Study of hidden mergers in double-peaked emission-line galaxies

Author: Alberto García Ribas
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Josep Maria Solanes Majúa

Abstract: We present a study of double-peaked (DP) emission-line galaxies at $z < 0.1$ based on a subset of 84,492 spectra extracted from the SDSS catalog. We have performed an automated selection procedure of DP spectra by adapting a previously existing routine developed for stars. A total of 5,682 galaxies of the parent database have been found to exhibit a significant double peak in at least one of the four emission lines traditionally used in the Baldwin, Phillips and Terlevich (BPT) diagrams to discriminate among different ionization mechanisms. A preliminary investigation on the possible existence of an anti-correlation between the ratios of fluxes from the [OIII]λ5007 line and velocity shifts of the red and blue components of the DP lines has produced results independent of the BPT classification of the galaxies. For most of the outliers of this relationship we find that the position of the strong [OIII] peak is significantly closer to the stellar velocity than for the weak [OIII] peak. These off-centered outliers could be linked to minor mergers in which central star formation has been enhanced or to objects with asymmetric gas outflows.

I. INTRODUCTION

The presence of narrow emission lines in galaxy spectra with a double Gaussian-like profile is known since the 1980s [1]. Initially it was mainly interpreted as an indicator of outflows from active galactic nuclei (AGN) or rotating narrow line regions (NLR) [2] within a single galaxy. Yet, the identification and study of samples of double-peak (DP) galaxies has gained considerable interest after the realization [3] that they can also come from binary galaxy systems that are in their ultimate stages of merger and, therefore, too close to each other to be seen as separate entities in photometric images. Thus, systematic searches for DP galaxies have become a fundamental element when it comes to discovering mergers and therefore shedding light on issues related to the evolution of galaxies and, especially, the formation of double AGNs (DAGN).

The main limitation of DP studies based on integrated spectra is the difficulty to unambiguously identify the origin of these double peaks due to the lack of information on the spatial location of the emission. In this work, we use the statistical power of what is currently the largest database on galaxy spectra to ameliorate this problem.

DP profiles can be associated with a variety of mechanisms that are briefly summarized subsequently: (i) A gravitationally bound close binary system of black holes that are co-rotating (i.e. a DAGN). This leads to an anticorrelation [3] between the relative velocity offset of the red and blue components of the lines, $\Delta v_{\rm r}/\Delta v_{\rm b}$, and their relative intensities represented by the fluxes, $F_{\rm r}/F_{\rm b}$. Assuming that the brighter emission-line component typically comes from the more massive black hole that would normally have the lower orbital velocity due to the conservation of the total linear momentum of the bound pair, it is expected that $F_{\rm r}/F_{\rm b} \propto 1/(\Delta v_{\rm r}/\Delta v_{\rm b})$, with a constant of proportionality independent on the masses of the galaxies and of the order of -1. (ii) The rotating disk of a NLR from a single black hole. (iii) Bi-polar or bi-conical outflows driven by starbursts or by AGN radiation pressure from a single black hole. (iv) The orbital motion of the nuclei of two galaxies in the final stages of an ongoing merger in which at least one of the members of the pair is a star forming (SF) or a starburst galaxy. (v) A single galaxy with two independent SF regions (e.g. a nuclear, non-rotating one and other in the disk) or with a unique extended SF region affected by extinction bias.

Assuming that NLR are circularly rotating homogeneous gaseous disks, case (ii) is expected to produce $F_{\rm r} \sim F_{\rm b}$ and $\Delta v_{\rm r} \sim \Delta v_{\rm b}$. This relationship would be also expected in case (iii) in the particular case that outflows have relatively symmetric masses and are under momentum conservation. On the other hand, asymmetric outflows free of external forces should give rise to a relationship of the sort expected for case (i) (assuming that the radiation from the outflows is proportional to the mass of the emitting gas). For all these cases, the ratios of fluxes and velocity offsets of the two components of the DP lines should be closely related.

Instead, AGN and starburst-driven outflows that do not conserve momentum (the usual case (iii)), as well as cases (iv) and (v), i.e. systems in which emission has a thermal component, can result in important deviations form the canonical relationship of slope -1 expected for DAGN. Non-conservative outflows should also produce differences in the velocity dispersion of the two lines.

