Versatile Emission Line Fitting Package Guideline

Zixuan Peng[1]

[1] *zixuanpeng@ucsb.edu*
[1]Department of Physics, University of California, Santa Barbara

1 Installation

For installing the `astroconda-lmfit` environment that is required to run the line-fitting algorithm, please follow these steps in the terminal:

1. Given the environment definition YML file (i.e., `environment.yml`), users can build the environment with:
   `conda env create -f environment.yml`
2. Activating an environment: Once you have built the environment, users can activate it with:
   `conda activate astroconda-lmfit`
3. To deactivate the environment, run:
   `conda deactivate`

Optional:

1. To delete the environment with all of its packages, run:
   `conda env remove -n <name>`
2. To rename the conda environment, run:
   `conda create --name new_name --clone old_name`

2 Running

Please modify the `run_line_fitting.py` file for running the line-fitting algorithm, especially the lines that load the 1D flux, error, and wavelength arrays. Also, users need to input the redshift of the specific galaxy and specify if the input spectrum is in "vac" (Vacuum) or "air" (Air) wavelength space. Please check the `Python` example below for the definitions of other crucial parameters of the line fitting algorithm.

```
# Step (1): Load the flux, error, and wavelength arrays
data_fits = current_direc + '/example_inputs/j1044+0353_addALL_icubes_wn.fits'
err_fits = current_direc + '/example_inputs/j1044+0353_addALL_icubes_wn_err.fits'
spec = fits.open(data_fits)[0].data
err = fits.open(err_fits)[0].data
wave = np.load(current_direc + '/example_inputs/wave_grid.npy')

# Step (2): Define the redshift of the galaxy
redshift = 0.01287

# Step (3): Specify whether the spectrum is in "vacuum" or "air" wavelength space
vac_or_air = 'vac'

# Step (4): Run the line-fitting algorithm
# The main parameters for the all_lines_result function are defined as follows:

# "n_iteration = 1000": Defines the number of iterations you want to run.
# "get_flux = True": If true, the return is the flux dict. This includes the
#                    flux of each line profile. Otherwise, the return is the
#                    best-fitting parameters.
# "get_error = True": Specifies if you want to calculate the error of
#                     each line flux.
```

```
# "get_corr = True": Specifies if you want the flux to be extinction-corrected.
# "get_ew = True": Specifies if you want to calculate the ew(s) of the selected
#                  emission lines, including emission and absorption ew(s).
# "save_flux_table = True": Defines if you want to save the best-fitting flux
#                           pandas table for each line.
# "save_ew_table = True": Specifies if you want to save the best-fitting
#                         equivalent width pandas table for each line.
# "save_sigma_table = True": Specifies if you want to save the best-fitting
#                            velocity width pandas table for each velocity component.
# "line_selection_method": Can be 'gui' or 'txt'. Defines how you want to select
#                          the lines for fitting, using a GUI or inputting a txt.

input_example_txt = current_direc + '/input_txt/line_selection_example.txt'
region = line_fitting_exec(redshift = redshift, vac_or_air = vac_or_air, E_BV = None,
                           fits_name = data_fits, line_select_method = 'txt',
                           input_txt = input_example_txt)
region.all_lines_result(wave, spec, err, n_iteration = 1000, get_flux = True,
                        get_corr = False, save_flux_table = True, get_ew = True,
                        save_ew_table = True, get_error = True, save_par_table = True)

# Step (5): Plot the fitting result
# "savefig = True" defines if you want to save the fitting result as a .pdf file.
region.fitting_plot(savefig = True)
```

The specific method for choosing the intended lines for fittings, controlled by the `line_selection_method = 'txt' or 'gui'` in the `line_fitting_exec` class, is illustrated in the following subsection.

2.1 Line Selection Methods

Users can select the intended lines for fitting either by using the line-selection GUI (`line_select_method = 'gui'`) or by inputting a text file (`line_select_method = 'txt'`) that includes all the lines for fitting. The description of each method is illustrated below respectively.

