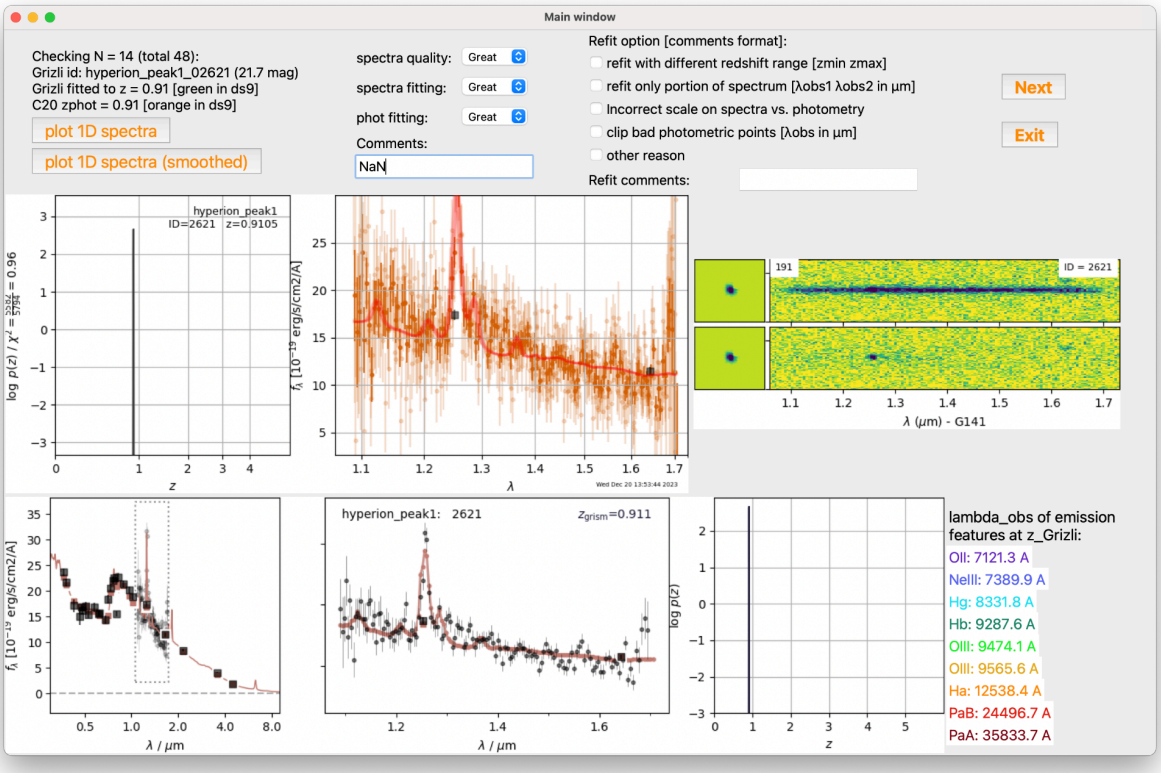
Running the GUI:

- python check_results.py -i <directory/inputfile> -n <start number> -l <grilzi id list>
    - The inputfile is the raw catalog containing objects to be classified in a given work package
    - All related files (i.e., *.png, *.fits) should be placed in the same directory as the inputfile.
    - When you start, click one of the dropdown menus for the first object before clicking next to prevent the GUI from crashing.
- Example:
    - Run the GUI from beginning of a catalog: **python check_results.py -i beta_test/beta_test.cat**

- Start from the 5th object in the input catalog: **python check_results.py -i beta_test/beta_test.cat -n 5**
- Running the GUI with only two objects, e.g., grizli id = "hyperion_peak1_13353" and "hyperion_peak1_09985: **python check_results.py -i beta_test/beta_test.cat -l '["hyperion_peak1_13353","hyperion_peak1_09985"]'**
- If both "start number" and "grizli id list" are given, "grizli id list" will be used.
- The input flags and comments will save in every 2 objects to the save input file. This is to prevent loss due to exit for any unknown reason. If the code exit itself, you can re-start 2 objects before the object you were working on to check.

