

# H $\alpha$ emitters from the Southern Photometric Local Universe Survey (S-PLUS)

L. A. Gutiérrez-Soto<sup>1</sup><sup>★</sup>, R. Lopes de Oliveira<sup>1,2,3</sup>, S. Akras<sup>4</sup>, D. R. Gonçalves<sup>5</sup>, C. Mendes de Oliveira<sup>1</sup>, F. Almeida-Fernandez<sup>1</sup>, F. R. Herpich<sup>1</sup>, A. Kanaan<sup>6</sup>, T. Ribeiro<sup>7</sup>, W. Schoenell<sup>8</sup>

<sup>1</sup>Departamento de Astronomia, IAG, Universidade de São Paulo, Rua do Matão, 1226, 05509-900, São Paulo, Brazil

<sup>2</sup>Departamento de Física, Universidade Federal de Sergipe, Av. Marechal Rondon, S/N, 49100-000, São Cristóvão, SE, Brazil

<sup>3</sup>Observatório Nacional, Rua Gal. José Cristino 77, 20921-400, Rio de Janeiro, RJ, Brazil

<sup>4</sup>Institute for Astronomy, Astrophysics, Space Application & Remote Sensing, National Observatory Athens, GR-15236, Athens, Greece

<sup>5</sup>Observatório do Valongo, Universidade Federal do Rio de Janeiro, Ladeira Pedro Antonio 43, 20080-090, Rio de Janeiro, Brazil

<sup>6</sup>Departamento de Física, Universidade Federal de Santa Catarina, Florianópolis, SC, 88040-900, Brazil

<sup>7</sup>NOAO, P.O. Box 26732, Tucson, AZ 85726

<sup>8</sup>GMTO Corporation 465 N. Halstead Street, Suite 250 Pasadena, CA 91107

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

The ongoing multi-colour survey performed by the S-PLUS project will have covered 9300 deg<sup>2</sup> of the Southern skies by the time it is completed. S-PLUS has a crucial feature: images over the whole ~~area~~<sup>field</sup> taken in the H $\alpha$  narrow-band. The H $\alpha$  transition provides a superb tool for the study of a number of important astrophysical processes and, in particular, it allows the classification of different types of astrophysical sources. Here we explore the S-PLUS data release 3, which covers 2000 deg<sup>2</sup>, including the Stripe-82 area, to highlight the potential of the survey for finding H $\alpha$  emitters using the (r - J0660) versus (r - i) color-color diagram and in distinguishing the red from the blue sources based on the (r - i) versus (g - z) diagram. Our H $\alpha$ -emitter catalog contains 8,446 objects that exhibit excess in the narrow J0660 band. For 8,039 of them, the excess is thought to be due to the H $\alpha$  emission line, while for the remaining, the excess may be due to redshifted lines. Unsupervised clustering machine learning approach reveals two distinct populations: one with an intense blue continuum and another one with a red continuum. The hierarchical agglomerative clustering algorithm (HAC) was compared with the hierarchical density-based cluster selection (HDBSCAN) in order to reinforce the robustness of the red and blue populations' classification. By adopting a so-called “soft” clustering approach, we assigned the probability of each emitter belonging to a given population, blue or red. Around 84% of emitters were successfully classified as blue or red sources. We use synthetic and observed spectra to emphasize the potential of color-color diagrams in distinguishing several classes of H $\alpha$  emission-line emitters that include planetary nebulae, H II regions, young stellar objects, symbiotic stellar systems, cataclysmic variables, blue compact galaxies, star-forming galaxies, and quasars. In summary, the method described in detail in this paper is shown to be an efficient tool to find new emitters and to classify them, using multi-color data.

**Key words:** surveys – techniques: photometric – stars: novae, cataclysmic variables – galaxies: dwarf – quasars: emission lines

## 1 INTRODUCTION

Atomic excitation followed by recombination in Balmer hydrogen emission lines may be ignited in different ways, thermal and non-thermal collisional excitation in shock-heated gas and energetic photons acting over a diffuse gas. As a practical result, and the Universe being hydrogen abundant, the observation of those electronic transitions offer an important window into the study of astrophysical objects. Among all the possible electronic transitions, the Balmer series represent extremely useful tools in Astronomy. Particularly, the H $\alpha$

emission line – rest-frame wavelength of @6564.614 Å at vacuum – that corresponds to the electron transition from the  $n = 3$  to the  $n = 2$  energy level, is the strongest one, in both emission or absorption, and the most widely used to identify various types of objects (e.g star-forming regions, H II regions, PNe, supernovae, novae, circumstellar enveloped among others). Hydrogen recombination lines trace a vast variety of sources such as young stellar objects (YSO), Herbig-Haro objects, circumstellar disks, post-asymptotic and asymptotic giant stars (AGB), red giant stars (RGB), active late-type dwarfs. Amongst massive stars, emission lines are observed in Be stars with decréation disks, Wolf-Rayet (WR) stars, interacting binary systems that

\* E-mail: gsoto.angel@gmail.com

experiencing mass exchange like symbiotic stars (SySt), cataclysmic Variables (CVs), among others.

At much larger scales, the H $\alpha$  emission line can also be emitted by planetary nebulae, H II regions or star-forming regions in galaxies, novae and supernova remnants, as well as other galaxies. In the case of high redshifted sources like starburst galaxies and quasi-stellar object (QSOs), the detection of an emission at 6563 Å is not associated with the recombination of H $\alpha$  but with other UV emission lines.

Most of the aforementioned classes of objects are not homogeneous and far from complete even in the local Universe, with some being highly populated while others being highly underrepresented. For example, there are  $\sim 320$  known SySt, with only  $\sim 65$  of those located in galaxies other than the Milky Way (Akras et al. 2019a; Merc et al. 2019). The number of known PNe in our Galaxy is of the order of  $\sim 3500$  (Parker et al. 2016), which may represent only 15–30% of the total population (Frew 2008; Jacoby et al. 2010).

H $\alpha$  surveys in a variety of angular resolutions, sky coverage, and sensitivity were carried out in the past. Some of them, with modest spatial resolutions, revealed spatially resolved, extended nebular emission to study supernova remnants, galaxy groups, and star-forming regions (e.g. Davies et al. 1976). Others with higher spatial resolution disclosed compact emission-line sources in the Galaxy and sources in nearby galaxies. Examples are the INT Photometric H $\alpha$  survey (IPHAS; Drew et al. 2005; Barentsen et al. 2014), the SuperCOSMOS H $\alpha$  Survey with the UK Schmidt Telescope (UKST) of the Anglo-Australian Observatory (Parker et al. 2005), and the ongoing VST Photometric H $\alpha$  Survey (VPHAS+; Drew et al. 2014).

Traditionally, H $\alpha$  emitters are revealed directly from images and in colour-colour diagrams from photometric surveys observing the sky with at most five – generally broad-band or H $\alpha$  – filters. For example, the ( $r - H\alpha$ ) versus ( $r - i$ ) colour-colour diagram or a similar diagram has been used to find CVs (Witham et al. 2006, 2007), YSOs (Vink et al. 2008), SySt (Corradi et al. 2008; Corradi & Giammanco 2010; Corradi et al. 2011; Akras et al. 2019b), early-type emission-line stars (Drew et al. 2008), and PNe (Viironen et al. 2009; Sabin et al. 2010; Akras et al. 2019c).

There are two ongoing multi-band surveys observing the sky in a systematic, complementary way, with 5 broad and 7 narrow-band filters, including H $\alpha$ : the Javalambre Photometric Local Universe Survey (J-PLUS<sup>1</sup>; Cenarro et al. 2019), covering the Northern celestial hemisphere, and the Southern-Photometric Local Universe Survey (S-PLUS<sup>2</sup>; Mendes de Oliveira et al. 2019), covering the southern sky with a twin 80 cm telescope. These are paving the way for an even more ambitious survey, the Javalambre Physics of the Accelerating Universe Astrophysical Survey (J-PAS; Benítez et al. 2014 and miniJ-PAS; Bonoli et al. 2021), which will observe the Northern sky with 56 narrow-band filters. As source hunters, the spectral energy distributions provided by these surveys enable an unprecedented source classification using photometry only. However, in the Big Data era, efficient investigation tools are required to deal with their massive imaging and catalogues production and machine learning techniques have been increasingly used to explore these data sets.

Here we present a census of H $\alpha$  emitters from the S-PLUS DR3 by using color-color diagrams and unsupervised machine learning techniques, classifying them as blue or red sources and also proposing a class to which they belong. Section 2 describes the observations related to the S-PLUS project, as well as important information on

the third data release. It also presents the technique implemented to select the H $\alpha$  emitters and machine learning approaches used to divide the sample into two populations based on their colors. In section 3 our findings are described and finally section 4 discusses our main results and conclusions.

## 2 METHODOLOGY

### 2.1 Observations: the S-PLUS project

This paper uses data from the S-PLUS DR3, available in the database of the project, splus/cloud, and it covers 2,000 square degrees. S-PLUS is being carried out by a dedicated 0.83m robotic telescope located at Cerro Tololo, Chile (Mendes de Oliveira et al. 2019). The project is surveying the Southern sky using the 12 filters from the so-called Javalambre filter system (Marín-Franch et al. 2012), that spans the wavelength range from 3000Å to 10000Å. The system includes seven narrow-band filters ( $J0378$ ,  $J0395$ ,  $J0410$ ,  $J0430$ ,  $J0515$ ,  $J0660$ , and  $J0861$ ) and five broad-band Sloan-like (Fukugita et al. 1996) filters (see Figure 1). The narrow-band  $J0660$  filter used in S-PLUS is centered at lambda 6614 Å and has a width of about 147 Å (Table 2 of Mendes de Oliveira et al. 2019), and therefore it covers both the H $\alpha$  and the doublet [N II]  $\lambda\lambda 6548, 6584$  spectral lines for sources up to a redshift of approximately 0.02.

The data set used for this study, DR3, includes about 60 million objects distributed over  $\sim 2,000$  deg $^2$  (of the total of  $\sim 8,000$  deg $^2$  of high Galactic latitudes fields with  $b > 30^\circ$  planned to be covered when the survey is complete). Note that the area including the Galactic disk and bulge was not included in this study (S-PLUS plans to cover  $\sim 1,300$  deg $^2$  of such areas), given that both Galactic areas will be available only in DR4. Amongst the different aperture photometry available in the catalog we have used the PStotal photometry, which is a 3-arcsec aperture corrected magnitudes (Almeida-Fernandes et al. 2022). In order to ensure that high-quality data are used in the present analysis, only objects detected in at least the  $r$ ,  $i$  and  $J0660$  bands, simultaneously, with errors less than 0.2 mag, were considered.

The first goal of this paper is the identification of H $\alpha$  emitters in the S-PLUS DR3. For this, we applied an iterative and automatic technique to select objects with an excess in the  $J0660$  band, which is consistent with the detection of the H $\alpha$  line in emission. Next, the sample of H $\alpha$  sources is divided into two subgroups: the blue and red one. This classification was made by employing optical colours in combination with unsupervised machine learning/statistical tools. These procedures are described in the following subsections.

### 2.2 Selection of H $\alpha$ emitters

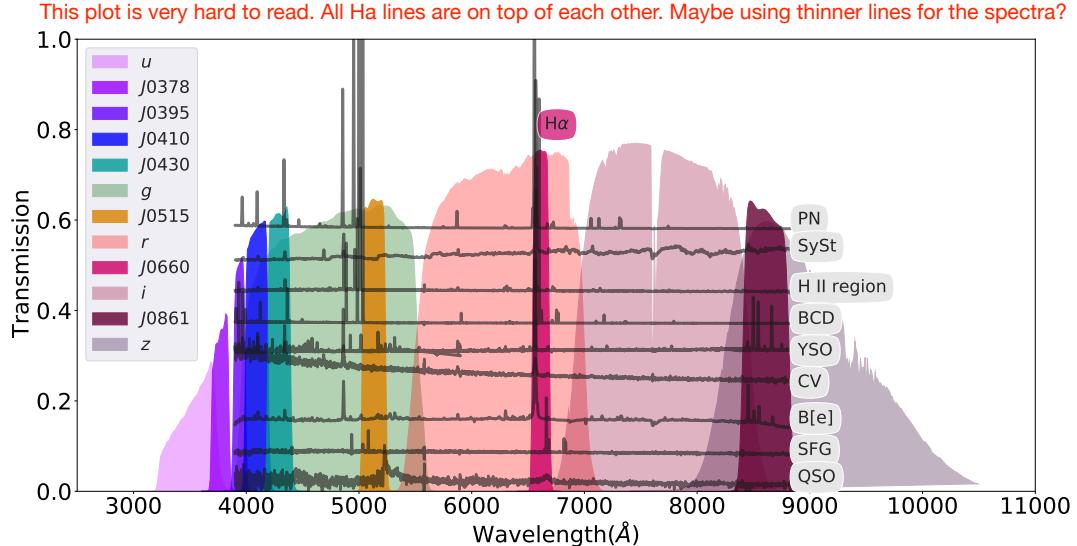
Before search for the H $\alpha$  emitters, we first divided our sample into four sub-samples based on their magnitudes in the  $r$ -band. Thus, we considered the following four sub-samples: (i)  $r$ -band  $< 16$ , (ii)  $16 \leq r < 18$ , (iii)  $18 \leq r < 20$ , and  $20 \leq r < 21$ . In this way, we avoid mixing up bright and faint sources with low and high uncertainties, respectively. Otherwise, the selection criteria would be potentially affected by the intrinsic scatter in the measurement of faint objects.

The identification of H $\alpha$  emitters is based on the method successfully applied by Witham et al. (2008) to the IPHAS project, given that similar filters to latter are also available in S-PLUS:  $r$ ,  $J0660$ , and  $i$  filters. The same technique was used by Scaringi et al. (2013) and Wevers et al. (2017) to reveal H $\alpha$  emitters.

We first generated the ( $r - J0660$ ) versus ( $r - i$ ) diagram for each

<sup>1</sup> <https://www.j-plus.es>

<sup>2</sup> <http://www.splus.iag.usp.br>



**Figure 1.** Transmission curves of the S-PLUS filter set. The narrow-band filter  $J0660$  includes the  $H\alpha$  emission line. Over-plotted are spectra of different classes of emission line objects. From top to bottom: a PN, a symbiotic star, an extragalactic H II region, a blue compact/H II galaxy, a YSO, a CV star, a B[e] star, a star forming galaxy and a QSO at a redshift of  $\sim 3.1$ .

magnitude bin and attempted to fit the loci mainly occupied by main-sequence and giant stars with a linear regression fit. We then implemented an iterative  $\sigma$ -clipping technique so that, by construction,  $H\alpha$  emitters should satisfy the condition:

$$(r - J0660)_{\text{obs}} - (r - J0660)_{\text{fit}} \geq C \times \sqrt{\sigma_s^2 - \sigma_{\text{phot}}^2} \quad (1)$$

where  $\sigma_s$  is the root mean squared value of the residuals around the fit and  $\sigma_{\text{phot}}$  is the error on the observed  $(r - J0660)$  colour index.  $C$  is a constant parameter with a value of 4 following Wevers et al. (2017). The fits were made by employing `astropy.modeling`<sup>3</sup>

Figure 2 illustrates the procedure applied. The solid black lines indicate the initial fit and the dashed lines show the  $4-\sigma$  clipping fit. The dotted lines correspond to the selection criteria for the  $H\alpha$  emitters,  $4-\sigma$  above of the final fit. It should be noted that these cut-off lines are only approximations because only the residual around the fit is taken into account. The photometric uncertainty of the  $r - J0660$  colour index for each individual point is also taken into account (see Equation 1).