In this work, we have adapted an existing procedure that automatically identifies DP narrow emission lines in stars to search for DP in galaxy spectra. In contrast to many previous studies of this kind, we do not constrain our search to dual AGN candidates or to a visual selection of galaxies. We base our work on the analysis of several tens of thousands of optical single-fiber spectra retrieved from the Main Galaxy Sample of the Sloan Legacy Survey [4].

The manuscript is organized as follows. In Section II, we describe the selection of data and the procedure implemented for double-peak identification. The criteria

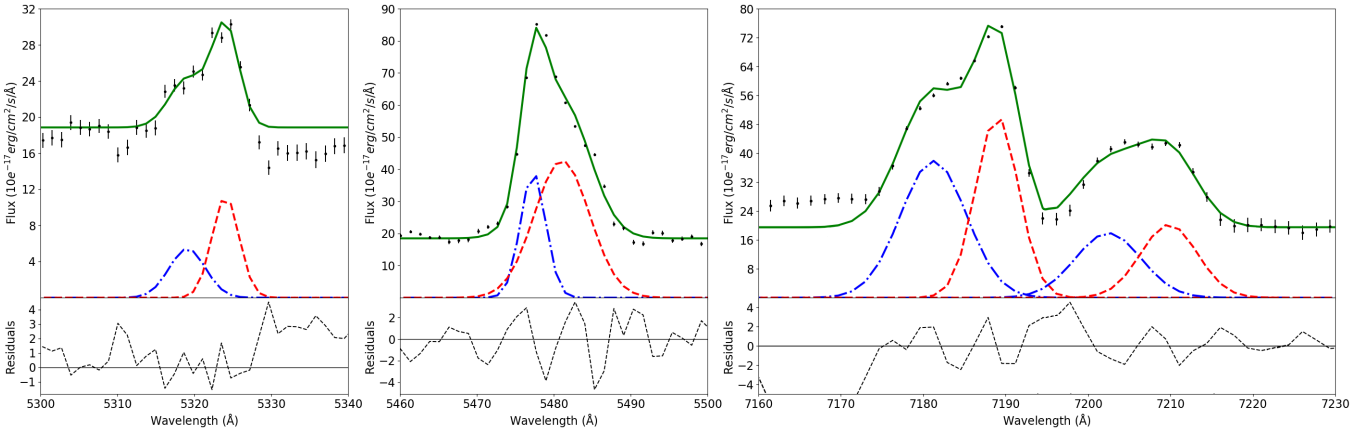

FIG. 1: Graphic representation of the DP emission-line fitting for the galaxy J143553.75+075652.1. On the upper section, blue and red dashed lines indicates the blue and redshift of each double-peak whereas the green line represents the fitting of both Gaussian functions. Residuals between the proposed model and the experimental data are showed on the lower section of each graph. Left and center image represents H_β and [OIII] emission lines respectively, both fitting a double Gaussian function; the right graph represents H_α and [NII] emission lines for the same galaxy, in this case both emission lines are also double-peaked.

used to identify the main DP candidates is also detailed. The relationship between the flux ratios of the redshifted (r) and blueshifted (b) components of the [OIII] line with the ratios of their corresponding velocity offsets in investigated in Sec. III, while the main results are discussed and summarized in Sec. IV.

II. DATA TREATMENT

A. Spectroscopic selection

For the present analysis, a sample of galactic spectra has been selected from the compendium of optical spectra given by the SDSS catalogue. The SDSS encompasses a family of single-fiber spectroscopic surveys comprising the Legacy Survey [5], the Baryon Oscillation Spectroscopic Survey (BOSS) [6], and the Sloan Extension for Galactic Understanding and Exploration (SEGUE) [7], each with different strategies for targeting objects for spectroscopy motivated by differences in their scientific goals. We have chosen to work only with the objects from the Legacy Survey because of homogeneity in target selection and the fact that we want to focus on relatively nearby galaxies ($z < 0.1$). Legacy spectra relate to a magnitude-limited sample of low-redshift galaxies, a near-volume-limited sample of galaxies called Luminous Red Galaxies, and a magnitude-limited sample of quasars. There are almost a million spectra for the first two classes of galaxies, and there are more than 120,000 spectra of quasars. Spectra for all these objects are available from the SDSS 3 Science Archive Server.