Gui Windows:

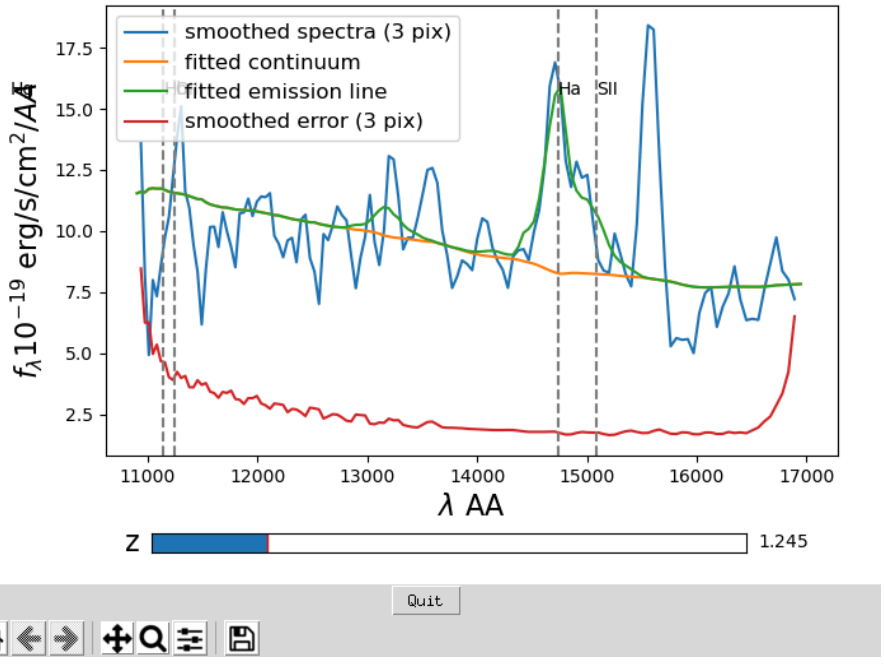
- **Main window:**



-
  - **Information about the objects (top left):**

- The number of this object in the input file and the total number of objects in the input file;
- Grizli id of this object and F160W/F140W AB magnitude
- Grizli best fitted redshift
- Photometric redshift from COSMOS2020 with the Classic photometry method and using EAzY;

- **Button "plot 1D spectra"**, click it to show the 1D spectra window ;
- **Button "plot 1D spectra (smooth)"** smoothed at 3 spectral pixels, which is approximately a matched filter to the G141 resolution (see example below). This window also contains a slider bar to set the redshift. The value on the slider bar is originally set to the Grizli fitted redshift, but can be slid to show other redshift solutions. The observed-frame location of various emission lines (most of those used by Grizli to do the redshift fit as well as few others) will be overplotted for a given redshift solution. If you have issues dragging the slider, click along the slider bar to select a redshift instead. Note that some features included in the 1d spectral fit that is performed *after* the redshift fit are not included in the displayed line list (see notes on Grizli redshift fitting near the end of this document).



- **flagging system:**

- "spectra quality": the quality of spectra. This quality is largely decoupled from our ability to recover a redshift from the spectrum and speaks mostly to the data quality. This quality can be determined from the 2d spectra plot in the main window or ds9 window as well as the 1d spectral plot.
    - "Great": objects with clear continuum and some emission line features, with no overlapping contamination or other reduction artifacts (spectrum cutoff on one end, hot pixels, etc.).
    - "Good": spectrum contains (1) one or more emission lines with no/faint continuum OR (2) decent S/N continuum (based on a visual inspection) without emission line features. No strong contamination or other reduction artifacts present.
    - "Unclear": (1) objects with low S/N continuum or no continuum and no emission lines, but contains no strong contamination or other reduction artifacts, OR (2) has obvious contamination in some wavelengths but a clear emission feature or decent S/N featureless continuum in an uncontaminated area (a note put in the comments)
    - "Bad": objects with contamination in the extraction window or severe reduction artifacts, which includes more than half of the spectral coverage being cutoff. Note that contamination is typically shown as very dark green in the 2d spectra plot.
    - Note that if the continuum is not fitted well (i.e., not subtracted in the continuum subtracted or over-subtracted), but visually there is spectrum, this is spectra fitting issue, not spectra quality issue and the spectral quality should be flagged accordingly. Example: #8 in the beta_test.
    - Even if there is contamination near the extraction window in the 2d spectrum, as long as it is not spatially coincident with the extraction window (or at the very edge where it is severely downweighted), it is fine. Can mark “co” (contamination offset from window) in the comments in this

case. If there is contamination within the extraction window, mark “cw” (contamination in the window) in the comment.