Once the list of  $H\alpha$  emitters was obtained, we proceeded with a visual inspection of their false-color images and their spectral energy distributions, constructed with 12 points corresponding to the 12 S-PLUS filter mean magnitudes for each source, hereafter called the S-spectra. The upper panel of Fig. 3 shows an example of what the S-spectrum of an  $H\alpha$  emitter looks like, while the bottom panel presents the SDSS spectrum of the same source. It is evident from the comparison of the two spectra that the excess in the  $J0660$  band is linked with the  $H\alpha$  emission line.

The distribution of the  $H\alpha$  emitters in the  $(r - J0600)$  versus  $(r - i)$  colour-colour plane is shown in Fig. 4. The loci of the main-sequence and giant stars derived from synthetic spectra (Pickles 1998) convolved with the transmission of the filters in the AB magnitude system (Oke & Gunn 1983) are also plotted. All sources located above the locus of the main and giant stars exhibit an excess in  $H\alpha$ . The wide distribution of sources across the  $(r - J060)$  and  $(r - i)$

colour-colour diagram indicates that several types of  $H\alpha$  emitters are selected. Sources with high  $(r - J0660)$  colour index are likely associated with PNe, H II regions, SySt or blue compact galaxies. On the other hand, the  $(r - i)$  colour index indicates redder sources such as SySt and YSO, while sources with strong blue continuum such as CVs and QSOs exhibit lower  $(r - i)$  values.

Fig. 5 displays the distribution of all  $H\alpha$  emitters in Galactic latitude and longitude. The density map regions represent the spatial positions of the objects on the sky. The surface density of  $J0660$ -excess objects is highest near the Galactic plane.

Our list of  $H\alpha$  emitters includes 8,446 sources. We now proceed to their classification into blue and red populations.

## 2.3 Unsupervised machine learning clustering approach

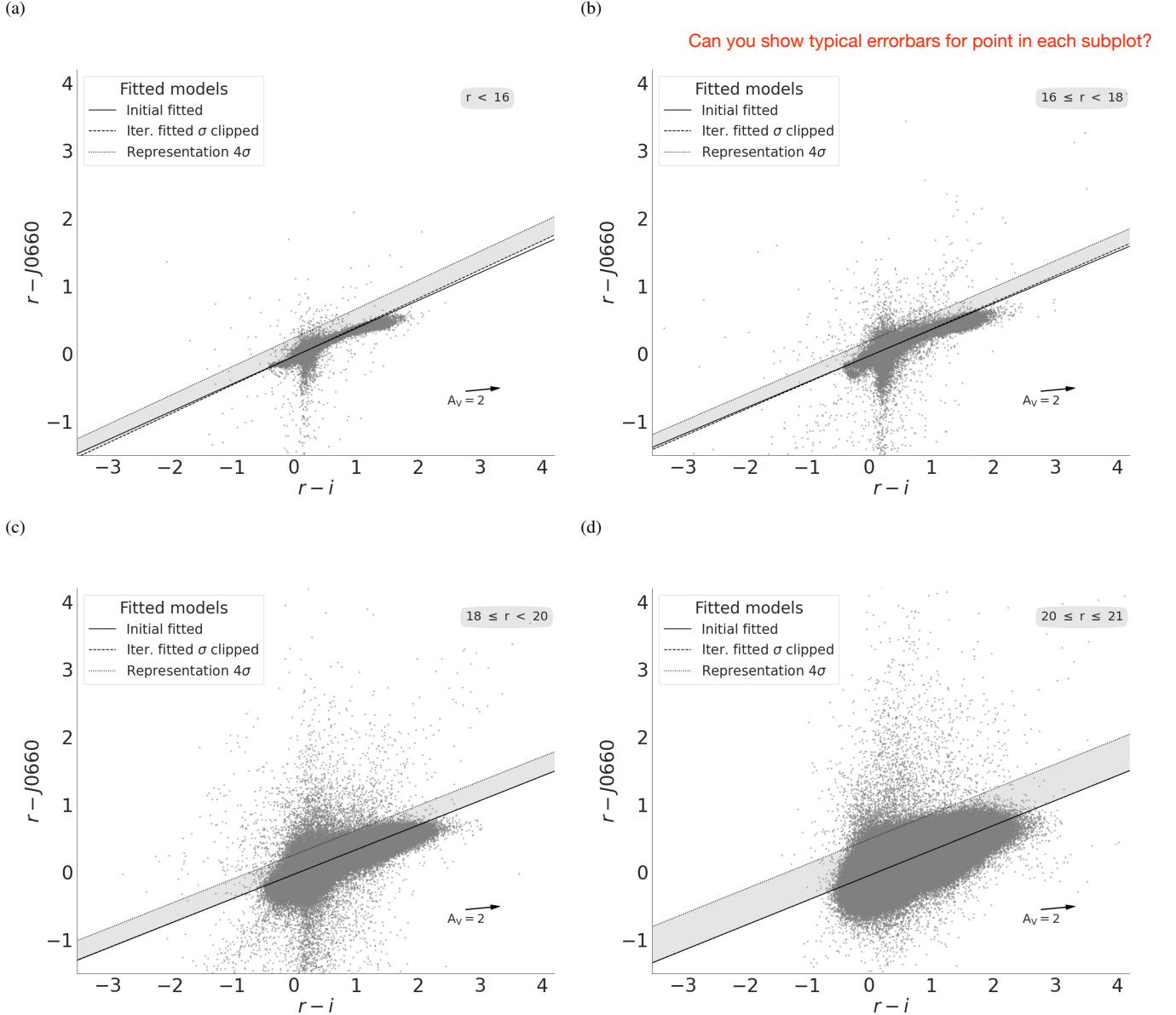
For the split of the sample of  $H\alpha$  emitters into two classes, the blue and the red populations, we follow an unsupervised machine learning approach implementing two clustering techniques: hierarchical agglomerative clustering and hierarchical density-based cluster selection, both based on the  $(g - r)$  and  $(z - g)$  colors, whose results are mutually compared.

### 2.3.1 Hierarchical agglomerative clustering

*Hierarchical clustering* (HC) belongs to the family of clustering algorithms of which clusters are constructed by recursively grouping and splitting the sources. Being an unsupervised algorithm, HC does not require a training sample or pre-conceived hypotheses. Data elements are grouped based on patterns in a given space of parameters and on the levels of similarity at which the groupings change (Jain et al. 1999). In the end, HC returns a diagrammatic representation of the groups as a tree – a dendrogram that follows an hierarchical structure.

There are two types of hierarchical clustering: the *hierarchical agglomerative clustering* (HAC; the one used in this work), which follows “bottom-up” approach, and the *hierarchical divisive clustering* that follows “top-down” approach. The HAC consists of building a binary merge tree, starting from each data element stored at the

<sup>3</sup> <https://docs.astropy.org/en/stable/modeling/index.html>



**Figure 2.** An illustration of the selection criteria used to identify strong emission-line objects via colour-colour plots. The data shown here are all from the S-PLUS DR3. The data are split up into four magnitude bins, as shown in the four panels. Objects with H $\alpha$  excess should be located near the top of the colour-colour diagrams. The thin continuous lines illustrate the original linear fit to all the data (grey points). The dashed lines represent the final fits to the stellar locus of points which were obtained by applying an iterative  $\sigma$ -clipping technique to the initial fit. The actual cuts used to select H $\alpha$  emitters are shown by the dotted lines. These correspond to  $4\sigma$  above of the final fit. Objects selected as H $\alpha$  emitters must be located above the dotted line. Note that the position of these lines (selection criteria) shown in the figure are approximated, given that the actual selection criterion also considers the errors on each source.

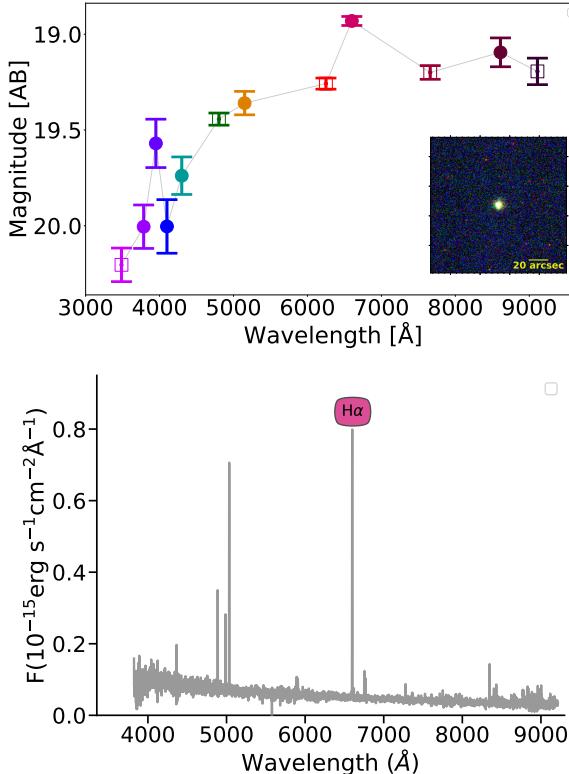
leaves (interpreted as individual clusters) and proceeds by merging two by two the “closest” sub-sets (stored at nodes) until the root – unique cluster – of the tree that contains all the elements of the data set is reached. The term “agglomerative” is used to point out that data elements are successively agglomerated into higher-levels. In each iteration, two “nearby” clusters are collapsed into a new, more populated group (Mann & Kaur 2013; Aggarwal 2015). Hence, each step reduces the number of clusters. The procedure may be summarized in three steps:

- (i) Initially, each data element represents one cluster, i.e. “leaves

of the tree”. This means that at the beginning, the total number of clusters/leaves is equal to the number of the elements in the data set.

- (ii) Through a looping process, the clusters are merged into new ones that are described by the maximum similarity between them.
- (iii) Finally, all the clusters belong to an unique cluster, “the root of the tree” structure.

On the other hand, the *hierarchical divisive clustering* algorithm follows a “top-down” approach. This means that the clustering starts from data element from only one cluster and then moves down recursively in the hierarchy to smaller groups. In simple words, hierarchical (agglomerative and divisive) clustering algorithms intend



**Figure 3.** Top panel: S-spectrum of a random emitting object found by the method explained in Section 2.2. Open squares represent the SDSS-like broadband filters. From left to right:  $u$ ,  $g$ ,  $r$ ,  $i$  and  $z$  magnitudes. Circles represent the narrow-band filters, which from left to right correspond to J0378, J0395, J0410, J0515, J0660 and J0861. The inset figure is the coloured image of the object which was produced by combining all twelve bands. Bottom panel: SDSS spectrum of the object - the H $\alpha$  line is marked.

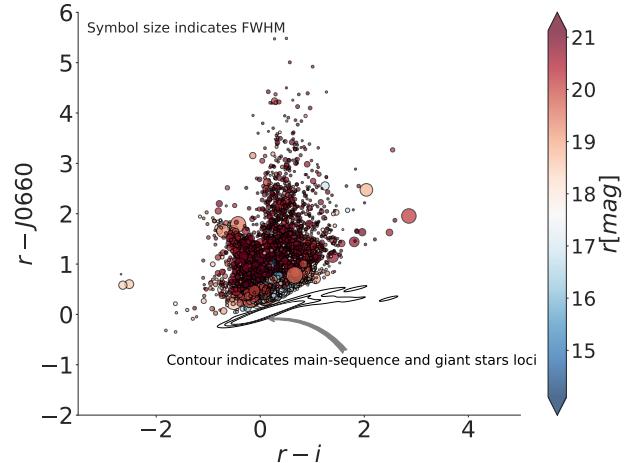
to gather similar objects into groups called clusters in the space of parameters which is investigated.

### 2.3.2 Hierarchical density-based cluster selection

Hierarchical density-based cluster selection (hereafter HDBSCAN; Campello et al. 2013) is another unsupervised machine learning algorithm that relies on clustering. It is based on a slightly modified version of density-based spatial clustering of applications (DBSCAN; Ester et al. 1996) which declares data points as noise. It assumes that clusters are characterized by “islands” of high density in the sea of the parameter space. HDBSCAN takes the DBSCAN concept forward by introducing a hierarchy to the clustering, with “persistent” clusters finally extracted from the hierarchical tree. The main advantage of HDBSCAN in comparison with its predecessor consists in the possibility of finding clusters of variable densities and different shapes. Following Malzer & Baum (2021) and Ntwaetsile & Geach (2021) it works as follows:

(i) HDBSCAN defines the “core” distance for a data point  $x$ ,  $\text{core}_k(x)$ , as the distance of an object to its  $k$ th nearest neighbour. This mean that lower values of  $\text{core}_k(x)$  represent higher densities and vice-versa.

(ii) The “mutual readability distance” between two points  $a$  and  $b$  is defined as  $d_m(a, b) = \min\{\text{core}_k(a), \text{core}_k(b), d(a, b)\}$ , where  $d(a, b)$  is the distance between  $a$  and  $b$  according, for instance,



**Figure 4.** Colour-colour diagram with all the emission-line objects selected from the S-PLUS DR3. The size of the symbols represent the measured FWHM assuming a Gaussian core (for more detail see Almeida-Fernandes et al. 2022). Coloured bar indicates the magnitude values of the  $r$ -band. The contours represent the S-PLUS synthetic photometry of main-sequence and giant stars loci from the library of stellar spectral energy distributions of Pickles (1998).

