



UNIVERSIDAD  
**COMPLUTENSE**  
MADRID



**THE UNIVERSITY**  
*of* **NORTH CAROLINA**  
*at* **CHAPEL HILL**

Faculty of Science  
Department of Physics and Astronomy

**A Binning Code for 2D Spectroscopy in the**  
***RESOLVE* Survey**

Trabajo Fin de Grado  
Grado Ciencias Físicas

Presented by:  
**Pablo Lechón Alonso**

Advisor:  
Dr. Sheila Kannappan

Curso 2018/2019



UNC  
COLLEGE OF  
ARTS & SCIENCES

THE UNIVERSITY  
of NORTH CAROLINA  
at CHAPEL HILL

DEPARTMENT OF PHYSICS  
AND ASTRONOMY  
272 PHILLIPS HALL  
CAMPUS BOX 3255  
CHAPEL HILL, NC 27599-3255

T 919.962.2078  
F 919.962.0480

[www.physics.unc.edu](http://www.physics.unc.edu)

May 31, 2019

Tribunal del Trabajo de Fin de Grado  
Departamento de Física de la Tierra y Astrofísica  
Universidad Complutense de Madrid  
Plaza Ciencias, 1, 28040 Madrid

Dear Evaluation Committee,

I was Pablo's Advisor during his exchange year at the University of North Carolina at Chapel Hill (UNC). During both the fall and spring semesters he worked in my research group through two courses: Research with Faculty Mentor I (PHYS 295, 3 credit hours) and Research with Faculty Mentor II (PHYS 395, 4 credit hours). Note that 1 credit hour is equivalent to 2 ECTS credits. His primary work for these courses was to develop the research project summarized in the report to which this letter is attached. Along the way, he kept a daily research log, met with me at least once per week, and regularly submitted a written progress report due every other week. He also read background literature and joined in research group meetings and discussions. In both semesters he presented excellent talks to our research group on his work, and near the end of the second semester he presented a formal poster on his work at UNC's "Celebration of Undergraduate Research" Conference. His final written report each semester summarized his results to that point, and he also wrote a tutorial at the end of the second semester to enable others in our research group to use the code he developed.

The work he is presenting to you is his final report on his project, which was turned in on May 3<sup>rd</sup>, 2019. As stated in the attached UNC transcript, Pablo's final grade for PHYS 395, which constitutes the culmination of his work with me at UNC, was an A, the highest grade UNC awards.

Sincerely,

Sheila Kannappan, PhD  
Associate Professor  
Department of Physics and Astronomy  
The University of North Carolina at Chapel Hill

# THE UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL



*[Signature]*  
University Registrar

RAISED SEAL NOT REQUIRED

**Name:** Lechon, Pablo  
**Student ID:** 730291691

**Birthdate:** 09/17/1996  
**Print Date:** 06/04/2019

## Academic Program History

**Program:** AS Non Degree UGRD Exchange  
**Active in Program:** 04/02/2018:  
College of Arts and Sciences  
Non-Degree Undergraduate Major  
Undergraduate Visitor Non Degree  
**04/02/2018:**

## Beginning of Undergraduate Record

### 2018 Fall

Course	Description	Attempted	Earned	Grade	Points
PHYS 295	RESEARCH WITH FACULTY I	3.000	3.000	A-	11.100
PHYS 405	BIOLOGICAL PHYSICS	3.000	3.000	A	12.000
PHYS 521	APPL OF QUINTM MECHS	3.000	0.000	W	0.000
PHYS 631	MATH MET THEO PHY	3.000	3.000	A	12.000
PHYS 660	FLUID DYNAMICS	3.000	3.000	A-	11.100

<b>Term GPA</b>	<b>3.850</b>	<b>Term Totals</b>	<b>15.000</b>	<b>12.000</b>	<b>12.000</b>	<b>46.200</b>
-----------------	--------------	--------------------	---------------	---------------	---------------	---------------

<b>Cumulative GPA</b>	<b>3.850</b>	<b>Cum Totals</b>	<b>15.000</b>	<b>12.000</b>	<b>12.000</b>	<b>46.200</b>
-----------------------	--------------	-------------------	---------------	---------------	---------------	---------------

**Term Honor:** Dean's List

### 2019 Spring

Course	Description	Attempted	Earned	Grade	Points
ASTR 519	OBSERVA ASTR	4.000	4.000	B	12.000
MUSC 112	GROUP LESSONS IN STRINGS	1.000	1.000	A-	3.700
Course Topic:	Beginning Classical Guitar				
PHYA 341	INTERMEDIATE VOLLEYBALL	1.000	0.000	A	4.000
Grading Basis:	EXCLUDE FROM HRS TO GRADUATION				
PHYS 395	RESEARCH WITH FACULTY II	4.000	4.000	A	16.000
PHYS 712	ELECTROMAG THEORY	3.000	3.000	B+	9.900

<b>Term GPA</b>	<b>3.508</b>	<b>Term Totals</b>	<b>13.000</b>	<b>12.000</b>	<b>13.000</b>	<b>45.600</b>
-----------------	--------------	--------------------	---------------	---------------	---------------	---------------

<b>Cumulative GPA</b>	<b>3.672</b>	<b>Cum Totals</b>	<b>28.000</b>	<b>24.000</b>	<b>25.000</b>	<b>91.800</b>
-----------------------	--------------	-------------------	---------------	---------------	---------------	---------------

End of Official Undergraduate Academic Record

# A Binning Code for 2D Spectroscopy in the *RESOLVE* Survey\*

Pablo Lechón

## Abstract

We have written a binning code <sup>1</sup> to improve the signal-to-noise ratio (S/N) of spectroscopic data of galaxies drawn from the *RESOLVE* survey by determining and performing the optimal binning in a specific way based on the user's input. The input data we used are fluxes and errors of several emission lines corresponding to different positions in the galaxy. The improved S/N data will initially be used to determine optimized Balmer decrements to be used for dust extinction correction (in which case the user will bin the whole galaxy possibly rejecting a central row contaminated by light from an active galactic nucleus (AGN)). Future uses will include obtaining optimized metallicities and metallicity gradients (binning in sets of rows that assure a certain S/N target) and getting matched spectroscopic parameters to combine with photometric profiles for analysis with stellar population synthesis code (binning only rows specified by the user). We have tested the code for the case of Balmer decrement measurements, finding a significant improvement in the number of galaxies with high enough S/N (by nearly a factor of 3 for S/N = 5). The measured Balmer decrements show discrepancies with SDSS values that could be fixed by modeling absorption lines and using a more accurate flux calibration.