2.1.1 Line Selection GUI

When users set the `line_select_method` parameter to `'gui'`, a line-selection window appears, displaying all available emission lines for potential fitting (Figure 1). Users should click on the "white box" adjacent to the desired line's name to opt for a particular line. Beyond the "`Free fitting`" approach, where both the velocity center and width of each component remain variable, users can also select from alternative strategies: "`Fix velocity centroid`" (all components share a common velocity center), "`Fix velocity width`" (uniform velocity width across components), or "`Fix velocity centroid and width`" (all components have matching velocity centers and widths). To finalize their lines and fitting strategy choices, users should click the "`Confirm`" button, moving on to subsequent line selection steps.

The subsequent step (see Figure 3 for details) requires users to characterize the chosen line(s) within the line-selection GUI, primarily by ascertaining if the lines exhibit multiple emission velocity components. The GUI will query users on the necessity of a double-Gaussian (indicating two emission velocity components) or a triple-Gaussian (implying three emission velocity components) model for the line's fitting. Subsequently, users will be prompted to determine if the line showcases broad wings, implying that the velocity width of the second or third emission component should be broader than the first's. Finally, another GUI window will surface, prompting users to inspect if the selected line(s) present absorption troughs, a feature commonly associated with Balmer lines.

2.1.2 Line Selection Input Text File

In scenarios where users wish to fit the same lines and employ a consistent fitting strategy across multiple galaxy spectra, repeating the line selection process through the GUI can become tedious and inconvenient. To streamline this, users can opt to provide a text file containing all the necessary line-selection information by setting the `line_select_method` parameter to `'txt'`. The required format for this text input can be found in `line_selection_example.txt` within the `input_txt` folder (see Figure 2 for details).

2.2 Fitting Parameter GUI

After the line selection procedure, another GUI will appear for users, regardless of the line selection method used. This GUI allows users to input the initial guess for each parameter and specify its range of variation

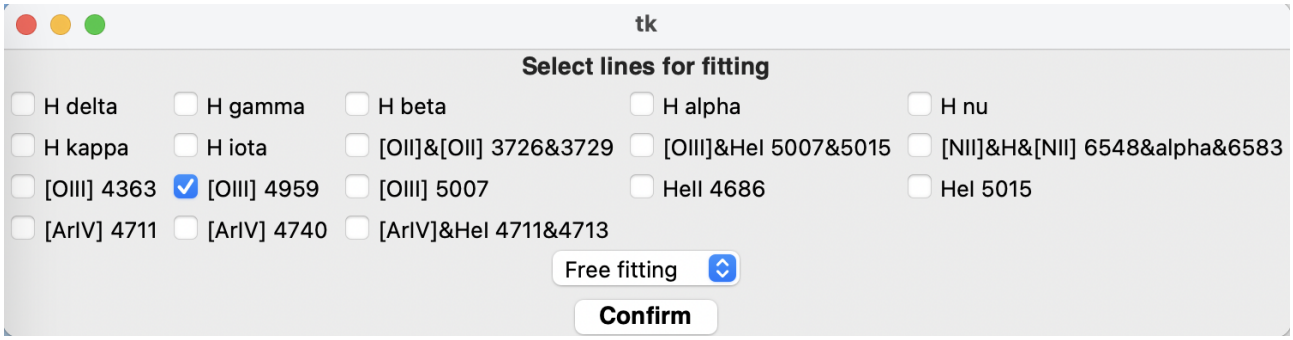

Figure 1: Demonstration of the line-selection GUI allowing users to choose the intended lines for fitting. The [O III] λ4959 line is highlighted for fitting with the `Free fitting` method (alternative options include `Fix velocity centroid`, `Fix velocity width`, and `Fix velocity centroid and width`). To proceed, users are prompted to click the `Confirm` button, advancing to further line selection processes. This includes evaluations for the presence of multiple emission components and absorption troughs (as detailed in Figure 3).