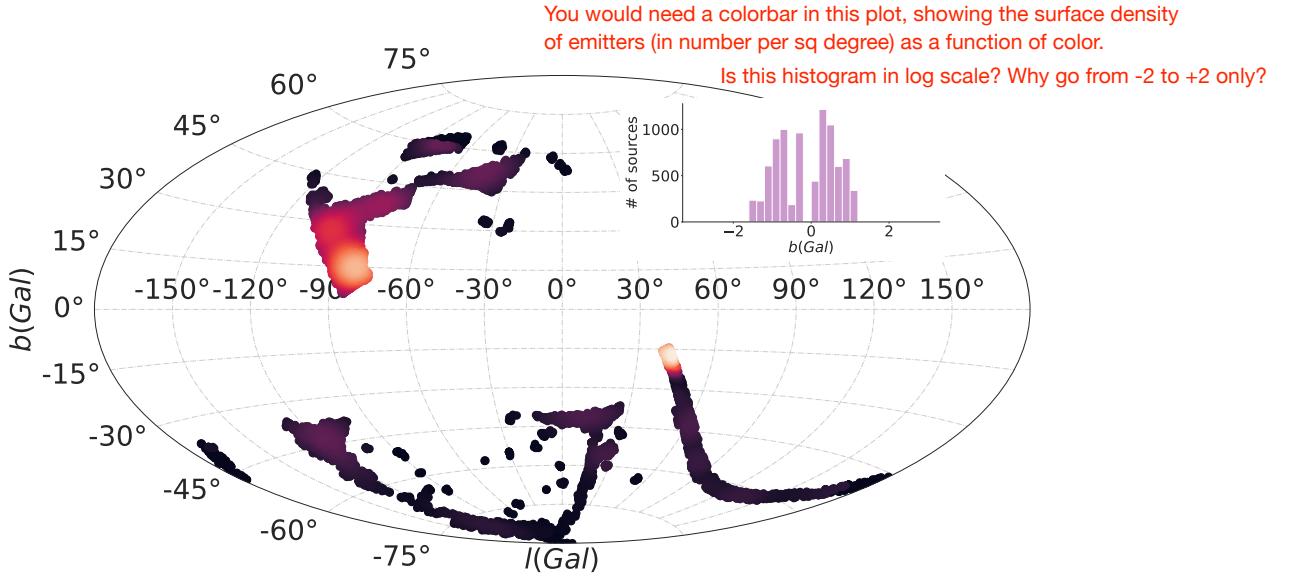
Euclidean metric. The mutual readability distance allows data points in dense regions to stay close together and those that are in less dense regions to move away.

(iii) The mutual readability plot is used to construct the minimum spanning tree, and sorting its edges by the mutual readability distance resulting in a hierarchical tree structure. The hierarchy of connected components is defined by sorting the edges of the tree by distance in reverse order, describing a dendrogram (the diagram explained in 2.3.1). This is the structure from which the cluster will be identified.

(iv) HDBSCAN allows extracting clusters of variable density, effectively, by cutting the dendrogram at different levels of grouping.

(v) The cluster tree is condensed into a simpler structure (see, for instance, Figure A1 of Appendix A). Considering the single main trunk which contains all the data points, the tree splits into branches. A condensed cluster hierarchy can be described by considering the number of points that are kept in each branch as it splits. It is important to mention that there is a key parameter called minimum cluster size. If a given branch splits into two, with one branch containing fewer points than the minimum cluster size, the larger branch “persists” and the smaller split branch “falls out” of the cluster. If a branch splits into two with both branches exceeding the minimum cluster size, both new branches are conserved.

(vi) The clusters are extracted on the notion of persistence in the hierarchy. The parameter  $\lambda = d_m^{-1}$  is defined, and each cluster has a  $\lambda_{\text{birth}}$  (the point at which the cluster split off) and  $\lambda_{\text{death}}$  (the point when the cluster split into other clusters). In each cluster, we have  $\lambda_p$  describing when each point fell out of the cluster (or was split off into a new cluster), so that  $\lambda_{\text{birth}} \leq \lambda_p \leq \lambda_{\text{death}}$ . Cluster stability  $S$  is defined as the sum of  $\lambda_p - \lambda_{\text{birth}}$  for all points in the cluster. To extract clusters, the following procedure is implemented. First, each leave constitutes a cluster. Then, moving through the hierarchy, it is considered the stability of a parent cluster  $S_p$  and its  $n$  descendants  $S_d^{0,1,2,\dots,n}$ : if  $S_p > \sum_{i=0}^n S_d^i$ , we unselect all the descendants; otherwise, the cluster stability is set as  $S_p = \sum_{i=0}^n S_d^i$ .



**Figure 5.** Distribution of the emission-line objects in galactic longitude and latitude coordinates. The inset figure represents the distribution of the objects in galactic latitude.

At the root node, we have our set of the selected clusters. Any data point in the sample that does not fall into one of the selected clusters is defined as noise.

(vii) By adopting the soft clustering or the fuzzy clustering technique it is possible to mitigate the need to establish or define the cluster membership limit. In fact, each source has a finite probability of belonging to every selected cluster. In this approach, all points (including noises) are not assigned to a cluster label, but are instead assigned to a vector of probabilities whose length is equal to the number of clusters found. Such an approach may solve the problem due to noise classification.

The HDBSCAN algorithm starts off much the same as DBSCAN: transforming the space according to density, exactly as DBSCAN does, and perform single linkage clustering on the transformed space. Instead of taking an epsilon<sup>4</sup> value as a cut level for the dendrogram, a different approach is followed: the dendrogram is condensed by viewing splits that result in a small number of points splitting off as points “falling out of a cluster”. This results in a smaller tree with fewer clusters that “lose points”. That tree can then be used to select the most stable or persistent clusters. This process allows the tree to be cut at varying height, picking our varying density clusters based on cluster stability. The immediate advantage of this is that we can have varying density clusters; the second benefit is that we have eliminated the epsilon parameter as we no longer need it to choose a cut of the dendrogram. Instead we have a new parameter `min_cluster_size` which is used to determine whether points are “falling out of a cluster” or splitting to form two new clusters.

Over the last few years, HDBSCAN has been used for different tasks in Astronomy. HDBSCAN was used to identify IR bubbles from Spitzer images (Jayasinghe et al. 2019). Webb et al. (2020) implemented HDBSCAN for discovering transients. Recently, Ntwaetsile & Geach

(2021) employed HDBSCAN to group radio sources into a sequence of morphological classes, illustrating a simple methodology to classify and label new, unseen galaxies in large samples. This approach was also implemented to identify stellar groups in Canis Major OB1 (Santos-Silva et al. 2021).

#### 2.4 Splitting the H $\alpha$ emitters into blue and red populations

For the selection of the blue and red populations in the sample of H $\alpha$  emitters, we first looked for the best colour-colour diagram by using the S-PLUS synthetic photometry of several classes of emission line objects. Such diagram, the ( $g - r$ ) versus ( $z - g$ ) color-color diagram is displayed in Fig. 6. SySt and YSOs span a wide range on ( $z - g$ ), from -0.5 to 6.0. On the other hand, the PNe, H II regions, CVs, QSOs, and emission line galaxies are located on the lower-right region of the diagram. The dashed line in Fig. 6 highlights the blue and red zones.

Fig. 7 displays the ( $g - r$ ) versus ( $z - g$ ) diagram from the list of H $\alpha$  emitters in S-PLUS. Obviously only such emitters with detections in the  $g$  and  $z$  filters are considered for this colour classification by making a cut in the magnitude errors at 0.2, totalizing 7086 objects. A bi-modal distribution is found for both colour indices (see inset plots of the Fig. 7). The two peaks on the ( $g - r$ ) and ( $z - g$ ) distributions have immediate correspondence with the blue and red zones pointed out from the synthetic diagram (Fig. 6). One can also see that the fraction of blue objects is considerable larger than that of the red ones.

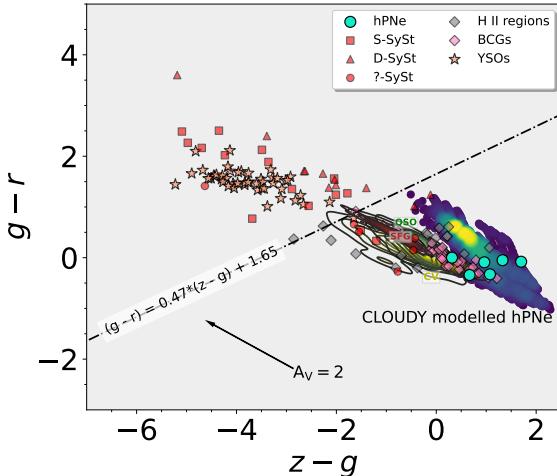
##### 2.4.1 HAC

The ideal way to choose the number of clusters is by displaying the **dendrogram diagram**. Firstly, the hierarchical cluster output dendrogram can be used to obtain the desired clustering. Secondly, the dendrogram allows a convenient way to establish the entity-relationship at all levels of granularity.

Fig. 8 illustrates the dendrogram based on the ( $g - r$ ) and ( $z - g$ )

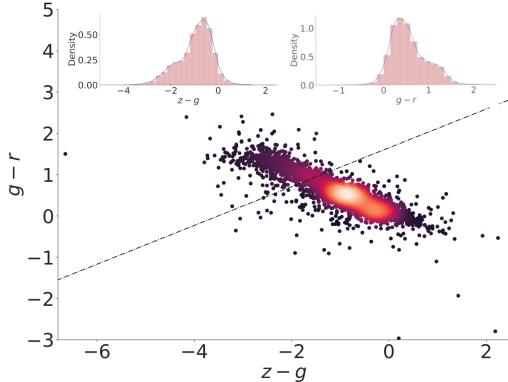
<sup>4</sup> Epsilon parameter in DBSCAN represents the maximum distance between two samples for one to be considered as in the neighborhood of the other. This is the most important DBSCAN parameter to choose appropriately for the data set and distance function.

How do you define these numbers? At this point they seem arbitrary.



**Figure 6.** The  $(g - r)$  versus  $(z - g)$  synthetic color-color diagram of several classes of emission lines objects. Included in the diagrams, there are families of CLOUDY modelled halo PNe spanning a range of properties (density map region). Cyan circles represent S-PLUS photometry from observed spectra of PNe. Grey diamonds represent H II regions in NGC 55. Red boxes and triangles display S- and D-type symbiotic stars, respectively. Red circles are SySt with no associated type. This group includes Galactic and external SySt from NGC 205 IC 10 and NGC 185. Yellow contours correspond to CVs from SDSS. Pink circles indicate blue compact galaxies (BCGs) from SDSS. Brown contours refer to SDSS star-forming galaxies (SDSS SFGs). SDSS QSOs at different redshift ranges are shown as green contours, and YSOs from Lupus and Sigma Orionis are represented by salmon stars. The diagonal dashed line represents a subjective criterion to separate the objects into two color types. The arrow indicates the reddening vector with  $A_V \approx 2$  mag.

circles or diamonds?



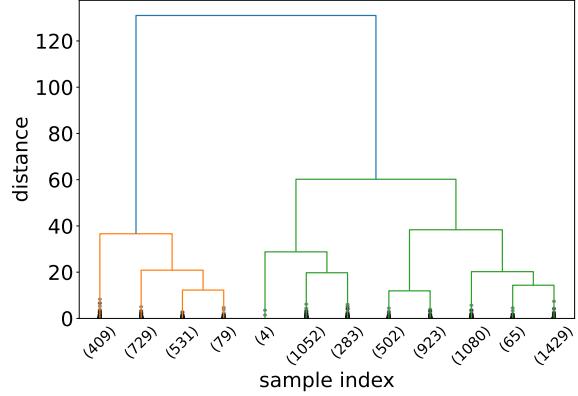
Why so many object in the intermediate region compared to Fig 6, do you know?

**Figure 7.** The  $(g - r)$  versus  $(z - g)$  color-color diagram with all the emission line objects selected in S-PLUS. The inset figures represent the  $(g - r)$  and  $(z - g)$  colour distributions.

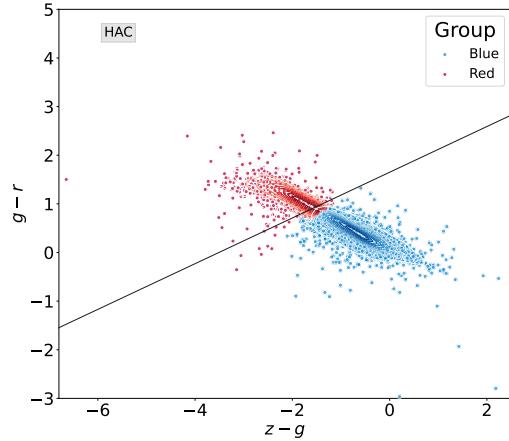
colours of H $\alpha$  emitters, and it highlights the order and distances of the groups in the hierarchical clustering, stopping at 12 nodes:

- The  $x$ -axis specifies the population in the nodes in a given level of grouping – that summed up correspond to the total number of elements under investigation.
- The  $y$ -axis represents the “distance”, which is a measurement of the closeness of the clusters or data points in different levels of clustering.

Reading the diagram from the top to the bottom, we see that all



**Figure 8.** Truncated dendrogram of complete-linkage hierarchical clustered based on  $(g - r)$  and  $(z - g)$  colours. The cluster sizes are exposed in the brackets for the 12 truncated clusters.



**Figure 9.** The  $(g - r)$  versus  $(z - g)$  color-color diagram with the two population found by implementing HAC algorithm. The blue and red symbols represent the sources with intense blue continuum and those with intense red continuum, respectively. The straight line is the same line of Figure 6.

Not a coincidence! Expected, given the bimodality and distributions shown in Figs 6/7

systems are divided after the very first level from the top already into (only) two groups: coincidentally, they correspond to the red and blue populations of H $\alpha$  emitters presented in Fig. 7 as it is showed in Fig. 9. From that point on, the groups were subdivided without evident distinction, and truncation was thus assumed when the 12-node level was reached. The truncation is an usual procedure when dealing with big data.

In this work, HAC was employed by using the library Scikit-learn<sup>5</sup> (Pedregosa et al. 2011). Then, the DENDROGRAM() function, which is included in the package scipy<sup>6</sup> was used in conjunctions wit the task *Dendrogram Truncation* to generate the (truncated) dendrogram. The following parameters must be taken into account when the algorithm is applied to data: n\_clusters,

<sup>5</sup> <https://scikit-learn.org/stable/>

<sup>6</sup> <https://www.scipy.org/>

**Affinity**, and **Linkage**. `n_clusters` defines the number of clusters expected by the user. Given that our goal is to divide our sample into two groups, `n_clusters` is set to "2". **Affinity** determines the "metric" that compute the linkage. We have found that a simplistic metric, the "Euclidean", is effective for our purpose. **Linkage** determines which distance to use between sets of observation. **Linkage** defines how the similarity between two clusters is calculated, by determining the distance between sets of observations as a function of the pairwise distances between elements. The algorithm merges the pairs of cluster that minimize this criterion. **Ward's method** minimizes the variance of the clusters that are being merged (Ward 1963). To implement this method, find the pair of clusters that leads to a minimum increase in total within-cluster variance after merging. Ward procedure uses the error sum of squares to measure this variance. The two clusters with the smallest error sum of squares will eventually form a new cluster.

At this point, our list of H $\alpha$  emitters is divided into two populations based on their continuum, with the blue population (5,338 objects) being larger than the red one (1,748).

#### 2.4.2 HDBSCAN

For the sake of comparison with the results from HAC, we also used HDBSCAN to distinguish the blue and the red sources. The main difference between these two algorithms is that HDBSCAN is more conservative in the sense that several data points are classified as noise. For this task, the Python implementation of HDBSCAN<sup>7</sup> (McInnes et al. 2017) was adopted.

Similarly to HAC, there are some key parameters that should be considered when the algorithm is applied. Regarding the metric, the "Euclidian" one is assumed. The two most critical parameters are the "minimum cluster size" and "minimum number of samples". The former refers to the smallest size of a group that it is considered as a cluster. The value of "80" has been adopted for the "minimum cluster size". The "minimum number of samples" provides a measure of how conservative our clustering method will be, expressed as the fraction of data classified as noise, and the value of "40" was adopted. With this model configuration, two clusters were identified. Several small clusters are found when the minimum number of samples values becomes smaller than 40.