---

\*This project has been typeset from a L<sup>A</sup>T<sub>E</sub>X file prepared by the author.

<sup>1</sup>This code has been entirely written using Python

# Contents

<b>1</b>	<b>Motivation</b>	<b>5</b>
1.1	Extinction Correction . . . . .	5
1.2	Metallicity and metallicity gradients . . . . .	5
1.3	Photometric profiles . . . . .	6
1.4	AGN detection . . . . .	6
<b>2</b>	<b>Details of the code</b>	<b>6</b>
2.1	Main module . . . . .	7
2.1.1	Input . . . . .	7
2.1.2	Best binning . . . . .	8
2.1.3	Output . . . . .	10
2.2	Secondary Module . . . . .	10
2.2.1	Alignment . . . . .	10
2.2.2	Perform binning . . . . .	13
2.2.3	Re-fitting . . . . .	13
<b>3</b>	<b>Impact of the code</b>	<b>13</b>
<b>4</b>	<b>Results</b>	<b>14</b>
4.1	Quality of measurements . . . . .	15
4.2	Improvement of measurements . . . . .	17
<b>5</b>	<b>Conclusions</b>	<b>17</b>
<b>6</b>	<b>Future Work</b>	<b>17</b>

# 1 Motivation

Improving the S/N in 2D spectroscopic data is key to obtaining useful measurements from galaxies that would not yield any results otherwise. By binning the data in a specific way depending on the science case, we can significantly improve (by a factor of 2 or even 3 in some cases) the number of galaxies with a satisfactory S/N. This code has several applications, namely, extinction correction, measurement of metallicity and metallicity gradients, determination of stellar population parameters, and more accurate AGN classification.

## 1.1 Extinction Correction

One of the multiple ways of quantifying the level of extinction of an HII region is by measuring the Balmer Decrement. This quantity is defined as the ratio between the intensities of the lines  $H_\alpha$  and  $H_\beta$ .

$$\frac{H_\alpha}{H_\beta} = 2.86 \quad (1)$$

These are allowed lines, which are characterized by the fast decay of excited electrons. An electron can deexcite in two ways; collisionally or radiatively. Due to the fast decay, the probability for the former to happen is very low, and only the latter takes place. The consequence of this is that the ratio of equation 1 is constant within the ranges of temperatures of the measured regions. Therefore, any divergence from that value is due to extinction that needs to be corrected before making any other calculations. Note that the Balmer decrement depends significantly on the temperature at very hot regimes [1] but ours is not the case. Using the binning code to improve the S/N of  $H_\beta$  ( $H_\alpha$  is always more intense, so it does not need to be binned) will increase the number of galaxies for which a Balmer decrement is available, allowing for a more complete analysis of *RESOLVE* data.

## 1.2 Metallicity and metallicity gradients

Measuring metallicities is the way we obtain information on the chemical composition of galaxies. A way to determine a metallicity is by computing the ratio

$$\frac{[N_{II}]}{H\alpha} \quad (2)$$

Both the quality of this ratio and the number of galaxies for which we can calculate it will be increased by improving the S/N of  $[N_{II}]$ .

Another type of binning that bins rows to a certain S/N could be used to calculate metallicity gradients. These are useful because they provide us with constraints on gas inflows and outflows, which could lead to conclusions about star formation history and overall galaxy evolution.

### 1.3 Photometric profiles

The binning code will also be used to bin data to match photometric profiles (see Figure 2).

Feeding spectroscopic data into an existing stellar population synthesis code, along with the photometric data it uses, will provide much better constrained parameters (stellar mass, Stellar formation rate (SFR), stellar metallicity...)

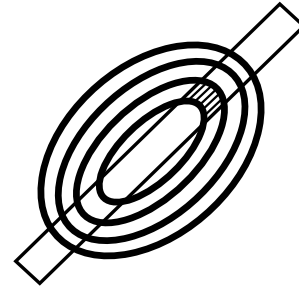


Figure 2: Matching the spectroscopic slit (rectangle) with the photometric annuli (ovals).

### 1.4 AGN detection

The determination of AGN candidates requires the calculation of line ratios such as  $[N_{II}]/H\alpha$  and  $[O_{II}]/H\beta$ . The number of galaxies from which we can obtain this ratios is expected to increase thanks to the binning code. Conversely, once AGN are identified we can reject their galaxies central row of data to improve the other measurements detailed above