- "spectra fitting": flag on the spectra, can be determined from (smoothed) spectra 1d plots, 2d spectral window in ds9, and the middle panels of 1st and 2nd row on the main GUI. See below about “Note on Grizli Redshift fitting”
    - "Great": spectrum continuum and emission lines are perfectly fitted and there is at least one clear feature fit;
    - "Good": a good fit (1) two possible emission features are fitted, but a high S/N continuum is present and is clearly not fit well. (2) a single feature is fit well and it's pretty clear the redshift is correct, but there appears other features in the grism spectra that may be real or continuum is not fitted or artifacts that are not accounted for or not accounted for well (3) a reasonably high S/N continuum (>~1-2) with a continuum-only fit when no strong emission lines are expected (these are Ha, OIII, Hb, OII in some cases, PaB in some cases) based on physics-based arguments either due to the rest-frame coverage of the grism or a red template fit ;
    - "Unclear" (1) only a continuum without features is fit and fit well, but it’s unclear whether emission features should be present, or (2) low S/N continuum where the presence of lower S/N features is difficult to determine, or (3) very noisy spectra fall in this category as well (as long as the model isn’t terribly off).
    - "Bad": emission lines and/or continuum obviously wrong.
- "phot fitting": flag on the SED photometric fit, shown in the left panel of 2nd row.
    - "Great": photometric data (black dots) are perfectly fitted (within normal scatter and any issues with F814W, see below).
    - "Good": Some photometric data points are significantly off (larger than data errorbar) from the fitted SED model, more than may be expected from typical Gaussian random errors.

Note that F814W is off in some cases, because F814W is not extracted the same way as other photometric data.

- "Unclear": cannot see the fitting result due to large errorbars in the spectra OR you think there may be another fit that might work equally well;
- "Bad": obviously wrong, a majority of the photometric data is off from the fitted SED model;

- "spec+phot fitting": this will be calculated automatically from your above inputs. Indicates the quality of the overall spectra and SED photometry fitting.
  - "Great": both “phot fitting” and “spec fitting” flags are ‘great’
  - "Good": both “phot fitting” and “spec fitting” flags are ‘good’ or one of the two is ‘good’ and the other is ‘great’
  - "Unclear": both “phot fitting” and “spec fitting” flags are ‘unclear’ or one of the two is ‘unclear’ and the other is ‘good/great’
  - "Bad": “phot fitting” and/or “spec fitting” flags are ‘bad’
- "comments:" put any comments/notes/reasons for the flagging here. *See standard comment codes at the end of this document. Note: commas should not be entered in the comments box as the final catalog entries are parsed by commas.*

- **re-fit flagging system:**
  - There are five refit options, of which you can choose more than one if relevant:
    - "refit with different redshift range", if there is emission line identified at another redshift, put the redshift range in the "refit comments" box such as "[2.0 3.0]";
    - “refit only portion of spectrum”, if part of the spectrum is contaminated or very noisy, only consider a range of wavelengths in the fit, put range in um in “refit comments” box, such as “[1.3 1.7]”;
    - "Incorrect scale on spectra vs. photometry" should be selected if the 1D spectra is offset in y (flux density) from the photometric data points at a common wavelength in the 1st and 2nd row middle panels;

- "bad photometric points", if there is any bad photometry data point in the SED fitting panel, put the wavelength in unit of micro-m in the "refit comments" box, such as "[0.8]". Only flag if there are severe offsets that may affect the fit (and can ignore F814W, see previous note);
- For other reasons, put a description in the "refit comments" box and check the “other reason” option;

■ *Note: do not put commas in the "refit comments" box, because the final catalog entries are parsed by commas.*

○ **Button next or exit**:

■ Next button will go to the next object. The input flags and comments will save in every two objects. This is to prevent loss of entries due to a crash or other issue.

■ Exit button will exit the GUI, save the current results, and print the number of the object that you were looking at when you exited the program.