Left panel of Fig. 10 shows the two clusters found with HDBSCAN. One cluster contains 192 red sources and the other one 3,825 blue sources. The number of objects classified as noise is 3,069. This result is fully consistent with those obtained from HAC. The two main clusters obtained with HDBSCAN are located in the same region in the ( $g - r$ ) versus ( $z - g$ ) diagram as those groups found based on the HAC. About 94% of the blue sources selected by HDBSCAN are in the list of blue objects identified by HAC. All the red sources selected by HDBSCAN were also classified by HAC as red objects. In fact, by applying the `condensed_tree_` to the data colors two clusters are selected. The `condensed_tree_` attribute is the equivalent dendrogram plot for HDBSCAN which displays the cluster tree mentioned in the section 2.3.2. (see Appendix A for more details about `condensed_tree_` attribute).

#### 2.4.3 Soft clustering for HDBSCAN

The main disadvantage of HDBSCAN is that several sources are labelled as "noise", so that they are not assigned to any cluster. As mentioned

<sup>7</sup> <https://hdbSCAN.readthedocs.io/en/latest/>

earlier, this comes from the conservative nature of HDBSCAN and the fact that these data points (data noise) are located far away of the clusters' cores. An alternative way to avoid outliers (data noises) classifications is the implementation of the "soft clustering" (see section 2.3.2). Soft clustering from HDBSCAN<sup>8</sup> was used to assign every object to a cluster that they most likely belong to. According to this approach, data points are not assigned in a deterministic way to a cluster but to a vector of probabilities as a measure of belonging to different clusters: the probability value at the  $i$ th entry of the vector is the probability that a data point is a member of the  $i$ th cluster. We can, then, simply assign cluster labels for every data point by taking the most likely cluster it belongs to, using probability thresholds. Therefore, soft clustering for HDBSCAN is achieved through an outlier score modification to consider how distant an outlier is from each cluster, which is based on the Global-Local Outlier Score from Hierarchies (GLOSH) algorithm (Campello et al. 2015). This is combined with a measure of distance from a given cluster to estimate the probability that a given data point belongs to any of the fixed groups drawn from the condensed tree.

The right panel of Fig. 10 shows which cluster the data points classified as the noise by HDBSCAN belong to. Blue and red points indicate those with the highest probability of being in the blue and red groups, respectively. This procedure fills out the clusters nicely. There were many noise points that most likely belong to the expected clusters in very good agreement with the results obtained from HAC. Indeed, our separation of the H $\alpha$  emitters into blue and red sources has been improved. Instead of forcing the algorithm to make a decision to which group a data point belongs to, just HAC does, we have quantified the likelihood of a given observation to belong to any of the two clusters found in our data set (see, for instance, the two last columns of Table B1).

## 3 RESULTS AND DISCUSSION

Our strategy that is focused on the identification of H $\alpha$  emitters in the S-PLUS footprint, exploiting the unique filter system of the survey returned 8,446 objects with excess in the J0660 band. The fractional contribution of different classes of H $\alpha$  emitters to the overall sample was evaluated by cross-matching the objects' list with the SIMBAD database<sup>9</sup>. Optical spectra available in the SDSS DR16 (Ahumada et al. 2020) and in the Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST; Wu et al. 2011) were also explored. In all cases, we assumed the angular distance on the sky-plane between sources considering positive matches those of mutually closest sources to each other within a given limit ( $d_{max,proj}$ ).

### 3.1 Matches with SIMBAD sources

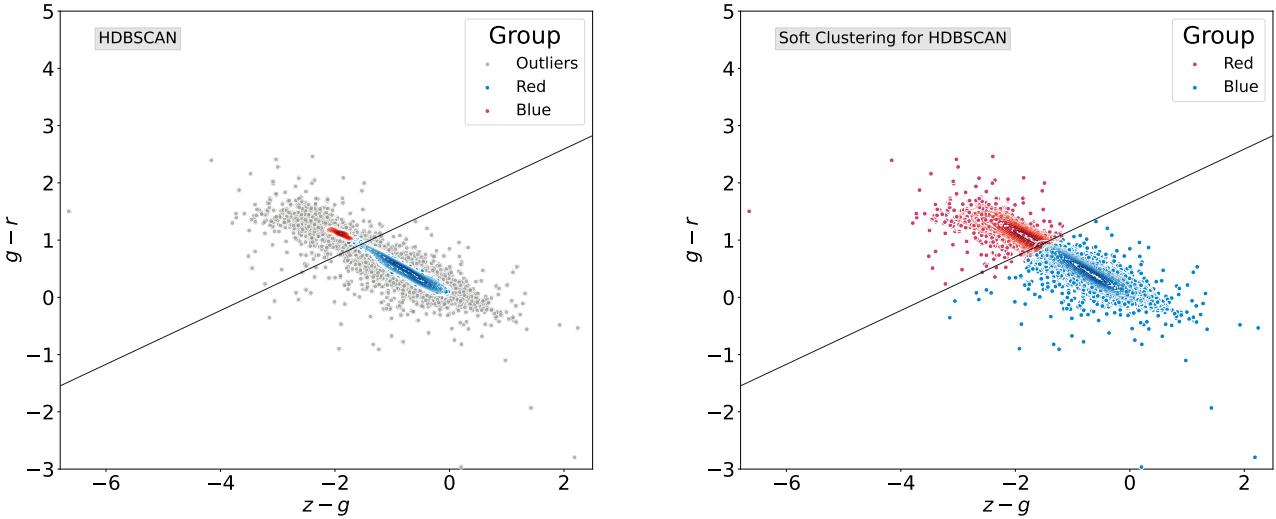
We found 1,042 positive matches between our catalog of H $\alpha$  emitters and SIMBAD database considering a radius of  $d_{max,proj} = 2$  arcsec. The results are described below and are listed in Table 1.

#### 3.1.1 Ionized nebulae

As it was mentioned, several classes of objects with diffuse appearance and/or nebular lines in our Galaxy and in nearby galaxies are

<sup>8</sup> [https://hdbSCAN.readthedocs.io/en/latest/soft\\_clustering\\_explanation.html](https://hdbSCAN.readthedocs.io/en/latest/soft_clustering_explanation.html)

<sup>9</sup> <http://simbad.u-strasbg.fr/simbad/>



**Figure 10.** *Left panel:* as Figure 9 but with the results after to apply HDBSCAN to the sample of emission-line sources. Blue and red symbols correspond to blue and red sources, respectively, with gray symbols representing sources classified as noise by HDBSCAN. The straight line is the same line as Figures 6 and 9. *Right panel:* results after apply a soft clustering to the HDBSCAN results.

**Table 1.** A summary of the results obtained of the positional cross-match between between the S-PLUS list of emission line objects and the SIMBAD database. We used a search radius of 2 arcsec. SIMBAD categories of objects are listed in the first column. The numbers of objects of each SIMBAD category are exposed in the third column.

Main type	Associated SIMBAD types	Number of S-PLUS objects with SIMBAD match
Nebulae	HII, PN, SN, Candidate_SN*, Nova	32
Stellar binary system	CataclyV*, Candidate_CV*, EB*, HMXB	51
Star	star, WD*, Candidate_WD*, Blue, BlueSG*, PM*, low-mass*, Cl*, GlCl	60
Variable star	RRLyr, Candidate_RRLyr, V*, PulsV*	23
Galaxy	EmG, HII_G, StarburstG, BlueCompG, IG, PartofG, GinPair, GinGroup, GinCl, LSB_G, BCIG, Galaxy	577
QSO	QSO, QSO_Candidate	209
AGN	AGN, AGN_Candidate, Seyfert_1, Seyfert_2, BLLac, RadioG	65
Other type	EmObj, FIR, MIR, MolCld, UV, Transient, Radio, X, Possible_lensImage, Unknown	25
Total		1,042

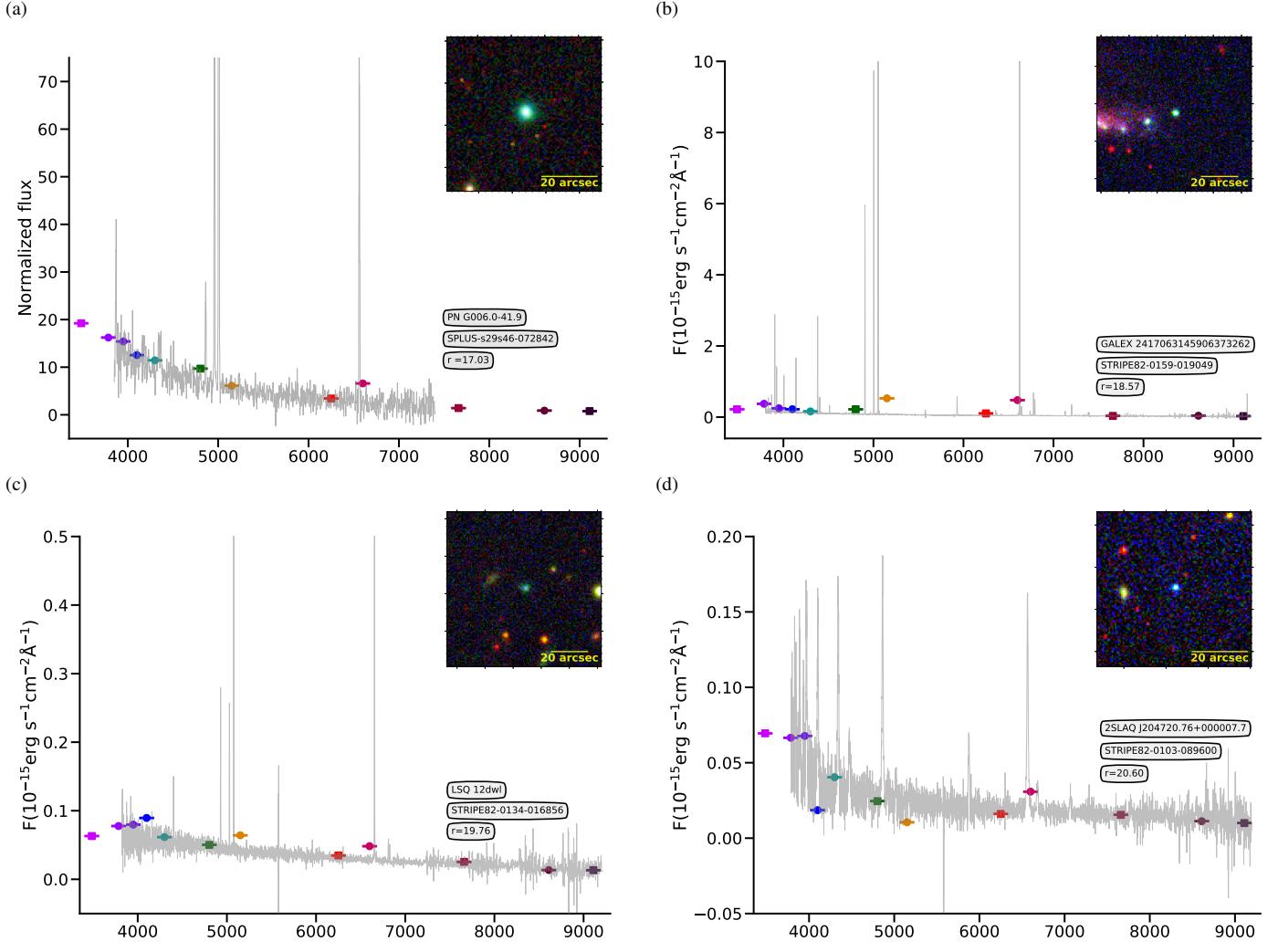
listed in our sample, most being H II regions, PNe, novae and SNe. H II regions are ionized by the UV light from early, massive stars (OB-type) and display an emission line spectrum. Unlike H II regions, planetary nebulae represent the final stages of low- and intermediate-mass stars from which the material has been previously ejected in the phases of AGB and post-AGB and is ionized by the high energetic radiation from a hot stellar remnant core. Supernovae or even novae also display emission line spectra and come from evolved stars through multiple channels. However, the energy-input mechanism is quite different in each case.

In our list of H $\alpha$  emitters only one PN is catalogued in SIMBAD. Its S-spectrum and SDSS spectrum are displayed on panel (a) of Fig 11. Emission lines like H $\alpha$  and [N II] are clearly perceptible in its spectra. This nebula belongs to the rare group of Galactic halo PNe. This rare group of PNe are of particular interest because they are characterized by low metallicity and present large velocities. 30

sources in our H $\alpha$  emitters list are catalogued as H II regions based on the SIMBAD database. The S-PLUS photometry and SDSS spectrum of the extragalactic H II region GALEX 2417063145906373262 is illustrated in panel (b). Panel c of Figure 11 also shows the SDSS and S-spectrum, as well as the coloured image, of an extragalactic SN, with evident emission lines.

### 3.1.2 Interacting binary systems

Following the classification available in Simbad, 20 known and 7 candidate CVs were found by our analysis. CVs are interacting binary systems of very short orbital period, in which a low-mass, early-type star fills its Roche lobe and transfers mass to a white dwarf companion (Patterson 1984). For the sake of illustration, Fig. 11 (panel d) shows the S-PLUS photometry overlapped to the SDSS spectrum of 2SLAQ J204720.76+000007.7, a CV, which we correctly classified as a blue



**Figure 11.** Summary of our selection results showing the spectrum (gray line en each panel) of different classes of emission line sources identified in our target list. A spectrum of a PN (a) from (Parker et al. 2016). The SDSS spectra of an external H II region (b), a super nova (c), a cataclysmic variable star (d). As in Figure 3, coloured square and circles symbols represent the S-PLUS photometry. All these objects show a significance excesses on the  $J0660$  filter in comparison with the broad-bands.

source. X-ray binary systems as well as eclipsing binaries are also listed on our catalog.

### 3.1.3 Stars

Vague Isn't this weird and unexpected? Comment? Several objects in our sample of  $H\alpha$  emitters have been categorized, by SIMBAD, as: normal stars, white dwarf (WD\*), white dwarf candidates (Candidate\_WD\*), blue stars, blue super-giant stars (BlueSG\*), high proper-motion stars (PM\*), variable stars of RR Lyr type and, low-mass star (low-mass\*;  $M < 1M_\odot$ ).