For this work, we use a sample of 84,492 galaxies of all Hubble classes selected previously from the twelfth Data Release of the SDSS (SDSS-DR12) [8] by J.L. Tous through the SDSS CasJobs service. The distribution of our data is shown in Table I. All these spectra correspond to objects automatically classified by the SDSS pipeline into the 'GALAXY' class and having a zero 'zWarning' flag in order to avoid potential problems with the spectra classification or its estimated redshift value.

Galactic Morphology	Galactic Spectra	Morphological Abundance (%)	Double-peak Candidates	Double-peak Abundance (%)
E	10000	11.84	558	5.58
S0	13493	15.97	2095	15.53
Sa	15000	17.75	1824	12.16
Sb	19999	23.67	930	4.65
Sc	26000	30.77	275	1.06
Total	84492	-	5682	6.73

TABLE I: Tabular representation of the different spectra proportions based on its galactic morphology. It is also provided the number and abundance of DP candidate galaxies detected.

B. Automatic selection of double-peak emission-line galaxies

The spectra provided by the SDSS survey have been processed in two stages to select potential DP emission-line candidates. (i) The first main program determines whether or not an emission line can fit two Gaussian functions. We select each emission line in a range of 60 pixels centered in the emission-line position and transform it into wavelengths. We restrict our analysis to four different emission lines: [OIII]λ5008, $H_\alpha\lambda$6565, $H_\beta\lambda$4863 and [NII]λ6550. We select objects with a maximum peak S/N > 10 in the [OIII] line to ensure the significance of this emission line in the spectrum using:

$$\max(S/N_i) = \frac{f_i}{\sigma_i} , \qquad (1)$$

with f_i defined as the flux over the baseline and σ_i defined as the uncertainty of f_i.

Then, we require each emission line to have a mean S/N > 5. For this calculation we have used

$$\overline{S/N}_i = \frac{\sum_j f_{ij}/N_i}{(\sum_j \sigma_{ij}^2/N_j)^{1/2}} \,, \tag{2}$$

where f_{ij} is the flux of the j pixel over which i line is unambiguously detected, N_i as the number of pixels over which $f_{ij} > 3\sigma_{ij}$ and σ_{ij} is defined as the uncertainty of f_{ij}. In our case, $j = 60$ pixels and i are any of the four studied emission lines.

Once we have filtered those objects that meet our S/N requirements, we define two 12Å wide windows within the 60 pixel region were our program try to fit the Gaussian functions. Using predetermined and arbitrary Gaussian amplitudes and sigma-values, the program runs four different fitting scenarios using the `lmfit` Python library. It will choose the best fit taking into account the lowest χ^2 value among the four scenarios. A mobile continuum has been used in order to get a better estimation of the Gaussian peaks. At this point, each emission line has a single and a double Gaussian function fitted.

Further selection criteria decides if DP candidates are valid or not: (a) $1/5 < A_1/A_2 < 5$ secures that none of the two possible peaks represents only noise or one of them are suppressed. (b) $\delta\lambda > 3,5$Å demands the separation of both peaks to be greater than the resolution of the spectroscopic observation, this values has been obtained through a trial and error fitting method. (c) $\chi^2_{double} < \chi^2_{single}$ compares the χ^2-value of single and double Gaussian scenario and discards poor fitting double peaks. For each emission line selected, the program returns the velocity of both single or double Gaussian functions for all spectra.

(ii) The second main program uses the velocity of each component and readjusts both Gaussian functions using a fixed continuum. The program returns the characteristic sigma-value for each Gaussian function, as well as the position and the amplitude of the peak. It also provides the value of the area under the curve (in relation to the flux) and the velocity of any given peak. The velocity of each component has been calculated following

$$v_j = c \cdot \left(\frac{\lambda_j}{\lambda_{ref}} - 1 \right) , \tag{3}$$

where c represents the speed of light ($c = 299,792.458$ m/s), λ_j and λ_{ref} stands for the position of the component and the position of the emission line, respectively. The j subscript represents either blue or red component.