○ **1st row plots (from left to right):**

■ **Left:** The marginalized redshift probability distribution function, p(z)

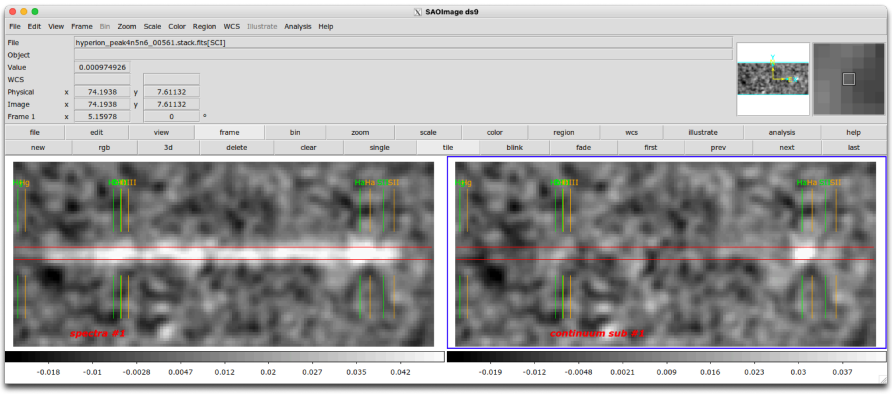
■ **Center:** The stacked 1D spectrum from all exposures is shown as orange dots with error bars. The 1D spectra from single exposures are shown in light orange. The best-fitted spectrum is shown in red lines. The black squares are photometric data within the spectra wavelength range.

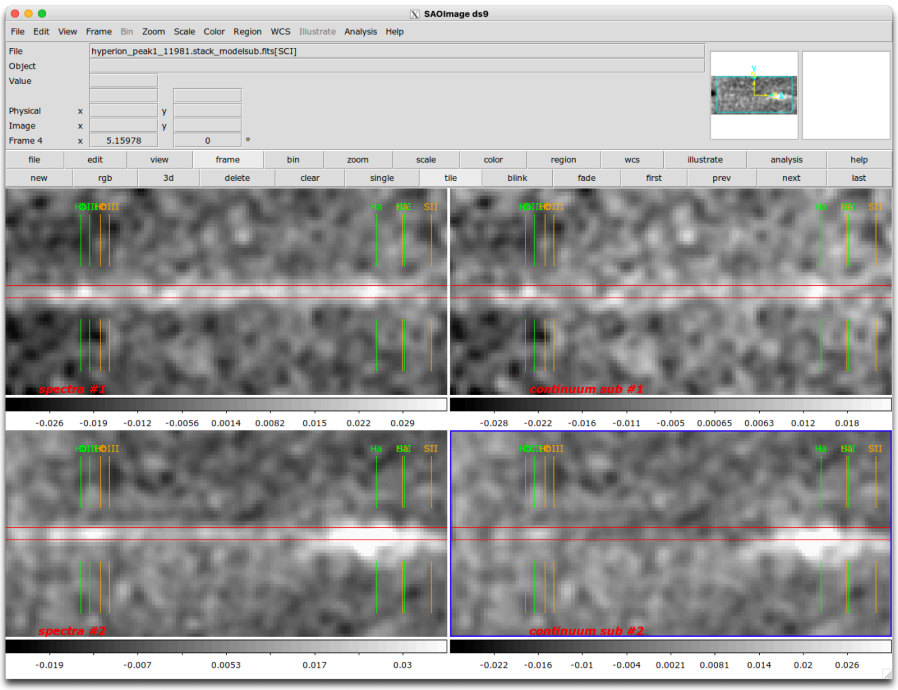
■ **Right:** *Left panels:* The objects in the direct image; *Right panels:* Top one or two panels: the stacked 2D spectra from each PA. The PAs are listed in the upper left corner. *Bottom:* the stacked 2d spectra from all PA with best-fitted continuum subtracted.

○ **2nd row plots (from left to right):**

■ **Left:** The SED photometric + spectral fitting : black squares are photometric data, maroon squares are the synthetic COSMOS2020 photometry as measured on the best-fit template, and the maroon line is the best-fit SED model. The 1d spectrum is also fit as part of the SED fitting process and is shown as gray dots in the gray dotted box.