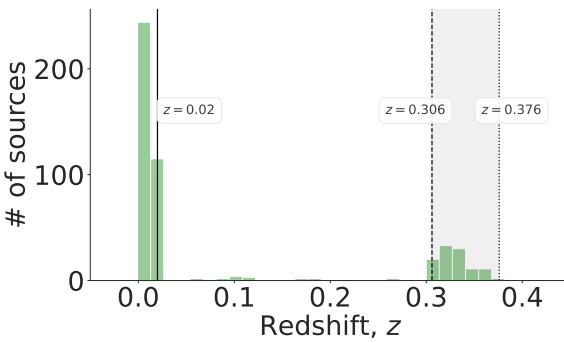
### 3.1.4 Galaxies

Galaxies are also included in our catalog: emission-line galaxies (EmG), blue compact galaxies (BlueCompG), H II galaxies (HII G), starburst galaxies (StarburstG), galaxies in clusters (GinCl) and in groups (GinGroup), low surface brightness galaxies (LSB G), radio-galaxies (RadioG), interacting galaxies (IG), part of a Galaxy (PartofG), Seyfert types- 1 and 2, and other type of AGNs. Since, we

are focusing on the  $H\alpha$  emission line, the emission line galaxies in the local Universe ( $z \sim 0.02$ ) are of particular interest because their  $H\alpha$  line still falls into the wavelength range covered by the  $J0660$  S-PLUS filter.

Fig. 12 shows the redshift distribution of the galaxies in our sample that have SIMBAD correspondents. About 60% of the galaxies have small red-shift values ( $z < 0.02$ ) probing that the emission detected in the  $J0660$  filters is associated to the real  $H\alpha \lambda 6563$  emission line. On the other hand, Fig. 12 also shows that there is an increment of galaxies with redshift between  $\sim 0.31$  and  $\sim 0.38$ . This particular population of  $H\alpha$  emitters is represented by AGN, Seyfert 1 and Seyfert 2 galaxies. In fact, at the accumulative red-shift range,  $0.306 < z < 0.376$ , represented by the filled area in the figure, the  $H\beta$  and  $[O III] 4959, 5007\text{\AA}$  emission lines are the ones detected by the  $J0660$  filter.

It should be noted here that the majority of the sources with a classification of “part of a Galaxy” (PartofG) in SIMBAD actually designates galaxies with Wolf-Rayet (WR) signature in the low redshift Universe also named “WR galaxy” (Osterbrock & Cohen 1982). The presence of WR stars in galaxies is perceptible from their spec-



**Figure 12.** Histogram of red-shift for the galaxies with SIMBAD coincidence. The black vertical continuous line indicates the threshold value ( $z \sim 0.02$ ) on which the H $\alpha$  emission-line is detected in  $J0660$  filter. The black vertical dashed and dotted lines represent the accumulative red-shift range where the H $\beta$  and [O III] 4959, 5007 Å emission lines are detected on the  $J0660$ -band.

tra, as suggested by strong emission lines such as H $\alpha$  and [N II]. However, these spectral features can also be indicative of extragalactic H II regions which are typically present in the outskirt of spiral galaxies. Panel (e) of Fig. 13 displays the S-PLUS and SDSS spectra of the galaxy with Wolf-Rayet signature [BKD2008] WR 14.

Some H $\alpha$  emitters are classified in SIMBAD as “galaxy”. These galaxies can in fact be spiral galaxies in which star formation is still active. Panel *c* of Fig. 13 shows the S-PLUS photometry and the SDSS spectrum of a star-forming galaxy. The spectrum clearly exhibits strong emission lines. Note that almost all these galaxies are classified as blue objects by HAC and HDBSCAN. However, most of the Seyfert type-2 galaxies and radio-galaxies as well as a handful of other galaxies are found to fill up the population of red sources.

Not a false positive, right? Just a different line?

### 3.1.5 QSOs: false positive detection of H $\alpha$ emission

Following the classification available in the literature, about 3% of blue H $\alpha$  emitters sources in our sample are found to be QSOs. We have to point out here that the excess in the  $J0660$  filter for QSOs is attributed to redshifted lines that fall in the wavelength range covered by that filter depending on the redshift of the QSOs – e.g., H $\beta$ , C IV 1550 Å, C III] 1909 Å, and Mg II 2798 Å (see Gutiérrez-Soto et al. 2020 and Nakazono et al. 2021). QSO 2SLAQ J220529.34-003110.6, shown in Panel *f* of Figure 13, is an example of a QSO at redshift  $\sim 2.45$ , for which the C III] line falls at the range covered by the  $J0660$  filter.

### 3.1.6 Other type of objects

### Other classifications

As it can be seen in Table 1, our sample also gathers a variety of objects without any previous classification. They may also be clusters of stars, far- and mid-infrared sources, molecular clouds, UV sources, among others, indicating the richness of the sample in nature and in physical properties.

## 3.2 SDSS and LAMOST: a spectroscopic validation

Finally, we also cross-matched out a sample of H $\alpha$  emitters in the S-PLUS with the SDSS DR16 (Ahumada et al. 2020). For doing this, we adopted a 2 arcsec as the cross-matching radius. In the case of the cross-match with LAMOST (Wu et al. 2011), the same radius was considered, and we ended up with 479 sources belonging to both

catalogues. **Approximately** And about 96% of them display strong emission lines spectra.

Most of the H $\alpha$  emitters with available spectroscopic information correspond to H II regions, CVs, SNe, emission-line galaxies (blue compact galaxies, H II galaxies, star-forming galaxies, among others), AGN (Seyfert 1 and 2), and QSOs. However, we emphasize that more detailed analysis is necessary to check which other types of objects are included in these samples of spectra – what is not in the scope of this paper. Also, it is worth noticing that part of the objects does not have a conclusive classification.

The spectra from both SDSS and LAMOST projects provide a good validation to our approach, clearly showing that the methodology is actually effective for selecting sources with the H $\alpha$  emission line in emission.

## 3.3 Magnitudes and color distributions

In Fig. 14, we demonstrate the distribution of the blue and red population of S-PLUS H $\alpha$  emitters in term of their  $r$  magnitude and their  $r - i$  and  $r - J0660$  colours.

Both, blue and red sources can be as bright as 16 mag in the  $r$  filter, while they show a peak at  $\sim 20$  mag. The fraction of blue sources in the  $16.0 \leq r \leq 19.0$  magnitude range is considerable higher in comparison with the red group. Therefore, the blue sample tends to be brighter than the red one in the  $r$ -band. Could this be observational bias?

Middle panel of Fig. 14 displays the  $(r - i)$  distribution of the blue and red H $\alpha$  sources which peak at distinct values of -0.9 and 0.5, respectively. This result is consistent with that obtained from Wevers et al. (2017) who also used the  $(r - i)$  colour index to select blue outliers from the Galactic Bulge Survey (GBS; Jonker et al. 2011).

Finally, the bottom panel of Fig. 14 shows the  $(r - J0660)$  colour index distribution of the blue and red objects with peak at 0.5 and 0.7, respectively. This result implies a strong H $\alpha$  emission in the red sources compared to the blue ones. Also bias?

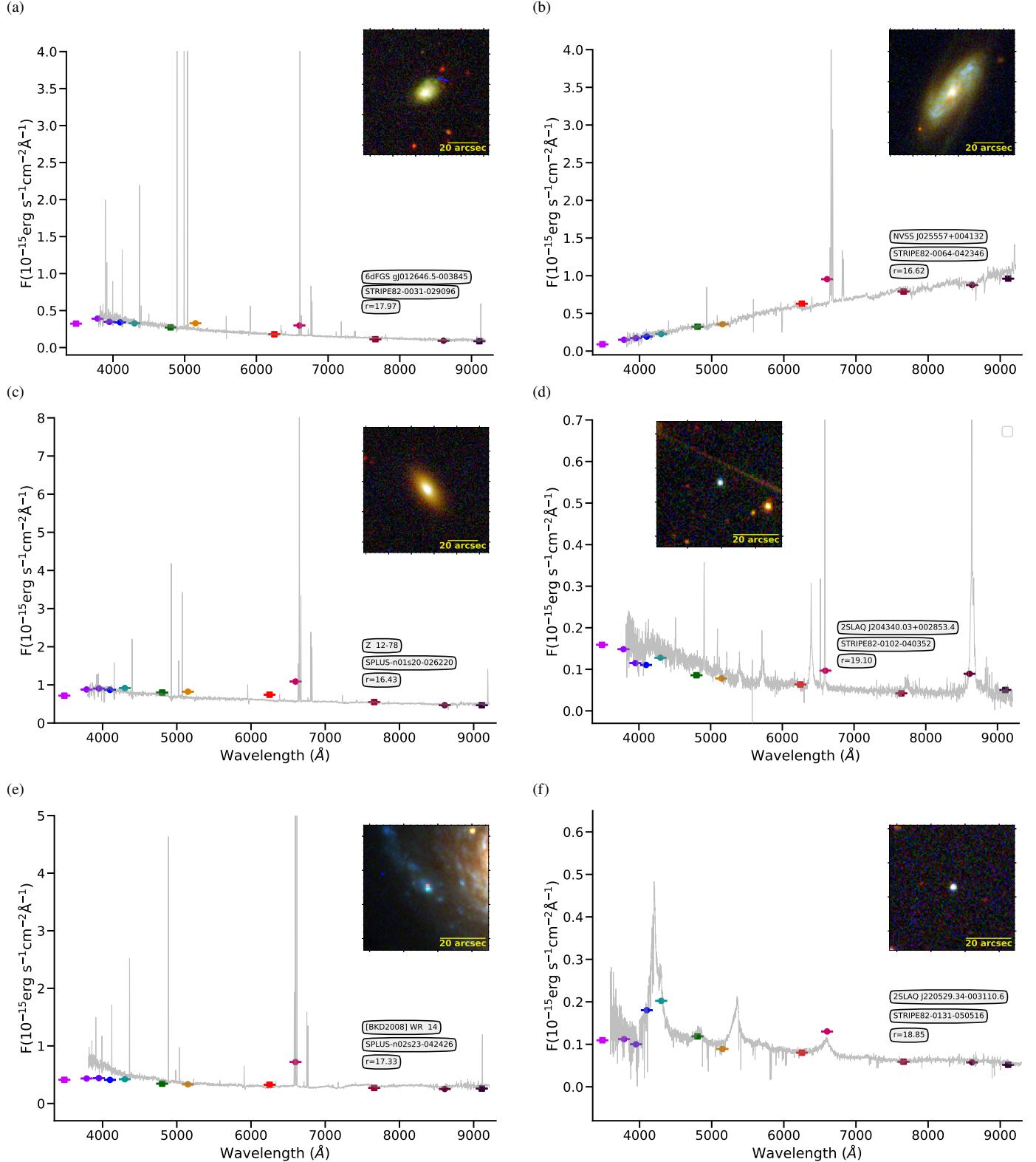
## 4 CONCLUSIONS

Here we exploited the capability of the S-PLUS project (Mendes de Oliveira et al. 2019) to survey H $\alpha$  emitters in the Southern Sky following a three-steps approach: identify H $\alpha$  emitters, distinguish the blue and red populations as a first diagnostic about the nature of the sources, and validate the results through spectroscopic databases.

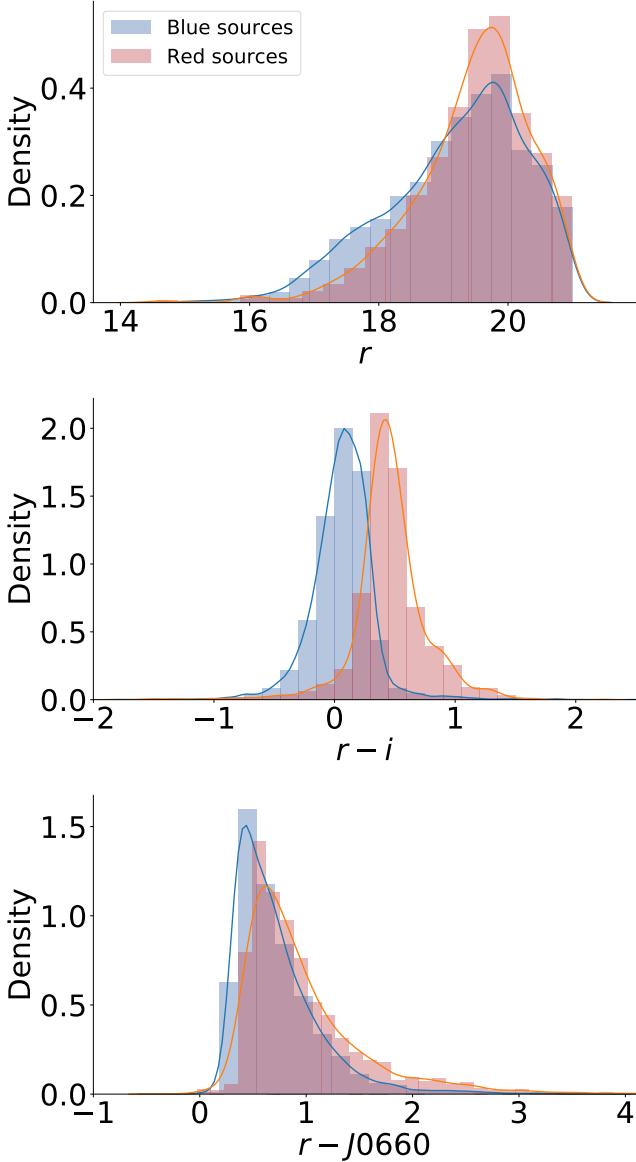
The H $\alpha$  emitters were identified by employing the  $J0660$  narrow-filter and  $r$  and  $i$  broad-filters available in the S-PLUS project. The  $(r - J0660)$  versus  $(r - i)$  colour-colour diagram was used to define the loci of the main-sequence and giant stars and disentangle objects in the local Universe with an H $\alpha$ -excess ( $r - J0660 > 0$ ) (see Fig. 4). 8,446 sources matched this criterion, with 407 of them claimed in the literature as QSOs and non-local galaxies, and therefore being false positive identifications of H $\alpha$  (see sections 3.1.4 and 3.1.5).

The  $(g - r)$  and  $(z - g)$  colour distributions of the H $\alpha$  emitters were found to be bimodal, indicating the presence of two distinct populations of bluer and redder sources with a narrow overlapping zone (Fig. 7). Two algorithms of unsupervised machine learning classification were used to distinguish the two populations: the HAC and the HDBSCAN clustering algorithms. Both algorithms ended up to very similar clusters, on the  $(g - r)$  and  $(z - g)$  colour indices space.

Given that HDBSCAN is considered as a conservative algorithm, many objects were labeled as noise data points while they did not by the HAC algorithm (Section 2.4.3). To overcome this problem, a so-called “soft” clustering approach for HDBSCAN was employed and the



**Figure 13.** SDSS spectra of a H II galaxy with  $z = 0.006$  (a), a radio galaxy with  $z = 0.014$  (b), a star-forming galaxy with  $z = 0.013$  (c). For this object, the H $\alpha$  line is responsible for the J0660 magnitude. And a Seyfert 1 with  $z = 0.317$  (d). For this last object, the excess on the J0660 is due to the [O III] 4959, 5007 Å emission lines. a WR in a galaxy (e) and a QSO (f) with red-shift of  $\sim 2.45$ . As in Figure 11, coloured symbols indicate the S-PLUS photometry.