## 2 Details of the code

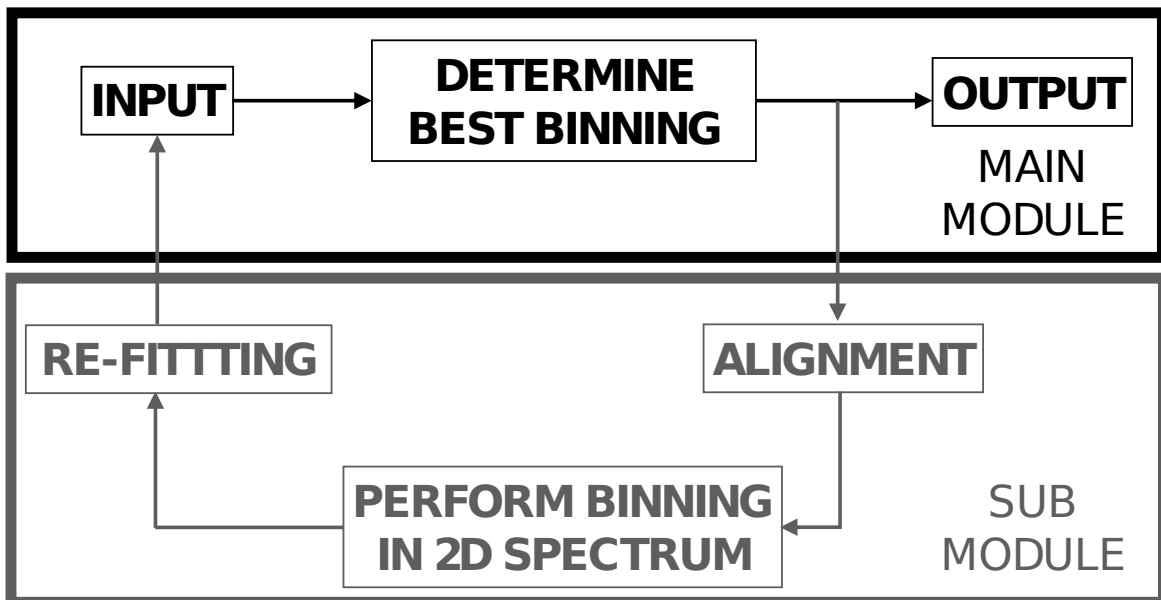


Figure 3: General flow of the code.

The big picture of the flow of the code can be seen in Figure 3. The main module of the code determines the best binning using data provided in the input. The second module is optional. If the user provides a 2D spectrum (i.e., one spectrum at each position of the galaxy), the code will be able to perform the previously determined best binning on the spectrum (after aligning it) according to the user specifications. The output will now include an image with the binned 2D spectrum. A re-evaluation of the fits based on the binned spectrum and consecutive feedback to the main module would improve the determination of the optimal binning.

In the following sections 2.1 and 2.2 we break up Figure 3 into parts (each corresponding to a box) and analyze them in detail.

## 2.1 Main module

In this section, we will analyze the form of the input, the different binning options (and their possible applications), and the output.

### 2.1.1 Input

First, there is a table with fluxes and errors from each emission line (columns) at each spatial position on the galaxy (rows). This table is obtained by fitting the lines in the spectrum with Gaussians. Computing the area under the curve for each line yields the flux value. The error is calculated as  $\sqrt{N_{chan}} \cdot \sigma_{chan}$ , where  $N_{chan}$  is the number of channels spanned by the line width, and  $\sigma_{chan}$  is the noise with no line. Along with this table the user needs to provide an array of strings with the names of the lines appearing in it. This is necessary so that the name of the line for which the optimal binning is going to be determined can be matched to a column from the table of fluxes and errors (more precisely, two columns; flux and error for the desired line). An array of velocities at each position of the galaxy is used to perform a reference frame correction on the fluxes. It is also crucial to de-Doppler-shift the spectrum if it is provided. The center of the galaxy needs to be specified too because it will be removed if the galaxy's AGN flag is set positive. The reason for this is that the Balmer decrement in AGN regions is not fixed, so it is useless if one wants to perform extinction correction. Finally, the user must specify the desired binning option they want to use.

Optional arguments needed for the second module can also be provided. The two most important are the 2D spectrum and a wavelength solution.

In Figure 4, the spectrum has not been Doppler-shift corrected yet. Each row is red-shifted due to the cosmological expansion, but it is also Doppler-shifted due to the rotation of the galaxy. The same line in different spatial positions of the galaxy will be shifted differently, giving us the rotation curve.

Some sort of wavelength solution also needs to be provided in order to perform the alignment.

There are other optional/default parameters such as the AGN flags, the paths where the input data can be found and the path where the output will be saved...



Table 1 summarizes the input arguments. The code takes 6 required arguments (bold face), and up to 11 optional ones (normal font), making a total of 17 possible arguments.

1. <b>Table of Fluxes &amp; errors</b>	4. <b>Velocities</b>
2. <b>Name of lines</b>	5. <b>Galaxy Center</b>
3. <b>Line to be binned</b>	6. <b>Binning option</b>
7. Input path	12. Wavelength solution
8. Output name	13, 14. Pre-binning row ranges
9. Output path	15. Tolerance
10. AGN flag	16. Print warnings
11. 2D Spectrum	17. Aligned spec

Table 1: Input items. Bold faced ones are required, whereas the others are optional

### 2.1.2 Best binning

Depending on the science purpose, there are several binning options. Based on the user’s choice, the code computes the best binning.

The first option involves **binning everything**. This option bins all the rows that improve the total S/N. It leaves out of the binning those rows that, when added, produce a worse S/N ratio because they are noise for the most part. This option would be useful to obtain an optimized Balmer decrement (see section 4 for a deeper analysis on this), and optimized metallicities.

Note that the S/N is always computed as

$$S/N = \frac{\sum_i F}{\sqrt{\sum_i \Delta F^2}} \quad (3)$$

where F is the fluxes obtained by fitting the spectrum, and  $\Delta F$  are the errors.

The second option bins only the **sets of rows** specified by the user. This option would be of particular interest if the user is trying to match a specific region of the spectrograph slit with its corresponding photometric profile (see Figure 2). If both kinds of data are available one could run then through a stellar population synthesis code and obtain better constrained parameters than by using only photometry.