■ **Center:** Zoom in on the 1D spectrum.

        - **Right:** The marginalized redshift probability distribution function, p(z);
- **ds9 window:**
    - displays the 2d spectra of this object. See below for two examples: the first example object has a single spectrum from a single roll angle (1 row), and the second example has two spectra, each of the same object, taken from two different roll angles (2 rows). For each row, the left shows the contamination-subtracted grism spectrum and the right shows the contamination-subtracted spectrum with a continuum model also subtracted. The orange vertical lines indicate potential features at their observed-frame location assuming the COSMOS2020 photo-z, the green vertical lines indicate the same for the Grizli best-fit redshift. The red horizontal line indicates the 1d extraction window as determined by a Gaussian measurement of the 1d collapsed kernel spatial profile (5th extension of the *.stack.fits MEF files). The bounds of the extraction window are set by the mean and FWHM of fit profile;
    - 

- 

Note on Grizli Redshift fitting

Grizli has two steps of spectra fitting:

(1) To determine a redshift, grizli uses 12 Flexible Stellar Population Synthesis (FSPS) templates, a simple stellar population (SSP) template, and four templates that contain a combination of lines (see the below figure). In each line template, the line fluxes ratios are fixed to the values given here (look for the array values

for the templates that have emission lines corresponding to the below entries).

```
Template   0: fsps/fsps_QSF_12_v3_nolines_001.dat (NZ=1).
Template   1: fsps/fsps_QSF_12_v3_nolines_002.dat (NZ=1).
Template   2: fsps/fsps_QSF_12_v3_nolines_003.dat (NZ=1).
Template   3: fsps/fsps_QSF_12_v3_nolines_004.dat (NZ=1).
Template   4: fsps/fsps_QSF_12_v3_nolines_005.dat (NZ=1).
Template   5: fsps/fsps_QSF_12_v3_nolines_006.dat (NZ=1).
Template   6: fsps/fsps_QSF_12_v3_nolines_007.dat (NZ=1).
Template   7: fsps/fsps_QSF_12_v3_nolines_008.dat (NZ=1).
Template   8: fsps/fsps_QSF_12_v3_nolines_009.dat (NZ=1).
Template   9: fsps/fsps_QSF_12_v3_nolines_010.dat (NZ=1).
Template  10: fsps/fsps_QSF_12_v3_nolines_011.dat (NZ=1).
Template  11: fsps/fsps_QSF_12_v3_nolines_012.dat (NZ=1).
Template  12: alf_SSP.dat (NZ=1).
Template  13: line Ha+NII+SII+SIII+He+PaB (NZ=1).
Template  14: line OIII+Hb+Hg+Hd (NZ=1).
Template  15: line OII+Ne (NZ=1).
Template  16: line Gal-UV-lines (NZ=1).
```

(2) Then, Grizli will fix the redshift to the best-fit redshift and fit the grism and photometric data with 12 FSPS templates, 1 SSP template and 13 emission line templates. See below the 13 emission lines that are considered in this step of the Grizli fitting process. This step will measure the line fluxes. Note that at this stage, emission line ratios are not constrained to correspond to physical values, and, as a result, many fits will result in line ratios that violate physics.

```
Template  13: line BrA (NZ=1).
Template  14: line BrB (NZ=1).
Template  15: line BrG (NZ=1).
Template  16: line PfG (NZ=1).
Template  17: line PfD (NZ=1).
Template  18: line PaA (NZ=1).
Template  19: line PaB (NZ=1).
Template  20: line PaG (NZ=1).
Template  21: line PaD (NZ=1).
Template  22: line HeI-1083 (NZ=1).
Template  23: line SIII (NZ=1).
Template  24: line OII-7325 (NZ=1).
Template  25: line ArIII-7138 (NZ=1).
Template  26: line SII (NZ=1).
Template  27: line Ha (NZ=1).
Template  28: line OI-6302 (NZ=1).
Template  29: line HeI-5877 (NZ=1).
Template  30: line OIII (NZ=1).
Template  31: line Hb (NZ=1).
Template  32: line OIII-4363 (NZ=1).
Template  33: line Hg (NZ=1).
Template  34: line Hd (NZ=1).
Template  35: line H7 (NZ=1).
Template  36: line H8 (NZ=1).
Template  37: line H9 (NZ=1).
Template  38: line H10 (NZ=1).
Template  39: line NeIII-3867 (NZ=1).
Template  40: line OII (NZ=1).
Template  41: line NeVI-3426 (NZ=1).
Template  42: line NeV-3346 (NZ=1).
Template  43: line MgII (NZ=1).
Template  44: line CIV-1549 (NZ=1).
Template  45: line CIII-1906 (NZ=1).
Template  46: line CIII-1908 (NZ=1).
Template  47: line OIII-1663 (NZ=1).
Template  48: line HeII-1640 (NZ=1).
Template  49: line NIII-1750 (NZ=1).
Template  50: line NIV-1487 (NZ=1).
Template  51: line NV-1240 (NZ=1).
Template  52: line Lya (NZ=1).
```