**Figure 14.** Distribution of  $r$  magnitude (top panel),  $(r - i)$  color (middle panel), and  $(r - J0660)$  color (bottom panel) for the blue and red sources of the sample of  $H\alpha$  emitters. The histogram heights show density normalization scales. The smooth curves represent a Kernel density estimation for both samples.

probabilities of each data point to belong to the “blue” and “red” subgroup were computed. The results from the HAC and HDBSCAN algorithms are mutually consistent. We, therefore, reckon that the  $(g - r)$  and  $(z - g)$  colours are ideal for separating objects into the bluer and redder populations and attribute their colour to the nature of the sources. In particular, the bluer objects were found to be mainly CVs, PNe, H II regions, dwarf compact galaxies, and QSOs, among others, while the redder sources are early type galaxies with emission lines (for instances, radio-galaxies and Seyfert 2 galaxies), probably young/active late-type stars or even symbiotic stars (in fact, evolved binary systems hosting a red giant star).

Finally, we also cross-matched our catalog of  $H\alpha$  emitters with spectroscopic databases (SDSS and LAMOST; see Section 3.2). This exercise demonstrated that at least 95% of the objects with available

spectroscopic information are genuine emission line sources, validating our approach to identify  $H\alpha$  emitters in the S-PLUS project. The spectroscopic sample of  $H\alpha$  emitters lists 240 sources of the local Universe (with  $z < 0.02$ ) indicating that the emission on the  $J0660$  filter corresponds to the  $H\alpha$  line, 239 sources with redshift larger than 0.02, indicating that they are very likely QSOs and non-local galaxies on which the excess of the  $J0660$  filter is due to C IV 1550 Å, C III] 1909 Å, and Mg II 2798 Å emission lines for the case of QSOs and H $\beta$  and [O III] 4959, 5007 Å emission lines for galaxies, those depending on their redshift.

As a practical result, here we make public a catalog from SPLUS/DR3 that can be explored by the community in the identification and investigation of sources in twelve photometric bands in a systematic and homogeneous way.

## ACKNOWLEDGEMENTS

LAG-S acknowledges funding for this work from FAPESP grants 2019/26412-0. RLO acknowledges financial support from the Brazilian institutions CNPq (PQ-312705/2020-4) and FAPESP (#2020/00457-4). DGR acknowledges the CNPq (428330/2018-5; 313016/2020-8) and FAPERJ (269312) grants.

The S-PLUS project, including the T80-South robotic telescope and the S-PLUS scientific survey, was founded as a partnership between the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Observatório Nacional (ON), the Federal University of Sergipe (UFS), and the Federal University of Santa Catarina (UFSC), with important financial and practical contributions from other collaborating institutes in Brazil, Chile (Universidad de La Serena), and Spain (Centro de Estudios de Física del Cosmos de Aragón, CEFCA). We further acknowledge financial support from the São Paulo Research Foundation (FAPESP), the Brazilian National Research Council (CNPq), the Coordination for the Improvement of Higher Education Personnel (CAPES), the Carlos Chagas Filho Rio de Janeiro State Research Foundation (FAPERJ), and the Brazilian Innovation Agency (FINEP).

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

Guoshoujing Telescope (the Large Sky Area Multi-Object Fiber Spectroscopic Telescope LAMOST) is a National Major Scientific Project built by the Chinese Academy of Sciences. Funding for the project has been provided by the National Development and Reform

Commission. LAMOST is operated and managed by the National Astronomical Observatories, Chinese Academy of Sciences.

Scientific software and databases used in this work include TOPCAT<sup>10</sup> (Taylor 2005), simbad and vizier from Strasbourg Astronomical Data Center (CDS)<sup>11</sup> and the following python packages: numpy, astropy, matplotlib, seaborn, scikit-learn.

## DATA AVAILABILITY

## REFERENCES

- Aggarwal C. C., 2015, Data Mining: The Textbook. Springer, Cham, doi:10.1007/978-3-319-14142-8
- Ahumada R., et al., 2020, *ApJS*, **249**, 3
- Akras S., Guzman-Ramirez L., Leal-Ferreira M. L., Ramos-Larios G., 2019a, *ApJS*, **240**, 21
- Akras S., Leal-Ferreira M. L., Guzman-Ramirez L., Ramos-Larios G., 2019b, *MNRAS*, **483**, 5077
- Akras S., Guzman-Ramirez L., Gonçalves D. R., 2019c, *MNRAS*, **488**, 3238
- Almeida-Fernandes F., et al., 2022, *MNRAS*, **511**, 4590
- Barentsen G., et al., 2014, *MNRAS*, **444**, 3230
- Benitez N., et al., 2014, arXiv e-prints, p. arXiv:1403.5237
- Bonoli S., et al., 2021, *A&A*, **653**, A31
- Campello R. J. G. B., Moulavi D., Sander J., 2013, in Pei J., Tseng V. S., Cao L., Motoda H., Xu G., eds, Advances in Knowledge Discovery and Data Mining. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 160–172
- Campello R., Moulavi D., Zimek A., Sander J., 2015, *ACM Transactions on Knowledge Discovery from Data*, **10**, 1
- Cenarro A. J., et al., 2019, *A&A*, **622**, A176
- Corradi R. L. M., Giannanco C., 2010, *A&A*, **520**, A99
- Corradi R. L. M., et al., 2008, *A&A*, **480**, 409
- Corradi R. L. M., Sabin L., Munari U., Cetrulo G., Englano A., Angeloni R., Greimel R., Mampaso A., 2011, *A&A*, **529**, A56
- Davies R. D., Elliott K. H., Meaburn J., 1976, *Mem. RAS*, **81**, 89
- Drew J. E., et al., 2005, *MNRAS*, **362**, 753
- Drew J. E., Greimel R., Irwin M. J., Sale S. E., 2008, *MNRAS*, **386**, 1761
- Drew J. E., et al., 2014, *MNRAS*, **440**, 2036
- Ester M., Kriegel H.-P., Sander J., Xu X., 1996, in Proc. of 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96). pp 226–231
- Frew D. J., 2008, PhD thesis, Department of Physics, Macquarie University, NSW 2109, Australia
- Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, *AJ*, **111**, 1748
- Gutiérrez-Soto L. A., et al., 2020, *A&A*, **633**, A123
- Jacoby G. H., et al., 2010, *Publ. Astron. Soc. Australia*, **27**, 156
- Jain A. K., Murty M. N., Flynn P. J., 1999, *ACM Comput. Surv.*, **31**, 264
- Jayasinghe T., et al., 2019, *MNRAS*, **488**, 1141
- Jonker P. G., et al., 2011, *ApJS*, **194**, 18
- Malzer C., Baum M., 2021, *Sensors*, **21**
- Mann A., Kaur N., 2013.
- Marín-Franch A., et al., 2012, in Navarro R., Cunningham C. R., Prieto E., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8450, Modern Technologies in Space- and Ground-based Telescopes and Instrumentation II. p. 84503S, doi:10.1117/12.925430
- McInnes L., Healy J., Astels S., 2017, *The Journal of Open Source Software*, **2**
- Mendes de Oliveira C., et al., 2019, *MNRAS*, **489**, 241
- Merc J., Gàlis R., Wolf M., 2019, *Eruptive Stars Information Letter*, **41**, 78
- Nakazono L., et al., 2021, *MNRAS*, **507**, 5847
- Ntwaetsile K., Geach J. E., 2021, *MNRAS*, **502**, 3417
- Oke J. B., Gunn J. E., 1983, *ApJ*, **266**, 713
- Osterbrock D. E., Cohen R. D., 1982, *ApJ*, **261**, 64
- Parker Q. A., et al., 2005, *MNRAS*, **362**, 689
- Parker Q. A., Bojičić I. S., Frew D. J., 2016, in Journal of Physics Conference Series. p. 032008 (arXiv:1603.07042), doi:10.1088/1742-6596/728/3/032008
- Patterson J., 1984, *ApJS*, **54**, 443
- Pedregosa F., et al., 2011, *Journal of Machine Learning Research*, **12**, 2825
- Pickles A. J., 1998, *PASP*, **110**, 863
- Sabin L., Zijlstra A. A., Wareing C., Corradi R. L. M., Mampaso A., Viironen K., Wright N. J., Parker Q. A., 2010, *Publ. Astron. Soc. Australia*, **27**, 166
- Santos-Silva T., et al., 2021, arXiv e-prints, p. arXiv:2108.06234
- Scaringi S., Groot P. J., Verbeek K., Greiss S., Knigge C., Körding E., 2013, *MNRAS*, **428**, 2207
- Taylor M. B., 2005, in Shopbell P., Britton M., Ebert R., eds, Astronomical Society of the Pacific Conference Series Vol. 347, Astronomical Data Analysis Software and Systems XIV. p. 29
- Viironen K., et al., 2009, *A&A*, **502**, 113
- Vink J. S., Drew J. E., Steeghs D., Wright N. J., Martin E. L., Gänsicke B. T., Greimel R., Drake J., 2008, *MNRAS*, **387**, 308
- Ward J. H., 1963, *Journal of the American Statistical Association*, **58**, 236
- Webb S., et al., 2020, *MNRAS*, **498**, 3077
- Wevers T., et al., 2017, *MNRAS*, **466**, 163
- Witham A. R., et al., 2006, *MNRAS*, **369**, 581
- Witham A. R., et al., 2007, *MNRAS*, **382**, 1158
- Witham A. R., Knigge C., Drew J. E., Greimel R., Steeghs D., Gänsicke B. T., Groot P. J., Mampaso A., 2008, *MNRAS*, **384**, 1277
- Wu Y., et al., 2011, *Research in Astronomy and Astrophysics*, **11**, 924

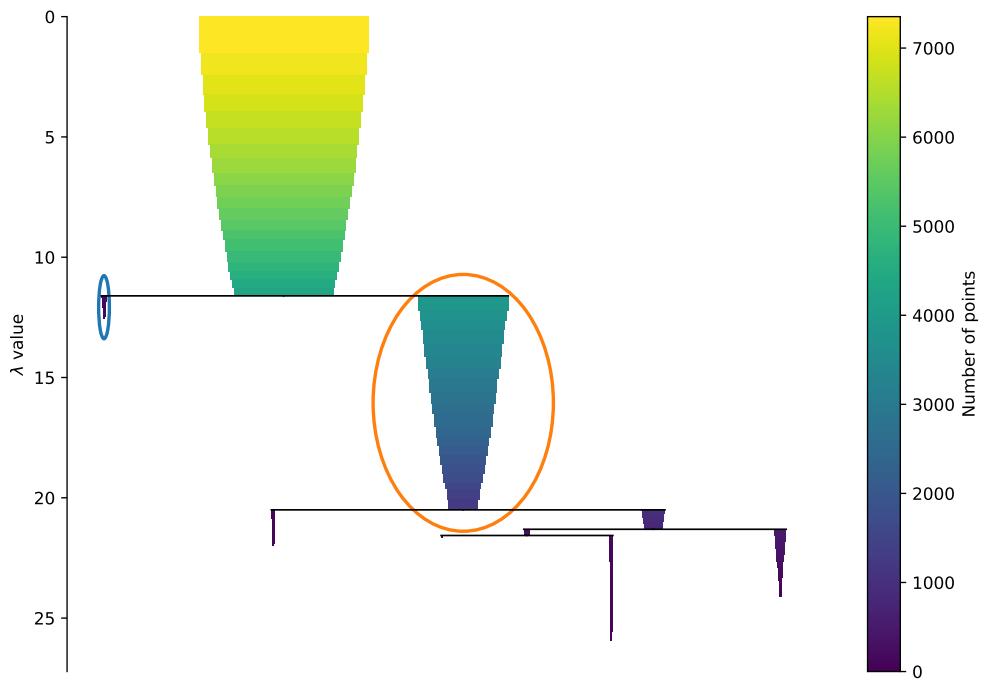
<sup>10</sup> <http://www.star.bristol.ac.uk/~mbt/topcat/>

<sup>11</sup> <https://cds.u-strasbg.fr/>

## APPENDIX A: CONDENSED TREES

The condensed Trees is a diagram for HDBSCAN that allows to see the cluster hierarchy as a dendrogram. It can be displayed via the `condensed_tree_` attribute of the `HDBSCAN` package. Figure A1 shows the condensed trees which was obtained by using the  $(r - g)$  and  $(g - z)$  colours as the the input parameters. It is possible to see that HDBSCAN has found two clusters in agreement with previous results. This means that they represent the blue and red sources.

## APPENDIX B: SIMBAD OBJECTS



**Figure A1.** The condensed Trees for our sample of  $\text{H}\alpha$  emitters. The width and color of each branch represent the number of points in the cluster at that level. The orange and blue ellipses represent the branches selected by the HDBSCAN algorithm.

Table B1: Objects from the SIMBAD data base. The first column presents the ID SIMBAD of the source in question. Right ascension and declination are shown in the second and third columns, respectively. The type is given in the fourth column. The red-shift, if exist in SIMBAD, is displayed in the fifth column. The colour-type classification performed with HAC algorithm is presented in the sixth column. The seventh and eighth columns show the probability estimated from HDBSCAN approach to being a blue and red source, respectively..