line_selection_example.txt

```
# example input file: selecting lines for fitting

# lines available for fitting (current version)
# 'H delta' 'H gamma' 'H beta' 'H alpha' 'H nu' 'H kappa' 'H iota'
# '[OII]&[OII] 3726&3729' '[OIII]&HeI 5007&5015' '[NII]&H&[NII] 6548&alpha&6583'
# '[OIII] 4363' '[OIII] 4959' '[OIII] 5007'
# 'HeII 4686' 'HeI 5015' '[ArIV]&HeI 4711&4713' '[ArIV] 4711' '[ArIV] 4740'

######################## Line Selection Keywords
# selected lines (needs quotation marks for the selected lines)
selected_lines = '[NII]&H&[NII] 6548&alpha&6583'

# fitting methods
# select from "Free fitting", "Fix velocity centroid", "Fix velocity width", "Fix velocity centroid and width"
fitting_method = "Free fitting"

# lines that have multiple emission components (needs quotation marks for the selected lines)
multi_emis_lines = 'H alpha', '[NII] 6548', '[NII] 6583'

# lines that need double-Gaussian fittings (needs quotation marks for the selected lines)
double_gauss_lines = 'H alpha', '[NII] 6548', '[NII] 6583'

# lines that need triple-Gaussian fittings (needs quotation marks for the selected lines)
triple_gauss_lines = None

# whether the double or triple Gaussian fittings have broad wings or not
double_gauss_broad = True # sigma_2 > sigma_1
triple_gauss_broad = True # sigma_3 > sigma_2 > sigma_1

# lines that have an absorption component (needs quotation marks for the selected lines)
absorption_lines = None
```

Figure 2: Example of the line selection input text file. Key parameters that will be used in subsequent analyses include `selected_lines` (lines selected for fitting), `fitting_method` (fitting strategy, as described in Section 2.1.1), `multi_emis_lines` (lines with multiple emission components), `double_gauss_lines` (lines requiring double-Gaussian fitting), `triple_gauss_lines` (lines requiring triple-Gaussian fitting), `double_gauss_broad` (indicating if the second emission component has a broader velocity width than the first), `triple_gauss_broad` (indicating if the second and third emission components have broader velocity widths than the first), and `absorption_lines` (lines with an absorption trough or component). Lines beginning with the symbol `'#'` are treated as comments. For the parameters `selected_lines`, `multi_emis_lines`, `double_gauss_lines`, `triple_gauss_lines`, and `absorption_lines`, the selected lines should be separated by commas, and each selected line needs quotation marks.

for each iteration. By modifying the default initial values, users can ensure more appropriate initial parameter guesses. Furthermore, if users want to set the amplitude ratios between some pair of lines, such as [O, III] λ4959, 5007 or [N, II] λ6548, 6583, to ensure more physical line fitting results, they need to click on the `Set Amplitude Ratio` button to select the specific pairs of lines and their respective line ratio values in a GUI (Figure 5).

Since the initial guesses are adjusted for each iteration based on the specified range of variation for each parameter, this method reduces the likelihood that the best-fitting result gets trapped in a local minimum within the multi-dimensional χ^2 space. For details, please check Figure 4 for the fitting-parameter GUI.

3 Saved Results

Users can review various line-fitting outputs. These outputs include a plot showcasing the best-fitting line profile, a CSV table detailing the best-fitting flux for each selected line, another CSV table displaying the best-fitting equivalent width (EW) of each selected line (and individual velocity components), and a final CSV table presenting the best-fitting parameter values for each velocity component. Users should enable the respective saving options to save each of these features. A summary of these options is provided below:

1. If `savefig = True` in the `fitting_plot()` function, the fitting result is saved as a .pdf file in the `plots` subfolder.
2. If `save_flux_table = True` in the `all_lines_result()` function, the best-fitting line flux table is saved in the `flux_tables` subfolder.
3. If `save_ew_table = True` in the `all_lines_result()` function, the best-fitting equivalent width table is saved in the `ew_tables` subfolder.
4. If `save_par_table = True` in the `all_lines_result()` function, the best-fitting parameter table is saved in the `parameter_tables` subfolder.

To show the versatility of this fitting package, the line-fitting examples in different scenarios are shown as follows.

1. Single-Gaussian fitting for a single line (see Figure 6).
2. Simultaneous single-Gaussian fittings for multiple lines (see Figure 7).
3. Triple-Gaussian fitting for a single line (see Figure 8).
4. Gaussian-Lorentzian fittings for the Balmer lines Hβ, Hγ, and Hδ that have obvious Balmer absorption troughs (see Figure 8).
5. Simultaneous fitting that includes a triple-Gaussian fitting for a strong line and two single-Gaussian fittings for two blended weaker lines (see Figure 10).