At this point every emission line of each spectra has a single or double Gaussian function fitted. We implemented two different conditions in order to decide which of the DP preliminary candidates can be considered DP galaxies. We decided that the presence of a double-peak in the strongest emission line is a sufficient condition to consider the candidate as a significant DP galaxy. In case this is not fulfilled, all those galaxies with two or more

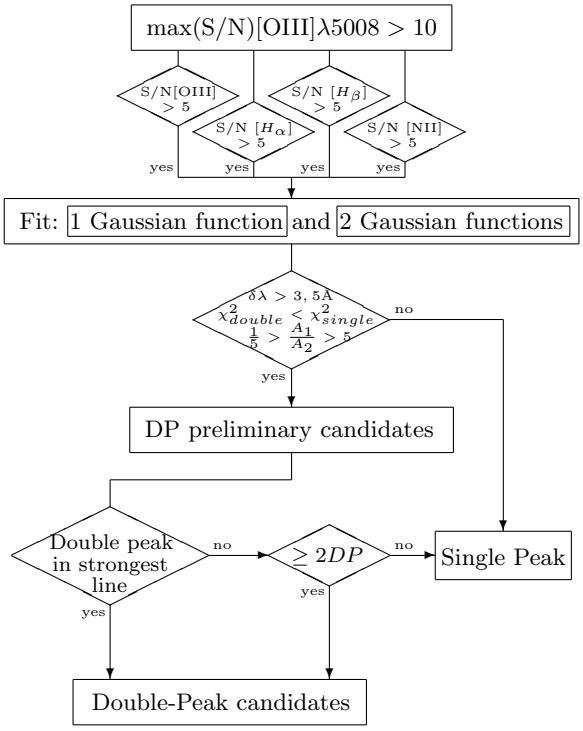

FIG. 2: Flowchart describing the main stages of our automatic fitting and selection process. Firstly, we select preliminary candidates using emission-line properties and the minimum χ^2 value among the single or double-peak scenario. The second stage filters the preliminary candidates according to our DP criteria (DP on the main emission-line or at least two DP emission-lines per candidate).

DP emission lines are also considered as DP galaxies. As a result of this procedure we reduced the original 84,492 sample to 5682 spectra, representing the 6.73% of the main sample.

C. Determination of the blueshift and redshift components

Each DP component experiences a velocity shift from the stellar velocity. To compute the value of the blue and red component velocity offset Δv_j^{gal}, we consider the contribution of each j component for all the DP emission lines identified in each spectrum. We use a weighted median calculation

$$\Delta v_j^{gal} = \frac{\sum_i v_{ji} \cdot \omega_i}{\sum_i \omega_i} \,, \tag{4}$$

where the i subscript refers to the number of DP identified, j subscript stands for the blue or red component and v_{ji} is the velocity shift of the given component calculated using equation (3). The weight ω_i is defined as

$$\omega_i = (\overline{S/N}_i)^2 \,, \tag{5}$$

with $\overline{S/N}_i$ as the average signal to noise value of the i peak, defined using equation (2). We decided to use a weighted average instead of an arithmetic average since it takes in consideration the contribution of every double-peak identified. Weighty components will contribute more to the overall velocity shift than the lighter ones.

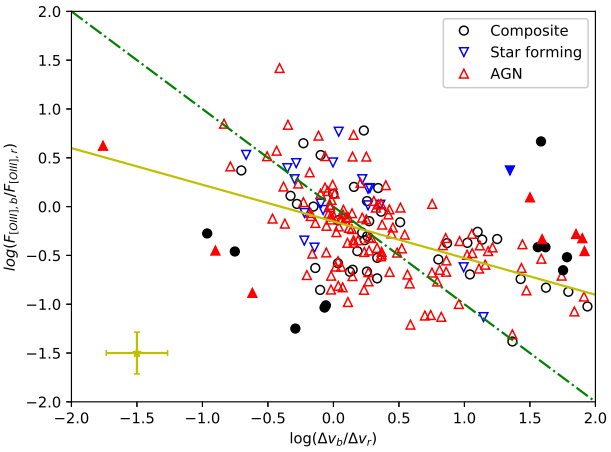

FIG. 3: Graphic representation of the correlation between the logarithm of blue/redshift velocity ratio and the logarithm of the blue/redshift [OIII] flux ratio. Different markers represent different types of galaxies according to the BPT classification. A linear fit of all the galaxies is plotted in yellow. Solid markers indicates the points furthest from the trend. The theoretical behaviour is represented by the dashed green line with a slope of -1. A yellow point with the mean error associated with the points is also showed.