Id Object	RA	Dec	Type	Redshift	Group		P(Blue) HDBSCAN	P(Red) HDBSCAN
					HAC	HDBSCAN		
CTLGD 9478	00:01:59.25	-29:18:40.4	Star	-	Blue	0.99	0.01	
QSO B2359+005	00:02:30.71	00:49:59.2	QSO	1.354	Blue	0.46	0.05	
CTLGD 2037	00:05:08.77	-30:51:04.2	Star	-	Red	0.09	0.45	
SDSS J000637.99-003656.2	00:06:37.99	-00:36:56.2	QSO	4.435	-	-	-	
LBQS 0004+0036	00:07:10.00	00:53:29.1	QSO	0.316	Blue	0.37	0.08	
2MASS J00075351+0017597	00:07:53.49	00:17:59.7	low-mass*	0.000	Red	0.05	0.10	
SDSS J000809.34+004935.5	00:08:09.34	00:49:35.3		3.293	-	-	-	
2SLAQ J000918.74-003907.2	00:09:18.76	-00:39:07.0	Galaxy	-	Blue	0.72	0.08	
[VV2006] J001040.1-294428	00:10:40.08	-29:44:27.3	QSO	1.361	Blue	0.41	0.08	
CTLGD 7291	00:10:48.73	-29:47:28.8	Star	-	Red	0.16	0.24	
2QZ J001055.3-304423	00:10:55.37	-30:44:23.5	Galaxy	0.307	Blue	0.60	0.08	
[VV2006] J001228.8-310241	00:12:28.78	-31:02:40.0	QSO	1.360	Blue	0.30	0.05	
LBQS 0010+0035	00:13:27.32	00:52:32.2	Seyfert 1	0.363	Blue	0.46	0.05	
[GPM2009] J0014-0044 1	00:14:28.79	-00:44:43.8	EmG	0.014	Blue	0.09	0.04	
2SLAQ J001455.99+001903.5	00:14:55.99	00:19:03.7	Star	-	Blue	0.80	0.07	
2SLAQ J001526.52+001813.2	00:15:26.52	00:18:13.4	QSO	1.362	Blue	0.51	0.09	
[VV2006] J001535.5+005355	00:15:35.55	00:53:56.1	QSO	1.358	Blue	0.75	0.05	
SDSS J001628.25+010801.9	00:16:28.24	01:08:02.0	Galaxy	0.010	Blue	0.43	0.08	
[VV2006] J001641.9-312657	00:16:41.87	-31:26:56.6	QSO	0.360	Blue	0.46	0.06	
2SLAQ J001731.27-004859.3	00:17:31.26	-00:48:59.2	QSO	1.357	Blue	0.44	0.06	
LEDA 1156	00:17:39.97	00:30:22.5	StarburstG	0.017	Blue	0.93	0.01	
SDSS J001753.82+005057.6	00:17:53.82	00:50:57.7		1.358	-	-	-	
2SLAQ J001912.39+000319.6	00:19:12.39	00:03:19.8	QSO	1.372	Blue	0.47	0.05	
2SLAQ J001940.23-005435.9	00:19:40.24	-00:54:35.8	QSO	1.374	Blue	0.71	0.05	
[VV2006] J001950.1-004040	00:19:50.06	-00:40:40.7	QSO	4.340	-	-	-	
2SLAQ J002237.90+000519.0	00:22:37.90	00:05:19.2	QSO	1.373	Blue	0.55	0.03	
UM 240	00:25:07.40	00:18:45.2	EmObj	0.011	Blue	0.46	0.08	
2MASX J00251994+0031312	00:25:19.92	00:31:31.7	Seyfert 1	0.014	Blue	0.82	0.04	
LEDA 3107905	00:27:53.84	-00:58:00.2	Galaxy	0.014	Blue	0.98	0.02	
SDSS J002916.79-010021.5	00:29:16.81	-01:00:23.1	Galaxy	0.013	Blue	1.00	0.00	
SDSS J002940.01+010528.5	00:29:40.02	01:05:28.7	QSO	1.387	Blue	0.54	0.06	
[VV2010c] J002951.5+004159	00:29:51.45	00:42:00.0	AGN	0.315	Blue	0.20	0.12	
SDSS J003117.70+001705.0	00:31:17.69	00:17:05.1	QSO	4.335	-	-	-	
2QZ J003137.5-292815	00:31:37.50	-29:28:15.3	Unknown	-	Blue	0.68	0.04	
2dFGRS TGS283Z142	00:31:50.70	-28:55:36.7		0.013	Blue	0.92	0.01	
2QZ J003152.5-293534	00:31:52.56	-29:35:33.3	Galaxy	0.313	Blue	0.90	0.07	
2SLAQ J003208.53-005303.7	00:32:08.53	-00:53:03.6	QSO	1.344	Blue	0.23	0.04	
SDSS J003234.62-001557.1	00:32:34.62	-00:15:57.1	QSO	3.243	Blue	0.25	0.05	
LEDA 559945	00:32:34.69	-42:40:10.4	Galaxy	-	Blue	0.71	0.05	
[VV2006] J003242.7+003111	00:32:42.74	00:31:11.1	QSO	0.360	Blue	0.13	0.04	
2dFGRS TGS365Z059	00:33:54.71	-29:56:12.7	Galaxy	0.006	Blue	0.79	0.03	
SWIRE J003517.14-420518.6	00:35:17.11	-42:05:19.0	AGN	0.320	Red	0.02	0.41	
[VV2006] J003545.9+002306	00:35:45.86	00:23:06.0	QSO	3.237	Blue	0.83	0.03	
2MASS J00362543-0029075	00:36:25.39	-00:29:07.1	AGN	0.308	Red	0.07	0.49	
2dFGRS TGS440Z027	00:36:38.44	-32:34:44.7	Galaxy	0.006	Blue	0.17	0.05	
[VV2006] J003714.1-005602	00:37:14.11	-00:56:04.0	QSO	4.361	-	-	-	
[VV2006] J003722.2-001140	00:37:22.17	-00:11:40.6	QSO	1.370	Blue	0.70	0.05	
UM 260	00:37:41.13	00:33:20.0	EmObj	0.014	Blue	0.68	0.08	
SDSS J003859.34-004252.2	00:38:59.35	-00:42:52.0	QSO	2.502	Blue	0.07	0.03	
FASTT 8	00:39:09.42	00:40:11.5	PulsV*	-	Blue	0.64	0.08	
SDSS J003930.30+012021.6	00:39:30.28	01:20:20.9	BlueCompG	0.015	Blue	0.18	0.04	

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
						HAC	HDBSCAN
IRAS 00370+0035	00:39:34.78	00:51:36.9	FIR	–	Blue	0.80	0.04
2QZ J004215.6-321257	00:42:15.62	-32:12:57.2	Galaxy	0.317	Blue	0.87	0.06
GALEX 2673249256393934953	00:42:43.87	01:17:02.2	QSO	1.366	Blue	0.47	0.05
2dFGRS TGS501Z235	00:43:21.88	-33:19:02.9	Galaxy	0.015	Blue	0.34	0.09
2SLAQ J004335.13-003729.7	00:43:35.16	-00:37:29.6	CataclyV*	–	Blue	0.32	0.09
SDSS J004415.83-004303.1	00:44:15.81	-00:43:03.1	QSO	3.248	–	–	–
[VV2006] J004544.4-315729	00:45:44.35	-31:57:29.2	QSO	1.344	Blue	0.48	0.06
2SLAQ J004626.30-011417.0	00:46:26.30	-01:14:16.8	Galaxy	–	Blue	0.84	0.04
[VV98] J004826.9-341340	00:48:26.97	-34:13:38.7	QSO	1.910	Blue	0.44	0.08
SDSS J004918.52+011308.9	00:49:18.52	01:13:09.1	QSO	1.339	Blue	0.37	0.06
LEDA 3034	00:51:49.42	00:33:53.8	Seyfert 1	0.015	Blue	0.86	0.03
QSO B0049-272	00:51:55.64	-26:57:43.3	QSO	2.484	Blue	0.13	0.04
[BKD2008] WR 353	00:51:59.72	-00:29:20.8	PartofG	0.005	Blue	0.56	0.08
ESO 411-27	00:52:51.61	-27:19:32.8	Galaxy	0.006	Blue	0.98	0.02
SDSS J005343.78+012147.6	00:53:43.76	01:21:47.5	QSO	1.358	Blue	0.37	0.05
RESOLVE rf554	00:54:15.54	-01:04:56.0	Galaxy	0.015	Blue	0.97	0.03
QSO B0051-2635	00:54:18.49	-26:19:04.5	QSO	2.010	Blue	0.08	0.03
2QZ J005440.1-320042	00:54:40.12	-32:00:42.2	EmG	0.324	Blue	0.57	0.04
QSO B0052-307	00:54:43.95	-30:30:54.1	QSO	2.450	Blue	0.41	0.06
[VV2006] J005532.1-311538	00:55:32.08	-31:15:37.8	QSO	1.350	Blue	0.46	0.05
[TYZ2012] II 11	00:55:41.32	-00:56:30.6	Galaxy	0.015	Blue	0.78	0.02
[CT83] 219	00:55:51.35	-30:56:42.8	UV	–	Blue	0.36	0.05
RGO 8439	00:55:53.16	-28:54:57.3	Star	–	Blue	0.99	0.01
2dFGRS TGS502Z028	00:55:53.31	-33:39:01.5	Galaxy	0.325	Blue	0.43	0.10
[VV2006] J005609.9-312209	00:56:09.93	-31:22:08.6	QSO	2.460	Blue	0.20	0.05
[VV2006] J005639.0-315759	00:56:39.05	-31:57:58.6	QSO	1.350	Blue	0.99	0.01
[GPM2009] 0057-0022	00:57:12.60	-00:21:57.7	Galaxy	0.010	Blue	0.45	0.08
[VV2000] J005840.5-300203	00:58:40.42	-30:02:00.1	QSO	1.361	Blue	0.22	0.04
LEDA 3530	00:59:04.10	01:00:04.2	GinCl	0.018	Blue	0.96	0.04
2dFGRS TGS503Z245	00:59:13.57	-34:19:15.7	Galaxy	0.012	Blue	0.97	0.03
LBQS 0057-0135	00:59:48.81	-01:19:05.2	QSO	0.325	Blue	0.27	0.05
QSO B0057-3948	00:59:53.21	-39:31:57.3	QSO	3.240	Blue	0.83	0.05
CAIRNS J005959.59-005157.2	00:59:59.58	-00:51:57.1	GinCl	0.166	Red	0.00	1.00
SCMS 679	01:00:04.44	-33:39:32.5	Star	–	Blue	0.51	0.05
2QZ J010009.9-320131	01:00:09.94	-32:01:31.1	Unknown	–	Blue	0.97	0.03
2dFGRS TGS561Z059	01:00:16.17	-34:57:40.6	Galaxy	0.113	Blue	0.52	0.07
2SLAQ J010121.76-000301.7	01:01:21.76	-00:03:01.8	Galaxy	–	Blue	0.08	0.04
QSO B0059-304B	01:02:14.65	-30:07:53.8	QSO	3.240	Blue	0.94	0.01
2SLAQ J010230.03-003206.8	01:02:30.02	-00:32:06.8	Seyfert 1	0.343	Blue	0.31	0.06
2MASX J01023175+0120363	01:02:31.78	01:20:36.1	GinCl	0.016	Blue	0.87	0.03
[VV2006] J010336.4-005508	01:03:36.39	-00:55:08.8	QSO	2.443	Blue	0.33	0.04
SDSS J010413.86-011552.1	01:04:13.86	-01:15:52.0	QSO	1.366	Blue	0.96	0.04
QSO B0103+00	01:06:19.23	00:48:23.4	QSO	4.435	Red	0.02	0.03
2MASS J01065344-3243420	01:06:53.44	-32:43:41.9	AGN	0.371	Blue	0.66	0.03
LIRAS J010658.95+010438.3	01:06:58.93	01:04:38.2	AGN	0.327	Red	0.08	0.46
[VV2006] J010705.6+000609	01:07:05.55	00:06:09.0	QSO	1.357	Blue	0.46	0.05
UGC 695	01:07:46.47	01:03:50.3	LSB G	0.002	Blue	0.89	0.07
SDSS J010748.62+004453.5	01:07:48.62	00:44:53.7	BCIG	0.266	Red	0.10	0.31
MCG+00-04-011	01:09:01.58	01:22:41.5	GinCl	0.018	Blue	0.76	0.08
SDSS J010907.59+000649.8	01:09:07.59	00:06:50.0	QSO	1.372	Blue	0.34	0.04
LEDA 1185205	01:09:07.95	01:07:15.5	HII G	0.004	Blue	0.46	0.08
SDSS J010918.56+005419.3	01:09:18.56	00:54:19.4	QSO	1.356	Blue	0.40	0.05
SDSS J010925.95-003739.0	01:09:25.96	-00:37:39.0	QSO	1.360	Blue	0.70	0.05
2QZ J011014.0-302445	01:10:13.97	-30:24:44.5	EmG	0.313	Blue	0.61	0.05
2QZ J011119.0-300019	01:11:19.02	-30:00:18.2	EmG	0.309	Blue	0.85	0.06
SDSS J011128.38+000143.7	01:11:28.35	00:01:43.3	QSO	0.765	Blue	0.07	0.06
2dFGRS TGS505Z356	01:12:12.64	-33:56:31.1	Galaxy	0.333	Blue	0.43	0.15
2SLAQ J011230.55+001441.5	01:12:30.55	00:14:41.7	QSO	3.259	Blue	0.82	0.04