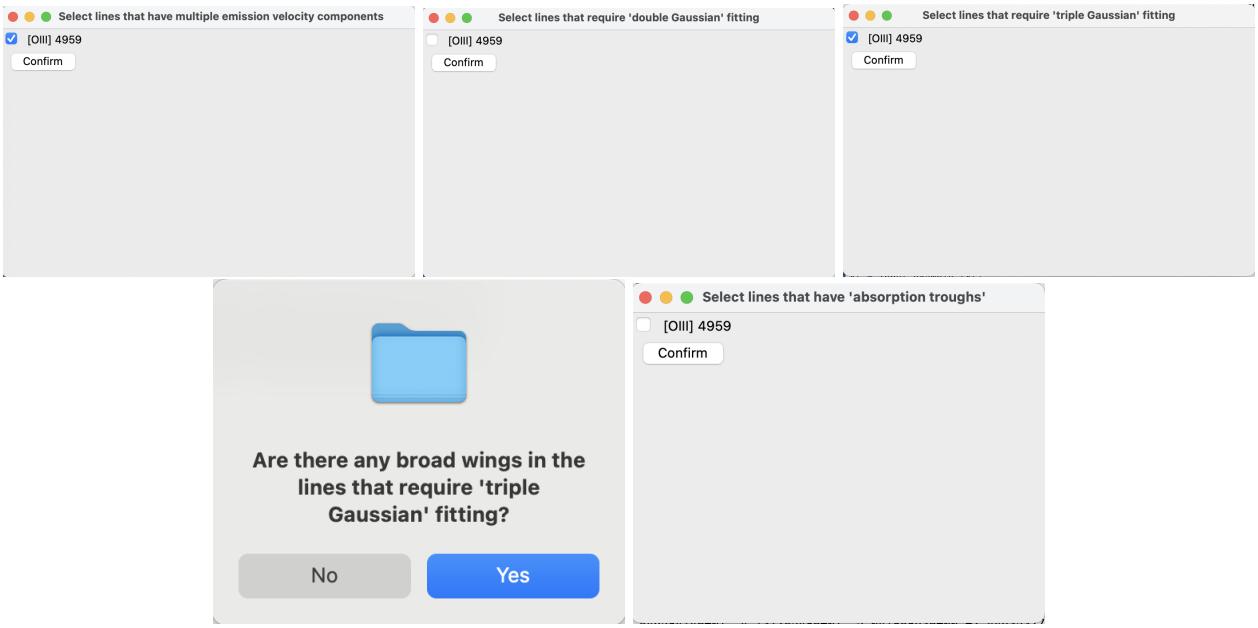

Figure 3: An illustration of the line-selection GUI, aiding users in determining the characteristics of the selected line(s) — specifically, [O III] λ4959 in this instance. *Top Left*: Assessing the presence of multiple emission velocity components. *Top Middle*: Establishing the need for a double-Gaussian fitting for its emission components. *Top Right*: Identifying the requirement for a triple-Gaussian fitting for emission components. *Bottom Left*: Deciding if the line exhibits broad wings, causing the velocity width of the second or third emission component to be constrained as greater than the first. *Bottom Right*: Checking for a pronounced absorption trough demanding a fitting via a Lorentzian profile, typically seen in Balmer lines.