The third option bins into new rows with a **S/N target** specified by the user. The optimal binning is determined to satisfy the target within a certain tolerance (20 % by

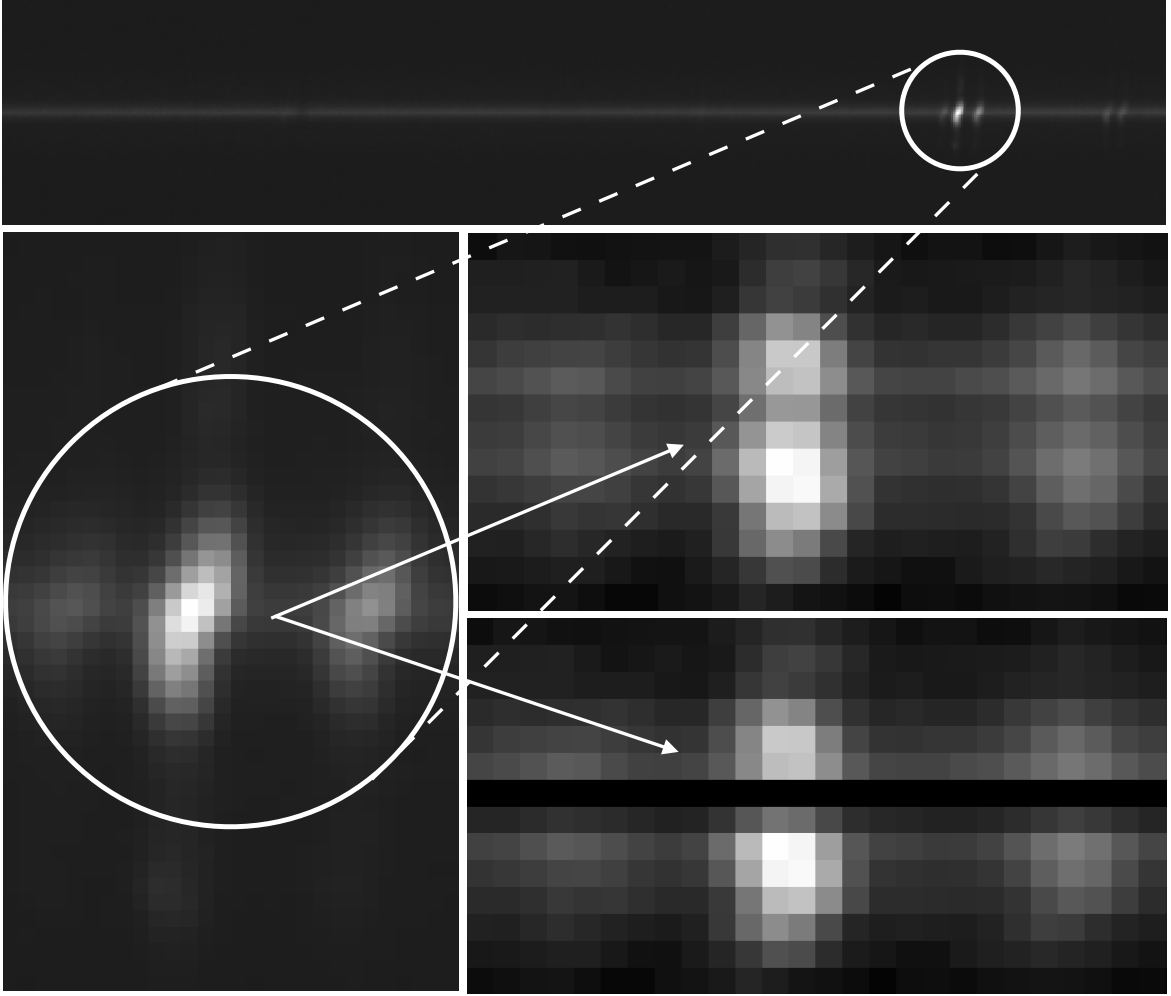


Figure 4: This figure is a graphical example of the core behavior of the code. A binning for the line  $H_\alpha$  (bottom left image) has been determined through the main module and applied through the secondary module to the 2D spectrum (upper image) of a galaxy, which has a spatial (vertical) dimension, and a wavelength (horizontal) dimension. On the center-right image the optimal binning has been performed. Below we can see the same result when an AGN has been flagged at the center position of the galaxy.

default) in each of the binned rows. The code starts binning the center symmetrically. By that, we mean that every time a row above the center row is added, another row from below the center is added too, even if it was not necessary because the S/N target had already been reached in the center row. By doing this, we are forcing the center of the galaxy to remain symmetric. After the center has been binned up, the rest of the rows above and below are binned separately in groups that also reach the required S/N. This option will be useful if the user is trying to calculate metallicity gradients. When the ends of the galaxy are reached, the lack of adjacent rows to bin may mean that the first and/or last position of the binned rows don't reach the target S/N within the given tolerance. In this case, they are thrown away instead of being added to the previous rows (even if they improve the S/N) If we added it that group of binned rows would be very delocalized which is not convenient when calculating metallicity gradients.

Prior to computing the best binning according to one of the above options, the center of the galaxy is removed if the AGN flag is set to a positive value. Line ratios in AGN regions are extremely different than in HII regions because the Balmer decrement depends on temperature significantly. Therefore we need to remove AGN regions when determining the optimal binning for most application other than analysis of the AGN themselves.

### 2.1.3 Output

The output of the code contains several items that are saved to a folder named `results_binning`, namely, a summary table and one to three files.

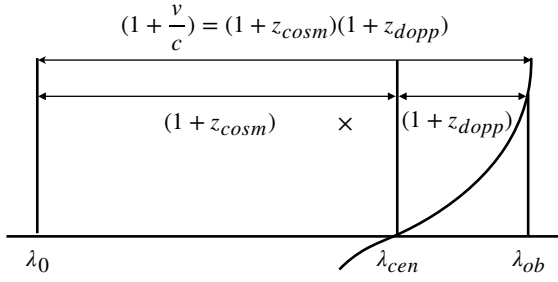
The summary table contains the S/N reached after the binning, the binned flux of the emission line for which the binning is being determined, and the set of rows that have been binned. Note that if one wanted to get the binned errors a simple division of the fluxes by the signal to noise would be sufficient. If the user provides a spectrum, a FITS image with the binned spectrum is also saved to the `results_binning` folder. There is also an optional argument that allows the user to save a FITS image of the aligned (but not binned) spectrum.