III. ANALYSIS

A. Classification according to the BPT diagnostic diagram

The BPT diagnostic diagrams [9] are a set of nebular emission line diagrams used to distinguish the ionization mechanism of nebular gas. The most used version makes use of the reddening-corrected flux (intensity) ratios [NII]6584/H_α versus [OIII]5007/H_β. Different positions on the graph represent different ionization mechanisms.

We now proceed to classify our DP candidates using the NSA spectral classification. We compare our sample with a catalog which has been classified according to the BPT diagnostic diagram, obtaining the proportions shown in Table (II). S0 and Sa class galaxies present a higher proportion of DP candidates, 15.53% and 12.16% of their respective totals. AGN type appear in a higher proportion in E and S0 than the other morphologies while SF type is more common in spiral galaxies (Sa, Sb and Sc). The proportion of Composite galaxies is similar for all morphological classes and Other-type galaxies abundance is marginal, representing only the 0.41% of all the DP candidate sample.

B. Correlation between velocity-offsets ratio and flux ratio

We focus our study on those candidates which present a significant double-peak on the [OIII] emission line in order to compare our results. This property has been associated with an indicator of AGN. Then, our sample of 5682 DP candidates is reduced to 221 galaxies which have at least a significant double-peak on the [OIII] emission line.

In order to analyze the properties and the origin of the double-peaked [OIII], we measure the ratios of velocity shifts and the flux ratio of both components. We define $\Delta v_b = |v_b - v_0|$ and $\Delta v_r = |v_r - v_0|$ as the shifted velocity-offset ratios between the stellar velocity v_0 of the host galaxy and its blue and red components, v_b and v_r respectively. Both velocity shift components have been calculated using equation (4). $F_{[OIII],b}$ and $F_{[OIII],r}$ are the flux of blue and red component of the fitted [OIII] emission line taken from the data obtained in section II B. We define the ratios of velocities and fluxes as $\Delta v_b/\Delta v_r$ and $F_{[OIII],b}/F_{[OIII],r}$ respectively. Fig 3 shows the correlation between velocity shifts and flux ratios:

$$\frac{F_{[OIII],b}}{F_{[OIII],r}} \propto \left(\frac{\Delta v_b}{\Delta v_r}\right)^{-1}. \tag{6}$$

Fig 3 show the presence of asymmetries towards the blushifted component on the velocity ratio. We identified the furthest decile from our linear fit and found that some of them match with the asymmetric behaviour. In order to interpret this anomaly, we plotted the velocity shift ratio between each strongest and weakest component of each galaxy with its respective sigma-value as it is showed in Fig 4. We found that most of the decile has a strong component with a velocity similar to the stellar one.

IV. CONCLUSIONS

We have identified 5,682 DP candidate galaxies from a subset of 84,492 objects of different morphologies with z < 0.1 extracted from the Main sample of single-fiber optical galaxy spectra of the SDSS. The identification has relied on an automated procedure that fits a single and a double Gaussian function to narrow emission lines that pass several selection criteria aimed at ensuring the reliability of the results. For a galaxy to be identified as a DP candidate we have required that its spectrum exhibits a significant double-peak in at least one of the following lines: $H_\beta\lambda4863$, [OIII]$\lambda5008$, $H_\alpha\lambda6565$, and [NII]$\lambda6586$.

Relying on a machine-learning-based morphological classification [10], we have found an important excess of S0 and Sa galaxies, that represent the 39.8% and 31.2%,

BPT	E		S0		Sa		Sb		Sc	
classification	Number	Fraction (%)	Number	Fraction (%)	Number	Fraction (%)	Number	Fraction (%)	Number	Fraction (%)
Composite	226	40.50	771	36.80	723	39.64	354	38.06	95	34.55
Star Forming	111	19.89	445	21.24	585	32.07	312	33.55	125	45.45
AGN	218	39.07	867	41.38	509	27.91	263	28.28	55	20.00
Others	3	0.54	12	0.57	7	0.38	1	0.11	-	-

TABLE II: Tabular representation of the different galaxy types according to the BPT diagnostic classification for each galactic morphology. In this chart are shown the number and proportions relative to its total for each morphological type.