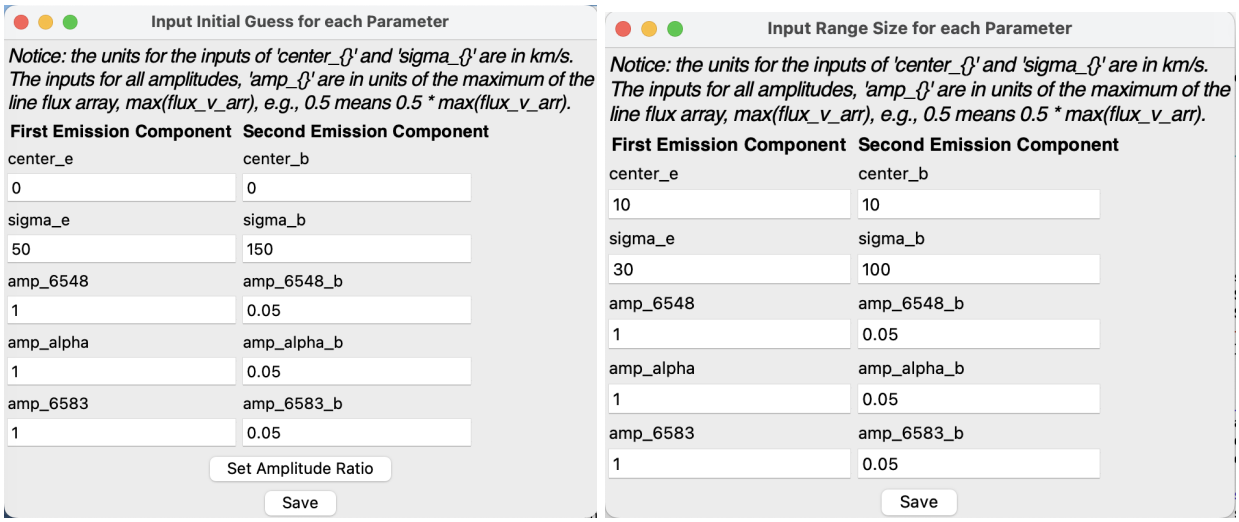

Figure 4: This is an example of the fitting-parameters GUI that allows users to input both the initial guess and the range of variation for each fitting parameter for the selected line(s). *Left*: the initial-guess window displays the default initial value for each fitting parameter. Each column header represents a parameter group (e.g., `First Emission Component`). A parameter name beginning with `center` denotes the velocity centroid of the Gaussian function, which defaults to `0` in units of km s^{-1}. Similarly, a name starting with `sigma` indicates the velocity width of the Gaussian function; its default value varies depending on the component and is also given in units of km s^{-1}. Parameters prefixed with `amp` define the amplitude of the Gaussian function. The default value for this varies among components and is given in units of the maximum flux, `max(flux_v_arr)`, from the extracted region surrounding the selected line. For instance, a value of 0.5 means `0.5 * max(flux_v_arr)`. Users can click the `Set Amplitude Ratio` button to fix the amplitude ratios between several selected line pairs (see Figure 5). *Right*: the parameter range-size window. Its columns and parameter nomenclature mirror those of the initial-guess window. The provided range values dictate the variation range for the initial guesses of fitting parameters throughout each fitting iteration.

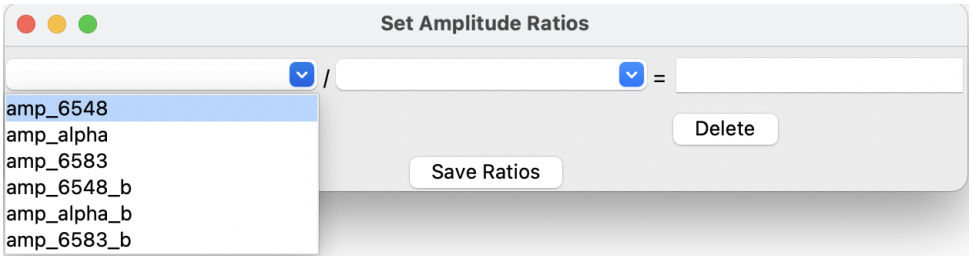

Figure 5: An example of the GUI that enables users to constrain the amplitude ratios between emission line pairs. Users need to select two lines, e.g., `amp_6583` and `amp_6548`, and their respective ratio, e.g., the intrinsic ratio is ~ 3 for this [N II] doublet. To add or delete one selected line par, users can click the `Add` or `Delete` button on the GUI's lower left or lower right corner. Last, users must click the `Save Ratios` button to set the amplitude ratios for the line-fitting process.