## 2.2 Secondary Module

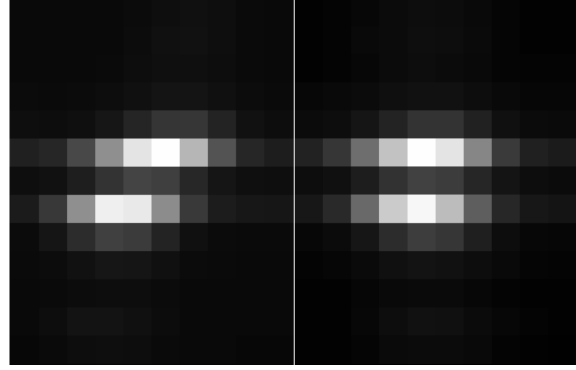
If the user inputs a 2D spectrum, then the code enters the secondary module. The optimal binning determined in the main module is performed on the provided spectrum. To do this successfully we need to align it first (see Figure 4). After the binning is performed we study the possibility of a re-evaluation of the fits.

### 2.2.1 Alignment

The positions of the lines in the spectrum of the galaxy is shifted differently depending on what part of the galaxy we are looking at (see bottom left image in Figure 4).



(a) Visualization of the Doppler correction. The curved line is the rotation curve of the galaxy.  $\lambda_0$  is the rest wavelength of the line.  $\lambda_{cen}$  is the wavelength of the line in the center of the galaxy. Since the center of the galaxy does not rotate, it has no Doppler shift, only cosmological redshift.  $\lambda_{ob}$  is the observed wavelength of the line in a position away from the center. The two correction factors are shown. To account only for the Doppler shift one needs to divide each  $\lambda_{ob}$  only by the Doppler shift correction.



(b) Comparison before and after the alignment. The positions right above and below the center are more intense than the center itself. This is due to the arbitrary pre-binning that was made in order to obtain fluxes and velocities. In section 6 we discuss the problems that this arbitrary pre-binning raises, and how to tackle them.

Figure 5

This is due to the rotation of the galaxy. On top of this, the whole galaxy is redshifted due to cosmological expansion. The code corrects only for the rotation curve, not for the redshift. The reason for this is that we want to modify our data as little as possible. This is a general rule that experience has demonstrated to be helpful when analyzing data.

The first thing that the code does once it enters into module 2 is straighten the lines so that they can be properly added together afterwards. This alignment process has two main steps; de-Doppler-shift (see Figure 5a), and resampling (see Figure 6).

To correct for the Doppler shift, we divide the 2D spectrum of by a spatial position dependent factor in order to bring every row to the velocity frame of reference of the center. This frame of reference is only affected by the redshift since the center of the galaxy is not rotating. To obtain the appropriate multiplicative factor for each row, we consider that the shift of the observed wavelength is due to both cosmological expansion and Doppler shift.

$$\begin{aligned}\lambda_{ob} &= \left(1 + \frac{v}{c}\right) \lambda_0 \\ &= (1 + z_{dopp})(1 + z_{cosm})\lambda_0\end{aligned}$$

However, since we are only correcting for the Doppler shift, we are interested in  $\lambda_{cen}$ ,

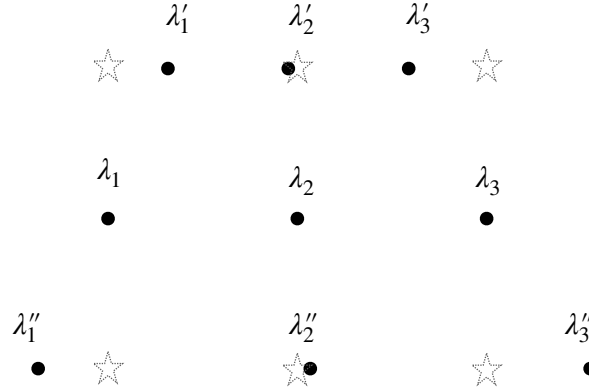


Figure 6: The rows above and below center (points) will be stretched or compressed after the correction. Resampling the data (stars) obtains the values for the fluxes at the wavelength positions of the center row, so we can sum rows properly.

not in  $\lambda_0$  (see figure 5a for clarity on this nomenclature).

$$\lambda_{ob} = (1 + z_{dopp})\lambda_{cen} = \frac{1 + \frac{v}{c}}{1 + z_{cosm}}\lambda_{cen} \quad (4)$$

As seen in equation 4 the correction factor is the ratio of the shift due to the total velocity (obtained from the fits to the spectrum) and the redshift (velocity of the center row).

There are some caveats to this. First, when we scale the wavelengths in one row by the correcting factor, we are not only aligning it to the central wavelength, but we are also rescaling the fluxes. The spectrum stretches or compresses depending on the correction factor being greater or less than 1. We have to compensate for that by multiplying the fluxes of each row by the inverse of the correction factor. This correction is tiny, but it is important in terms of working in a conceptually well-built frame and it may enable more general uses of the code, e.g. stacking spectra of galaxies with different cosmological redshifts

The other caveat on the process of straightening is resampling. Due to the rescaling of the rows in the 2D spectrum, these need to be resampled onto the wavelength grid of the center row with a flux conserving interpolation algorithm [2]. This is the same as saying that the integral over the resampled curve (total flux in the emission line) should yield the same value as before the re-sampling.

More precisely, when we de-redshift the rows, the photons lose energy. This implies that the energy flux is not conserved. However, the photon flux is because the number of photons is conserved.

### 2.2.2 Perform binning

After aligning the 2D spectrum we are in good shape to bin it according to the output of the first module. One of the columns of the summary table the main module outputs contains the sets of rows in the physical image of the 2D spectrum. This column will be used to bin it.