Note that stellar absorption (for example, absorption at the wavelengths of Ha or Hb due to stellar photospheric absorption) are accounted for in Grizli as a result of using the FSPS and SSP templates.

In the 1d spectra window (smoothed) and ds9, we marked lines that are included in the four combined line templates from step 1. Also we plot “Strong” lines ('OII', 'NeIII', 'Hg', 'Hb','OIII', 'Ha', 'PaB', 'PaA') in darker grey and “Weak” lines in light grey. Strong lines are more likely to be observed as compared to Weak lines and Strong lines make up the bulk of the emission lines that go into the initial fitting process. Many of the Strong lines are also listed in the bottom right of the main GUI. *Generally the presence/absence of a feature at a line designated as Weak should not be given a lot of credence when determining flags except when a feature is clearly present at high S/N*. You might see other fitted emission line features that are not marked, those are even weaker lines that are considered only in step 2. This does not mean they are not real, if they are observed (6300A [OI]) is a good example) but you should be careful.

List of emission line (Strong lines are marked in bold and blue. Combinations of emission lines in step 1 are marked in different colors.)

Emission line	Wavelength (A)
Ly-alpha	1215.4
N V	1240.8
N IV	1487
C IV	1549.48
Si IV+OIV	1398
He II	1640.4
OIII-1663	1663
N III	1750
CIII-1906 CIII-1908	1906.7 1908.7
Mg II	2799.1
Ne V	3343.5
Ne VI	3426.9
OII	**3727.1, 3729.9**
NeIII	**3869.87**

H10	3799.014
H9	3836.511
H8	3890.191
H7	3971.236
H-delta	4102.936
H-gamma	4341.731
OIII-4363	4364.436
H-beta	**4862.738**
OIII	**5006.8, 4960.3**
HeI-5877	5877.2
OI-6302	6302.0
NII	6549.86, 6585.27
Ha	**6564.697**
SII	6718.29, 6732.67
ArIII-7138	7137.8
OII-7325	7321.9, 7332.2
SIII	9071.1, 9533.2
HeI-1083	10832.1, 10833.3
Pa-Delta	10052.2
Pa-Gamma	10941.2
Pa-Beta	**12821.7**
Pa-Alpha	**18756.3**
Pf-Delta	32970.0
Pf-Gamma	37405.7
Br-Gamma	21661.3
Br-Beta	26258.8
Br-Alpha	40522.8

List of Comment codes

Comments are on a best-effort basis. Try your best to include at least the most important comments, don't worry about covering everything. These comments are primarily going to be used as:

1) **a guide for galaxies with good or great spec+phot flags to parse for features and any issues with the data quality when measuring spectral quantities/stacking**
2) **a guide for galaxies that may have salvageable redshifts on a re-fit**

Beyond that, the codes will be rarely used. Please keep this in mind when you're classifying so that you can conserve your energy for the long haul. E.g., if the flags are bad, bad, bad, for a given object and there's no re-fit that will be attempted/no way the data can be salvaged, you can choose to save your energy and not comment beyond maybe "su" or "cw_full" (see below)