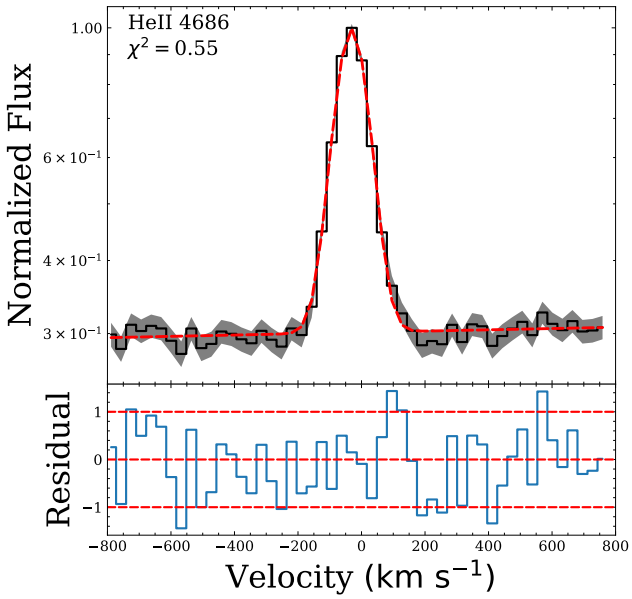

Figure 6: Example of single-Gaussian fitting for He II λ4686 line.

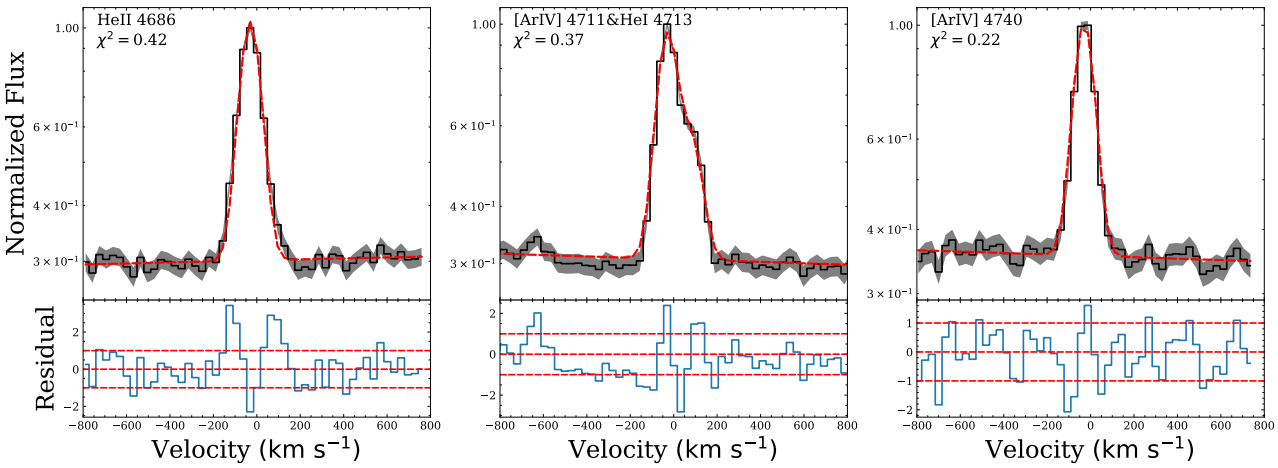

Figure 7: Example of simultaneous single-Gaussian fittings for He II λ4686, [Ar IV] λ4711, He I λ4713 line, and [Ar IV] λ4740. The velocity width and centroid are tied together in the fittings.

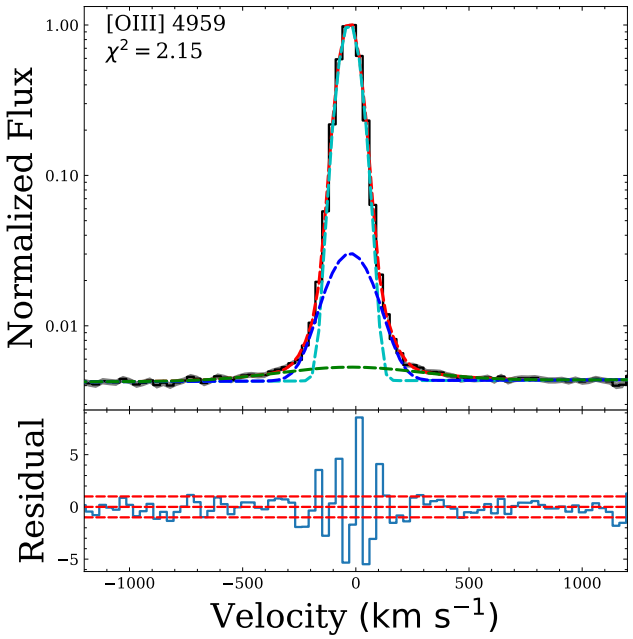

Figure 8: Example of triple-Gaussian fitting for [O III] λ4959 line such that the second and the third emission components have larger velocity widths than the first.