Note that there is a difference between determining the best binning and performing the binning. To determine the best binning, one only needs the initial S/N ratios. These can be obtained from the 2D table of fluxes and errors for each line that is provided as part of the input (first item in Table 1). There is no need to modify the 2D spectrum of the galaxy. However, one performs the binning on the spectrum (if provided) according to the optimal binning obtained from the fitted fluxes and errors.

### 2.2.3 Re-fitting

After the 2D spectrum has been binned, a reevaluation of the fits can be carried out. The new binned 2D spectrum would be fed to the code that fits it in order to generate a new table with fluxes, errors, and velocities. With the new data, an improved optimal binning and consequent optimal spectrum can be calculated. This step would close the loop and make the code iterative. However, this option has not been implemented yet.

## 3 Impact of the code

This code will interact with the work of several members of the group, providing them with results and acquiring data from their work to improve its performance.

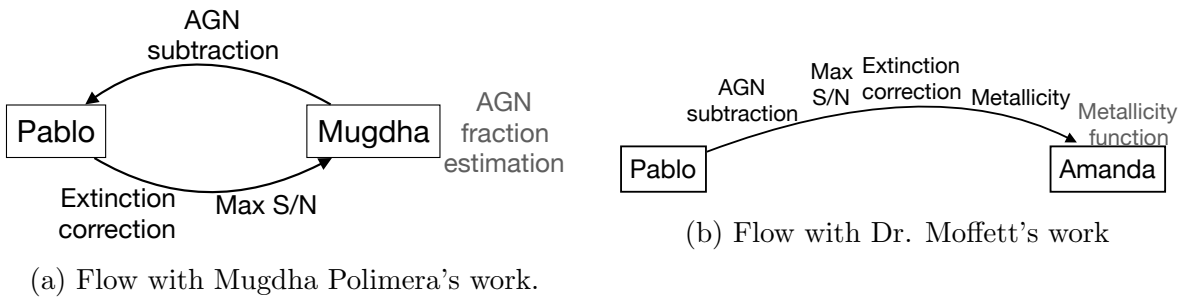


Figure 7

One of the inputs of the code is the AGN flags. Mugdha Polimera et. al are currently working on an improved method for AGN classification. In order to estimate the fraction of galaxies that may have AGN in the local Universe, Polimera et. al (in prep) use optical emission line diagnostic diagrams [3–5], that use multiple optical line fluxes.

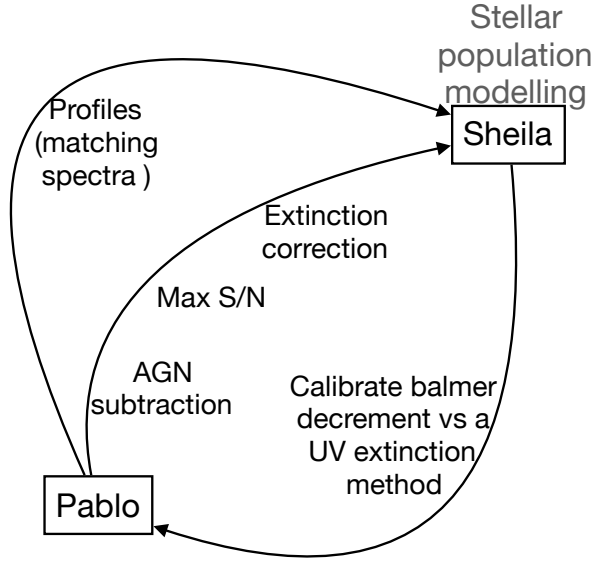


Figure 8: Flow with Dr Kannappan’s work

The improved S/N offered by the binning code will help measure the relevant emission line fluxes. Prior to this, the code will provide her with a higher number of galaxies with available Balmer decrements with which it is possible to perform extinction correction, whereupon she can use the corrected and improved S/N lines in her calculations. As a result, the binning code will receive more accurate AGN flags as inputs in the future. This will, in turn, generate more accurate results. Note that this interaction can be recursive.

There will also be an interaction with the work of **Dr. Amanda Moffett** work. In this case, the binning code will provide her with AGN subtraction, Balmer decrements, and finally, the optimally binned lines used to compute metallicities for each galaxy, so that she can apply this to her work on a metallicity function. Finally, the presented code will also interact with the work of **Dr. Sheila Kannappan** in stellar population synthesis analysis. This interaction has two levels. First, the code will provide Dr. Kannappan with AGN subtraction and Balmer decrements. One of the outputs of Dr. Kannappan’s code is the calibration of the Balmer decrement to a UV extinction method. We can use this method to extinction correct galaxies for which there is no Balmer decrement available. This process can also be recursive.

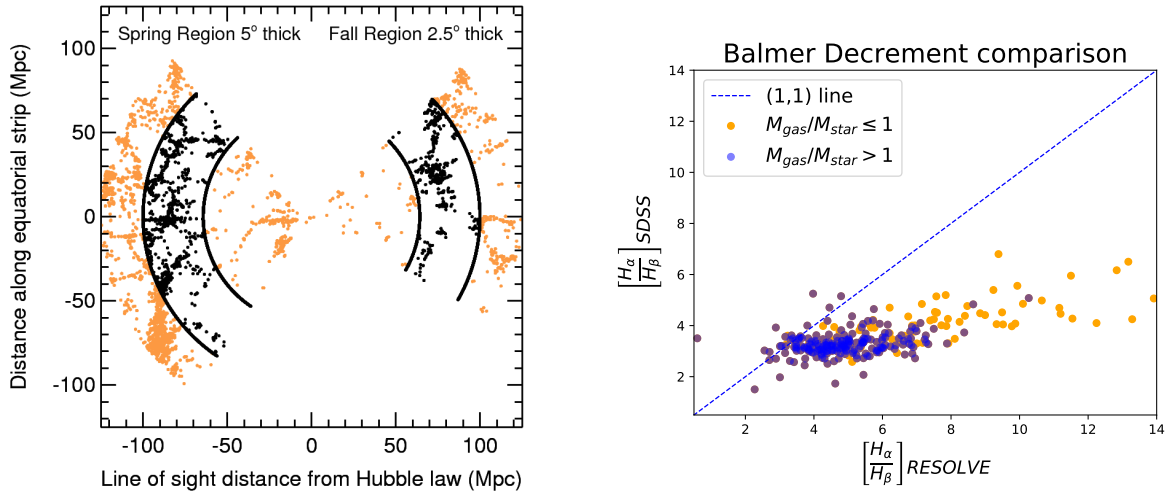
On a second level, the code will provide Dr. Kannappan’s code with spectroscopic data matching the photometric data she already has, as shown in Figure 2