- All comment codes below can be used either generally, meaning they apply to the whole observation, or for one PA (if there are observations at two different roll angles). For the latter, use: "**PA_XXX_commentcode**" where XXX is the PA as given in the upper left hand portion of the 2d spectra shown on the main GUI and commentcode = one of the comment codes below, e.g., "**PA_60_cw_1.5**".
- "**co**" = "contamination offset from window" when there is contamination spatially offset from the extraction window or near the edge where it's severely downweighted.
- "**cw_XXX**" = "contamination in the window" contamination spatially coincident with the extraction window. XXX gives information on where the contamination occurs. If the spectrum is contaminated by a single emission or multiple narrow emission features, XXX gives the wavelengths in microns (e.g., **cw_1.5**, **cw_1.5_1.7**). If the contamination in the entire occurs in a near continuous manner blueward/redward of a given wavelength, XXX indicates the spectral region where the contamination occurs (e.g., **cw_b1.5** for contamination blueward of 1.5 um, **cw_r1.5** for contamination redward of 1.5 um). **cw_full** for full spectral window contamination.

- “**cut_XXX**” = “spectrum cutoff” where XXX indicates the cutoff wavelength and whether the cutoff is blueward or redward of this wavelength (e.g., **cut_b1.5** for a spectrum that shows no data blueward of 1.5 um)
- "**su**" = "spectrum unusable" for when the spectrum is hopelessly contaminated or there are other issues that prevent the spectrum from being able to provide any reliable constraints
- "**ss**" = "same spectrum” for when the continuum subtracted spectrum is the same as the original continuum
- “**supser**” = “superimposed serendipitous detection” for when there appears to be a second object within the extraction window in addition to the main source. If you have a redshift guess, use “**supser_zXX**” (e.g., supser_z2.7)
- “**Ha**” “**SII**” “**OIII**” “**Hb**” “**Hg**” “**Hd**” “**NeIII**” “**OII**” “**PaB**” “**cont_break**” “**cont**” = indicates that a given feature is present in the spectrum (cont_break is for the Balmer/4000A break, cont indicates a strong continuum). Multiple features can be combined by “**/**”, e.g., “**Hb/OIII**”. Use **“cont”** only when there are no other features, but you’re marking the spectral quality and fit flag “Good”
- “**marg_**” = is used as a modifier for a feature or a contamination to indicate a weak feature/contaminate, e.g., **“OIII/marg_Hb”**
- “**wide_”** = is used when a feature is larger than the line spread function of the WFC3/G141 grism (i.e., ~>100A), e.g., **“OIII/wide_Hb”**. This will indicate a type-1 AGN.
- **“ex_”** = is used to indicate that there is spatially extended emission past the extraction window, e.g., **“OI/ex_Ha”**
- **“!”** or **“!!”** = If any feature or issue is particularly strong, add emphasis. For example, **“cw_1.5!!”** or **“Hb/OIII!/Ha”**.
- **“ef_XXX”** = “possible emission line present was not fit” where XXX indicates the wavelength of emission line in micron
- **“ppp”** = “photometry fitting plot problem where dynamic range is too large and cannot assess fit”
- **“oc_XXX” =** “oversubtracted continuum”, for when the model is a little overzealous. Can be combined with _XXX indicating the region where the continuum is oversubtracted (i.e., **oc_b1.6** for when the continuum is oversubtracted blueward of 1.6um)
- **“uc_XXX”** = “undersubtracted continuum”, as **“oc”** above but for cases where the continuum model leaves residual continuum emission.

- **“zc”** = redshift is correct irrespective of flags, put this when any flag is bad/unclear but you still believe the redshift should be used
- **“borderline_sq_XXX”/”borderline_sf_XXX”/”borderline_pf_XXX”** = to be used when there is a case where you are at the border between two Spectra Quality (sq), Spectra Fitting (sf), or Photometric Fitting (pf) flags, indicate the flag you didn’t mark in the XXX. E.g., if you thought the spectral fitting is just barely on the good side of the Good/Unclear divide, mark “Good” for the Spectra Fitting flag and write “borderline_sf_unclear”.
- **Anything else?** If you have additional notes that don’t follow one of the above comment codes, put it in! If you find yourself typing something many times, let everyone know and we can add it as a standard comment code
- *Remember: commas should not be entered in the comments box as the final catalog entries are parsed by commas.*