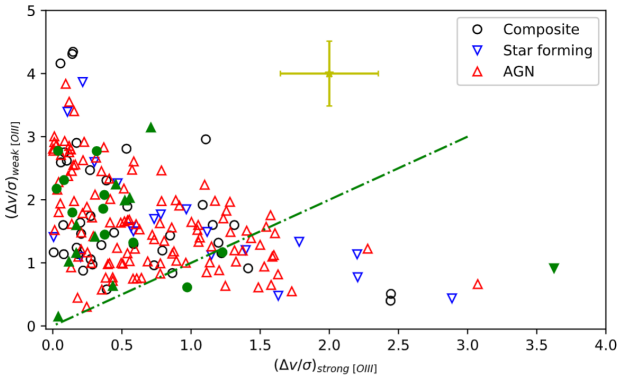

FIG. 4: Graphic representation of the ratio between the velocity offsets and its sigma-value for each weak and strong component for all double-peaked [OIII] emission lines. Different markers represent different type galaxies according to the BPT classification and solid green markers represents the outlier ones. A yellow point with the mean error associated with the point is also showed.

respectively, of the DP galaxies when one adopts an equal number of representatives of each Hubble type. These fractions are larger than the usual fractions observed in magnitude-limited samples, but are fully consistent with the findings of a similar analysis by D.Maschmann, 2020 [11]. Regarding the BPT classification, the most abundant active class in our DP sample is Composite (38.2%), followed by AGN (33.7%), being the S0-AGN the most frequent kind of galaxy.

We have also selected a subset of 221 galaxies with double-peaked [OIII]λ5008 profiles to investigate the existence of a possible correlation between the ratios of the double peak fluxes and of the velocity shifts of the blueshifted and redshifted lines. This correlation has been extensively studied in the literature as it can be naturally explained by the Keplerian relation predicted by models of co-rotating DAGNs. A plot comparing these two ratios as a function of the BPT class of the galaxies shows that there are no significant differences in the distribution of the data. Taking into account the errors in the measurements, we have found that most galaxies follow reasonably well the canonical relationship of slope -1 expected for DAGN, while detecting a tendency of DP candidates to accumulate in the region where these ratios tend to the unit.

We have also compared the velocity offsets of the two [OIII] emission-line components relative to the stellar velocity of the host galaxy in units of their velocity dispersion. For the majority of the galaxies the position of the strong peak is found closer to the stellar velocity than that of the weak peak. Applying ranking statistics to the residuals from the canonical relationship between fluxes and velocity-shifts ratios, we have identified as potential outliers those measurements that fall into the upper decile. Our analysis shows that the outliers are also frequently found among the most asymmetric velocity offsets for which the position of the strong [OIII] essentially matches the systemic velocity. In the absence of a more detailed investigation, we tentatively conclude that these off-centered outliers could be linked to minor mergers in which central star formation has been enhanced or to objects with asymmetric gas outflows.

Though this crude study seems to favor that most of the [OIII] double-peaked sources are DAGNs, further work should be done with larger statistical samples of optical spectra and with more emission lines in order to tackle the elusive nature of the DP feature in galaxies.

Acknowledgments

I would like to thank everyone who helped or supported me in any way, specially J.M. Solanes, J.L. Tous for their guidance and suggestions and J. Carbajo for giving us the original code, without his help this work could not have been done. To everyone who asked about this work and showed interest in my explanations.

[1] Whittle, M. 1985, MNRAS, **213**, 1
[2] Greene, J.E.; Ho, L.C. 2005, ApJ, **627**, 721
[3] Wang, J.-M.; et al. 2009, ApJ, **705**, L76
[4] Strauss M.A.; et al. 2002, AJ, **124**, 1810
[5] Abazajian, K.N.; et al. 2009, ApJS, **182**, 543
[6] Dawson, K.S.; et al. 2013, ApJ, **145**, 10
[7] Aihara, H.; et al. 2011, ApJS, **193**, 29
[8] Alam, S.; et al. 2015, ApJS, **219**, 12
[9] Baldwin, J.A.; et al. 1981, PASP, **93**, 5
[10] Domínguez Sánchez, H.; et al. 2018, MNRAS, **476**, 3661
[11] Maschmann, D.; et al. 2020, A&A, **641**, A171