## 4 Results

The binning code has been applied to a subset of 535 galaxies from the *RESOLVE* survey [6]. This is a volume-limited survey containing more than 1600 galaxies.

We have binned the  $H_\beta$  line for all the galaxies using option 1 (bin everything). We

have applied the optimal binning for  $H_\beta$  to the  $H_\alpha$  line as well, allowing us to compute Balmer decrements. We then have compared *RESOLVE* measurements with SDSS measurements [?] because the latter are already final and fully corrected [4, 7, 8]. From this comparison we have arrived at several conclusions regarding the quality of our measurements and the improvements after the binning.



(a) The volume limited *RESOLVE* Survey (black points) is shown within a portion of the larger SDSS Redshift Survey from which it was drawn (orange points). *RESOLVE* comprises 1600 galaxies in two equatorial strips. (Figure from <http://resolve.astro.unc.edu>.)

(b) *RESOLVE* Balmer decrement plotted against SDSS Balmer decrements. Galaxies dominated by gas (blue points) are in the region that is closer to the expected (1,1) line. The orange points are galaxies dominated by stars.

Figure 9

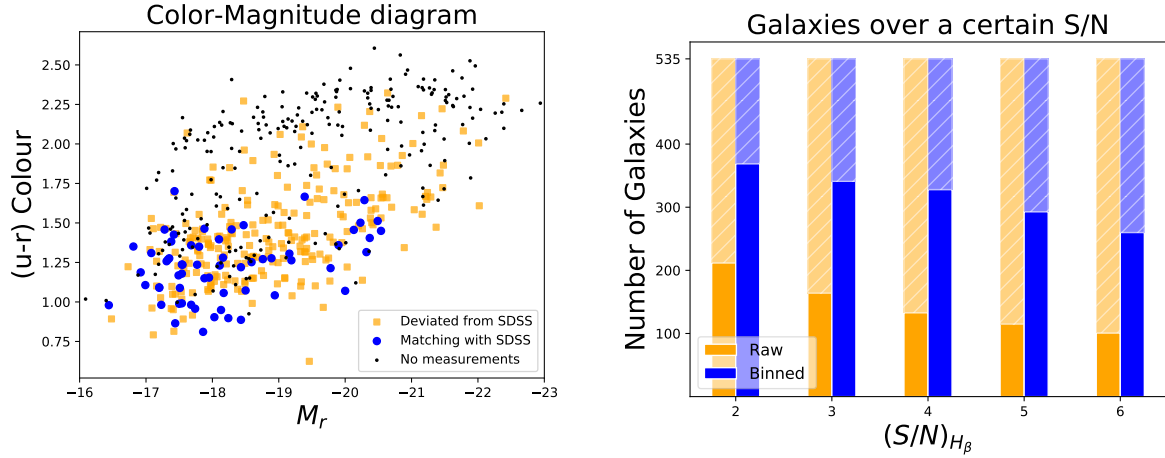
## 4.1 Quality of measurements

To determine the quality of our data we have run two diagnostics. First, we have plotted our Balmer decrement measurements against SDSS Balmer decrements (see Figure 9b). Second, we have plotted a color-magnitude diagram of all 535 galaxies (see Figure 10a).

In both diagnostics, we can see that Balmer decrements from blue and/or gas-dominated galaxies tend to be closer to SDSS results than Balmer decrements from red and/or star-dominated galaxies.

When white light generated inside a star passes through its chromosphere, absorption lines are generated. Those photons with exactly 10.2 eV of energy will not make it through the star chromosphere, because they will be absorbed by hydrogen atoms jumping from the ground state energy level to the first excited energy level. Likewise,





(a) The color-magnitude diagram shows that galaxies with a *RESOLVE* Balmer close to the SDSS measurement (blue points) are in the blue-cloud (mostly spirals dominated by gas). The orange squares are galaxies whose Balmer decrement doesn't match the SDSS measurement. The black points are galaxies with no Balmer decrement available.

(b) Comparison of the number of galaxies above a given S/N (x-axis) before the binning (orange) and after the binning (blue). For the raw data, the S/N was calculated at the center position of the galaxy for the line  $H_\beta$ . The striped lighter bars are the number of galaxies below the target  $(S/N)_{H_\beta}$ .

Figure 10

no light with an energy of 1.89 eV will make it through; those photons will be absorbed by hydrogen atoms jumping from the first excited energy level to the second excited energy level. The light that the star chromosphere absorbs shows up as dips (absorption lines) in the measured spectrum.

It is common to see spectra with emission and absorption lines of different intensities at the same wavelength. This makes the task of fitting a line a bit harder than just a Gaussian fit, which is what the code used to supply our initial inputs performs. Absorption lines are not accounted for in our fits, meaning that the flux of those galaxies with strong absorption lines (red galaxies dominated by stars) will be miscalculated. Incorrect fluxes will lead to erroneous Balmer decrements.

In Figure 9b we can see that gas-dominated galaxies are closer to the expected line than star-dominated galaxies. The disagreement between our measurements and SDSS measurements is explicit in figures 9b and 10a.