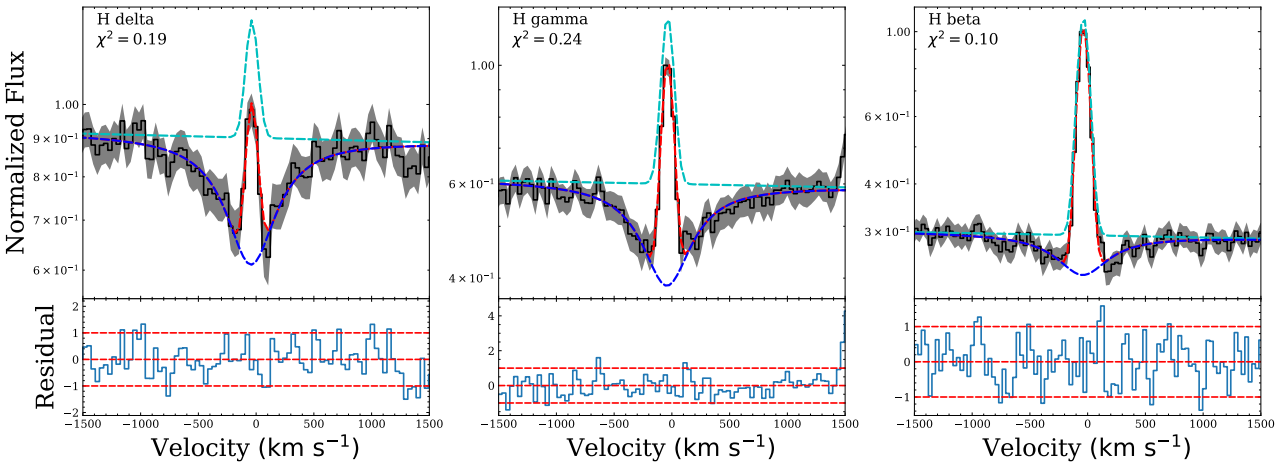

Figure 9: Example of Gaussian-Lorentzian fitting for Balmer lines that have obvious absorption troughs.

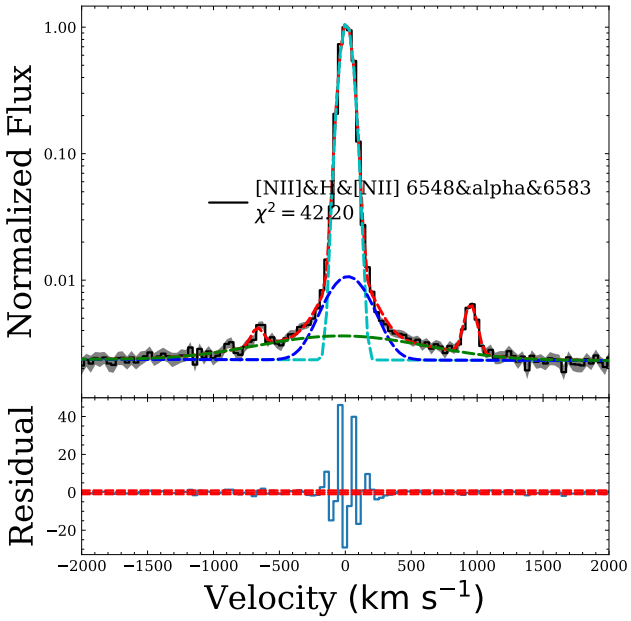

Figure 10: Example of a simultaneous fitting that includes a triple-Gaussian fitting for the strong line Hα and two single-Gaussian fittings for the weaker blended doublet [N II] $\lambda 6548, 6583$.