Note that a good check that confirms our code to be working correctly is that most of the galaxies without measurement are situated in the upper region of the color-magnitude diagram, the 'red sequence'. This is where the 'dead galaxies' are. They don't yield a measurement after the binning because have no emission lines.

Another source of error we suspect might be affecting our results is an inaccurate flux calibration due to a mismatched response function.

In spectroscopy, a response function aims to model and correct for instrumental and atmospheric effects by measuring the spectrum of a standard star and dividing by the known spectrum of that star. This will leave us with the combined response function, which can be used to correct any other spectroscopic measurement taken at close to the same time. The response function we are currently using was made based on a standard star measured in 2014. Preparing response functions with standard stars taken closer in time to the galaxy spectra will improve our flux calculations and therefore our Balmer decrement results.

## 4.2 Improvement of measurements

We have calculated the number of galaxies with an  $H_\beta$  S/N greater than 2, 3, 4, 5, 6 before the binning and after the binning. As expected, the latter is greater than the former. Averaging the increment of galaxies with a signal to noise greater than 2, 3, 4, 5, 6, after the binning yields averages a factor of  $\sim 2$  in comparison with the number of galaxies above those S/N before the binning. Details of this comparison can be seen in Figure 10b

## 5 Conclusions

After the binning we found a that the number of galaxies with a S/N higher than 3 increases by a factor of 2.1. This factor rises up to 2.3 for galaxies with a S/N higher than 5.

Two sources of error have been detected in our diagnostics. First, absorption lines need to be modeled in order to successfully compare SDSS Balmer decrements with the ones from RESOLVE. Second, a more recent standard star calibration needs to be carried out.

## 6 Future Work

There are 3 main future goals for this code to develop further.

There is an arbitrary prebinning that happens before the code is applied. The rows in the 2D spectrum are binned into 21 rows before the fits are performed. This is done to obtain a high enough S/N in the different spatial position of the galaxy allowing for good first guesses of where the lines are. Applying this prebinning to all galaxies is not the best practice, because while it might be convenient for some galaxies, it is not for all of them. In fact, the goal of this code is to determine what the optimal binning for each galaxy is. The 21 rows prebinning would be particularly inconvenient, for example, for galaxies with high S/N. By over-binning its rows, we would be losing much spatial information very valuable if one interested in calculating metallicity gradients.

The first future goal is to stop applying this prebinning while still being able to have

good first guesses for the observed wavelength of the lines at each spatial position of the 2D spectrum. The pixels will be binned in sets of 3 rows or more if needed. This is justified because our seeing is  $\sim 0.8\text{--}0.9$  arcsec. That value for the seeing implies that the minimum spatial scale of information contained in 3 pixels, meaning that within a group of 3 pixels, one would not obtain new information.

Due to this new binning, some of the lines might have very low S/N. By implementing a S/N threshold we would bin the minimum amount of rows in order to satisfy it. Another way to get good first guesses for the velocity at each positions with low S/N would be to interpolate based on the neighbor rows.

The second goal is to implement the reevaluation of the fits to the spectrum depicted in figure 3.

The binnig code only excludes the center row if the AGN flag is set to a positive value. However, there are simulations suggesting that super massive black holes in dwarf galaxies can wander, and not be exactly situated in the middle (up to 5 kpc off the center of the galaxy). The third future goal would then be to implement an option that allows the user to exclude off-center AGNs.

The last future goal is to apply this code to real science cases, namely, using the new Balmer decrements for a better extinction correction and bin different lines to obtain optimized line ratios useful for AGN classification.

## References

- [1] D. Ilić, L. Č. Popović, G. La Mura, S. Ciroi, and P. Rafanelli. The analysis of the broad hydrogen balmer line ratios: Possible implications for the physical properties of the broad line region of AGNs. *Astronomy & Astrophysics*, 543:A142, jul 2012.
- [2] D. Thomas, O. Steele, C. Maraston, J. Johansson, A. Beifiori, J. Pforr, G. Strömbäck, C. A. Tremonti, and D. Wake and. Stellar velocity dispersions and emission line properties of SDSS-III/BOSS galaxies. *Proceedings of the International Astronomical Union*, 8(S295):129–132, aug 2012.
- [3] Kewley, L. J., M. J. Geller, and R. A. Jansen. [OII] as a sfr indicator and the cosmic star formation history. *American Astronomical Society Meeting Abstracts*, 2003.
- [4] Guinevere Kauffmann, Timothy M. Heckman, D. M. Simon White, Stéphane Charlot, Christy Tremonti, et al. Stellar masses and star formation histories for  $10^5$  galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 341(1):33–53, may 2003.
- [5] L. J. Kewley, M. A. Dopita, R. S. Sutherland, C. A. Heisler, and J. Trevena. Theoretical modeling of starburst galaxies. *The Astrophysical Journal*, 556(1):121–140, jul 2001.
- [6] Sheila J. Kannappan, Lisa H. Wei, Robert Minchin, and Emmanuel Momjian. Galaxy gas fractions, characteristic mass scales, and the RESOLVE survey. In *AIP Conference Proceedings*. AIP, 2008.
- [7] Christy A. Tremonti, Timothy M. Heckman, Guinevere Kauffmann, Jarle Brinchmann, Stephane Charlot, Simon D. M. White, Mark Seibert, Eric W. Peng, David J. Schlegel, Alan Uomoto, Masataka Fukugita, and Jon Brinkmann. The origin of the mass-metallicity relation: Insights from 53,000 star-forming galaxies in the sloan digital sky survey. *The Astrophysical Journal*, 613(2):898–913, oct 2004.
- [8] J. Brinchmann, S. Charlot, S. D. M. White, C. Tremonti, G. Kauffmann, T. Heckman, and J. Brinkmann. The physical properties of star-forming galaxies in the low-redshift universe. *Monthly Notices of the Royal Astronomical Society*, 351(4):1151–1179, jul 2004.