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group		P(Blue)	P(Red)
					HAC	HDBSCAN		
PB 6318	01:12:58.01	00:58:37.0	Star	–	Blue	0.40	0.03	
UGC 772	01:13:40.42	00:52:39.0	LSB G	0.004	–	–	–	–
2SLAQ J011402.35-004750.9	01:14:02.35	-00:47:50.8	Seyfert 1	0.350	Blue	0.34	0.04	
[VV2006] J011405.3-310903	01:14:05.25	-31:09:02.8	QSO	1.333	Blue	0.13	0.03	
2dFGRS TGS505Z120	01:14:36.11	-32:38:41.0	Galaxy	0.020	Blue	0.68	0.08	
SDSS J011531.90-005144.5	01:15:31.89	-00:51:44.3	HII G	0.005	Blue	0.89	0.06	
2SLAQ J011533.07-005134.9	01:15:33.06	-00:51:34.8	Galaxy	–	Blue	0.55	0.08	
[BKD2008] WR 354	01:15:33.79	-00:51:31.5	HII G	0.006	Blue	0.14	0.05	
SDSS J011542.18+002300.2	01:15:42.18	00:23:00.4	QSO	1.373	Blue	0.70	0.05	
2dFGRS TGS505Z064	01:16:38.40	-32:55:39.1	Galaxy	0.018	Blue	0.95	0.02	
2dFGRS TGS505Z018	01:17:40.21	-33:04:40.7	Galaxy	0.018	Blue	0.68	0.08	
2dFGRS TGS377Z137	01:17:56.44	-30:26:25.8	Galaxy	0.018	Blue	0.79	0.08	
2dFGRS TGS506Z276	01:18:05.71	-33:03:09.1	Galaxy	0.012	Blue	0.88	0.07	
SDSS J011818.13+001455.2	01:18:18.12	00:14:55.5	QSO	1.372	Blue	0.60	0.05	
2SLAQ J011829.63+004549.4	01:18:29.62	00:45:49.4	Seyfert 1	0.314	Blue	0.40	0.09	
2dFGRS TGS506Z243	01:18:49.15	-33:20:13.1	Galaxy	0.012	Blue	0.53	0.05	
2MASX J01195427-3414599	01:19:54.23	-34:15:00.0	EmG	0.019	Blue	0.99	0.01	
2dFGRS TGS506Z158	01:20:09.99	-33:14:10.7	Galaxy	0.011	Blue	1.00	0.00	
SDSS J012110.74-005037.2	01:21:10.74	-00:50:37.1	QSO	1.352	Blue	0.75	0.02	
[HB93] 0119-341B	01:21:52.19	-33:56:15.8	Star	–	Blue	0.20	0.12	
SDSS J012213.85+005731.4	01:22:13.87	00:57:31.6	HII G	0.008	Blue	1.00	0.00	
2dFGRS TGS565Z149	01:22:17.09	-34:02:41.6	Galaxy	0.012	Blue	0.84	0.05	
SDSS J012226.76+000327.5	01:22:26.75	00:03:27.9	QSO	2.480	Blue	0.84	0.04	
QSO B0120-002	01:23:01.78	00:03:23.6	QSO	1.356	Blue	0.47	0.06	
2dFGRS TGS297Z222	01:23:50.87	-29:11:46.4	Galaxy	0.000	Blue	0.03	0.02	
MCG+00-04-113	01:23:54.75	00:16:56.4	GinCl	0.018	Blue	0.71	0.05	
SDSS J012356.34+001230.6	01:23:56.35	00:12:31.0	Galaxy	–	Blue	0.99	0.01	
ESO 352-67	01:23:57.47	-33:48:07.5	Galaxy	0.005	Blue	0.89	0.06	
SDSS J012405.73+005905.0	01:24:05.73	00:59:04.9	Galaxy	0.007	–	–	–	
QSO B0121-324	01:24:16.18	-32:12:21.7	QSO	1.358	Blue	0.37	0.05	
2dFGRS TGS507Z113	01:24:30.16	-33:38:45.5	Galaxy	0.305	Red	0.25	0.27	
QSO B0122-3232	01:25:04.59	-32:17:14.6	QSO	2.450	Blue	0.26	0.06	
2QZ J012526.2-304433	01:25:26.24	-30:44:32.8	EmG	0.311	Blue	0.35	0.04	
2QZ J012549.3-280944	01:25:49.29	-28:09:43.6	Galaxy	0.324	–	–	–	
LEDA 1180903	01:26:27.03	00:58:51.9	Galaxy	0.008	Blue	0.97	0.03	
2dFGRS TGS566Z338	01:26:37.73	-34:35:13.8	Galaxy	0.012	Blue	0.92	0.02	
LAMOST J012644.22-011833.1	01:26:44.21	-01:18:33.0	Galaxy	0.016	Blue	0.83	0.04	
6dFGS gJ012646.5-003845	01:26:46.51	-00:38:44.7	HII G	0.006	Blue	0.54	0.08	
ESO 413-7	01:27:59.31	-29:05:12.0	GinCl	0.005	Blue	1.00	0.00	
6dFGS gJ012926.6-011159	01:29:26.54	-01:11:59.0	GinCl	0.016	Red	0.25	0.16	
SDSS J013034.18-002106.6	01:30:34.17	-00:21:06.5	QSO	3.234	Blue	0.65	0.05	
2dFGRS TGS509Z295	01:31:21.84	-33:06:06.2	Galaxy	0.017	Blue	0.82	0.07	
2dFGRS TGS508Z142	01:31:45.65	-32:56:56.8	Galaxy	0.017	Blue	0.98	0.02	
LEDA 679811	01:31:47.24	-33:10:55.1	Galaxy	–	Blue	1.00	0.00	
2dFGRS TGS509Z242	01:32:53.43	-33:26:42.7	Galaxy	0.017	Blue	0.39	0.08	
SDSS J013303.18+005102.6	01:33:03.19	00:51:02.8	Galaxy	0.313	Blue	0.48	0.12	
2MASS J01330450+0003553	01:33:04.52	00:03:56.1	low-mass*	0.000	Red	0.09	0.33	
2SLAQ J013400.41-010358.2	01:34:00.46	-01:03:59.2		–	Blue	0.95	0.01	
RESOLVE rf246	01:34:52.04	-00:38:55.2	Galaxy	0.017	Blue	0.40	0.08	
[VV2006] J013500.8-004054	01:35:00.83	-00:40:54.2	QSO	1.007	Blue	0.10	0.03	
FBQS J0135-0019	01:35:17.53	-00:19:39.0	Seyfert 1	0.312	Blue	0.28	0.07	
2QZ J013531.1-313651	01:35:31.16	-31:36:51.0	Galaxy	0.320	Blue	0.28	0.05	
SDSS J013701.72-012059.3	01:37:01.71	-01:20:59.1	QSO	2.496	Blue	0.53	0.04	
[VV2006] J013729.4-320715	01:37:29.40	-32:07:15.7	QSO	1.368	Blue	0.57	0.05	
[VV2006] J013837.3+002818	01:38:37.28	00:28:18.5	QSO	1.348	Blue	0.83	0.08	
[VV2006] J013847.3-290547	01:38:47.30	-29:05:47.9	QSO	1.349	Blue	0.97	0.03	
SDSS J013951.07+002537.9	01:39:51.07	00:25:38.0	QSO	1.342	Blue	0.96	0.04	
LEDA 6178	01:40:17.06	-00:50:03.0	Seyfert 1	0.334	Blue	0.17	0.03	

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
						HAC	HDBSCAN
SDSS J014125.63+000755.6	01:41:25.64	00:07:55.8	QSO	0.322	Red	0.02	0.53
[VV2006] J014224.7-320414	01:42:24.73	-32:04:13.7	QSO	2.460	Blue	0.34	0.06
[VV2006] J014303.6-295255	01:43:03.49	-29:52:54.8	QSO	2.450	Blue	0.28	0.05
ESO 353-36	01:43:18.29	-34:12:22.4	EmG	0.013	Red	0.04	0.23
SDSS J014721.12-004505.3	01:47:21.12	-00:45:05.3	QSO	1.348	Red	0.18	0.25
[VV2006] J014739.2-285259	01:47:39.21	-28:52:59.2	QSO	0.360	Blue	0.47	0.03
SDSS J014806.25-002841.6	01:48:06.25	-00:28:41.6	Galaxy	0.215	Blue	0.73	0.05
2QZ J014844.1-275610	01:48:44.17	-27:56:11.0	Star	–	Blue	0.73	0.07
IC 1734	01:49:17.03	-32:44:33.1	Galaxy	0.017	Blue	0.60	0.07
[VV2006] J014921.5-003220	01:49:21.53	-00:32:20.9	QSO	1.379	Blue	0.47	0.06
[RGD2013] J015239.744+010557.768	01:52:39.74	01:05:57.9	Galaxy	0.331	Red	0.07	0.39
[RGD2013] J015253.854+011215.480	01:52:53.85	01:12:15.5	Galaxy	0.184	Red	0.05	0.88
2QZ J015257.7-284838	01:52:57.76	-28:48:37.8	Seyfert 1	0.326	Blue	0.31	0.05
2SLAQ J015331.85+002252.8	01:53:31.85	00:22:53.0	QSO	1.367	Blue	0.95	0.01
SDSS J015400.48-004509.5	01:54:00.50	-00:45:10.0	HII G	0.016	Blue	0.68	0.05
2SLAQ J015409.27+002645.2	01:54:09.27	00:26:45.3	QSO	1.355	Blue	0.72	0.05
SDSS J015440.44-000643.9	01:54:40.45	-00:06:43.6	EmG	0.019	Blue	0.22	0.05
2SLAQ J015526.89+000615.4	01:55:26.87	00:06:15.8	Galaxy	0.016	Blue	0.82	0.02
2SLAQ J015529.12-003927.3	01:55:29.07	-00:39:27.1	Galaxy	–	Blue	0.41	0.04
SDSS J015813.75+010143.5	01:58:13.75	01:01:43.4	RRLyr	–	Blue	0.85	0.02
[VV2006] J015832.1-301703	01:58:32.16	-30:17:02.7	QSO	1.380	Blue	0.65	0.04
[VV2006] J015850.2-300438	01:58:50.22	-30:04:38.1	QSO	1.351	–	–	–
[VV2006] J015935.4+000401	01:59:35.48	00:04:01.5	QSO	3.277	Blue	0.60	0.08
SDSS J020025.40+002916.5	02:00:25.40	00:29:16.8	QSO	0.313	Red	0.03	0.42
[VV2006] J020055.0-293527	02:00:55.02	-29:35:26.5	QSO	1.349	Blue	0.61	0.05
ESO 414-22	02:01:14.49	-31:43:42.9	GinGroup	0.019	Blue	0.68	0.10
[VV98] J020115.4+003136	02:01:15.53	00:31:35.1		0.362	Blue	0.21	0.04
2SLAQ J020200.06-000921.2	02:02:00.06	-00:09:21.2	QSO	1.359	Blue	0.63	0.05
[VV96] J020435.5-455923	02:04:35.46	-45:59:24.0	QSO	3.240	Blue	0.76	0.05
LEDA 1193771	02:05:00.83	01:24:03.7	Galaxy	–	Blue	0.60	0.07
2dFGRS TGS514Z164	02:07:20.33	-33:01:54.3	Galaxy	0.011	–	–	–
2SLAQ J020804.48-000023.2	02:08:04.49	-00:00:23.0	QSO	1.339	Blue	0.54	0.05
2SLAQ J020827.06-005208.1	02:08:27.07	-00:52:07.9	QSO	1.341	Blue	0.51	0.05
SDSS J020921.99-005455.5	02:09:22.00	-00:54:55.4	QSO	1.367	Blue	0.96	0.04
2dFGRS TGS515Z070	02:12:25.05	-33:04:59.0	Galaxy	0.106	–	–	–
2dFGRS TGS515Z311	02:14:24.22	-33:14:52.0	Galaxy	0.012	Blue	0.79	0.09
2dFGRS TGS461Z092	02:14:47.63	-32:42:35.2	Galaxy	0.012	Blue	0.79	0.06
2SLAQ J021529.02-005314.8	02:15:29.02	-00:53:14.9	QSO	1.369	Blue	0.61	0.06
2dFGRS TGS387Z025	02:16:13.75	-30:50:56.8	Galaxy	0.012	Blue	1.00	0.00
SDSS J021617.19-011046.9	02:16:17.19	-01:10:46.7	QSO	3.264	Blue	0.94	0.04
2SLAQ J021810.52-010147.4	02:18:10.52	-01:01:47.2	QSO	1.353	Blue	0.42	0.04
V* AX For	02:19:28.00	-30:45:46.0	CataclyV*	–	Blue	0.28	0.04
2QZ J022112.5-302559	02:21:12.54	-30:25:59.0		0.315	Blue	0.49	0.05
2SLAQ J022316.91-010049.7	02:23:16.93	-01:00:49.6	Galaxy	–	Blue	0.53	0.04
SHOC 120	02:24:17.14	00:06:26.1	Seyfert 1	0.060	–	–	–
LEDA 667000	02:24:52.74	-34:06:34.3	Galaxy	–	Blue	0.78	0.06
SDSS J022530.93-005007.0	02:25:30.92	-00:50:07.1	Galaxy	0.059	Blue	0.65	0.08
[BKD2008] WR 346	02:26:28.28	01:09:37.6	PartofG	0.005	Blue	0.99	0.01
LEDA 546974	02:26:46.27	-43:35:29.6	Galaxy	–	Blue	0.41	0.05
SDSS J022714.48+010536.1	02:27:14.47	01:05:36.3	EmG	0.349	Blue	0.42	0.05
RESOLVE rf668	02:27:19.29	01:01:32.2	Galaxy	0.015	Blue	0.98	0.02
[VV2006] J022738.3-313627	02:27:38.28	-31:36:26.4	QSO	1.350	Blue	0.28	0.04
[VV2006] J022758.2+000226	02:27:58.20	00:02:25.6	QSO	1.066	–	–	–
LCRS B022613.7-392927	02:28:14.52	-39:16:04.0	Galaxy	–	Blue	0.73	0.05
[BKD2008] WR 315	02:28:28.73	-01:08:58.6	PartofG	0.005	Blue	0.88	0.02
2SLAQ J022945.34+000856.2	02:29:45.34	00:08:56.4	Star	–	Blue	0.33	0.06
2QZ J022954.6-303558	02:29:54.69	-30:35:58.4	Seyfert 1	0.372	Blue	0.21	0.04
Pul -3 180355	02:29:57.00	-01:00:32.3	Star	–	Red	0.02	0.04

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group HAC	P(Blue)	P(Red)
						HDBSCAN	HDBSCAN
2dFGRS TGS463Z130	02:30:09.77	-31:36:18.2	Galaxy	0.015	Blue	1.00	0.00
SDSS J023020.93+001355.5	02:30:20.93	00:13:55.8	Seyfert 1	0.335	Blue	0.13	0.04
UGC 2009	02:32:09.34	-01:23:10.3	Galaxy	0.015	Red	0.02	0.53
2dFGRS TGS518Z089	02:32:18.08	-33:50:43.0	Galaxy	0.016	Blue	0.35	0.09
SDSS J023230.63-011654.5	02:32:30.63	-01:16:54.5	QSO	1.364	Blue	0.59	0.05
6dFGS gJ023241.8-391744	02:32:41.88	-39:17:42.8	Galaxy	0.005	Blue	0.52	0.08
SDSS J023248.71+005138.8	02:32:48.71	00:51:38.8	Galaxy	0.344	Blue	0.35	0.04
V* HP Cet	02:33:22.62	00:50:59.4	Nova	-0.000	Blue	0.56	0.09
NGC 986	02:33:34.29	-39:02:40.9	EmG	0.007	Blue	0.93	0.07
[VV2006] J023335.4-010744	02:33:35.37	-01:07:44.6	QSO	0.367	Blue	0.31	0.04
SDSS J023628.77-005829.7	02:36:28.75	-00:58:30.0	Galaxy	0.008	Blue	0.76	0.08
[VV2006] J023635.7-003203	02:36:35.69	-00:32:03.4	QSO	1.362	–	–	–
SDSS J024059.15+004545.8	02:40:59.14	00:45:45.9	QSO	3.233	Blue	0.96	0.04
[VV2006] J024235.0-010351	02:42:34.91	-01:03:51.9	QSO	1.373	Blue	0.57	0.04
[EKS96] NGC 1068 91	02:42:46.94	00:01:26.2	HII	–	Blue	0.08	0.04
[ZBF2015] NGC1073 1	02:43:35.61	01:22:37.9	HII	–	Blue	0.17	0.04
[ZBF2015] NGC1073 16	02:43:37.69	01:22:22.5	HII	–	Blue	0.94	0.06
[ZBF2015] NGC1073 21	02:43:42.74	01:21:34.4	HII	–	Blue	0.49	0.07
[ZBF2015] NGC1073 10	02:43:44.03	01:22:40.5	HII	–	Blue	0.51	0.03
6dFGS gJ024605.3-330500	02:46:05.28	-33:04:59.4	Galaxy	0.017	Blue	0.82	0.01
2MASS J02462415-0029539	02:46:24.14	-00:29:52.9	Star	–	Blue	0.78	0.03
Gaia DR2 2497764348684940160	02:46:24.75	-00:30:16.3	QSO	–	Blue	0.34	0.05
[BKD2008] WR 316	02:46:25.42	-00:30:09.8	PartofG	0.005	Blue	0.56	0.05
2SLAQ J024626.59-003000.2	02:46:26.57	-00:30:00.4	Star	–	Blue	0.99	0.01
2SLAQ J025100.64+001707.2	02:51:00.64	00:17:07.3	QSO	2.466	Blue	0.44	0.04
2SLAQ J025216.75+001741.2	02:52:16.73	00:17:41.2	Galaxy	0.005	Blue	0.89	0.06
2SLAQ J025252.02-002211.7	02:52:52.00	-00:22:11.6	QSO	1.370	Blue	0.83	0.05
SHOC 143	02:54:26.13	-00:41:22.7	Seyfert 1	0.015	Blue	0.88	0.08
NVSS J025557+004132	02:55:57.24	00:41:33.5	RadioG	0.014	Red	0.08	0.22
QSO B0253+0058	02:56:07.25	01:10:38.8	QSO	1.349	Blue	0.40	0.05
HBQS 0253+0022	02:56:25.32	00:34:29.4	HII	0.013	Blue	0.97	0.03
LEDA 1170514	02:56:28.43	00:36:28.2	Galaxy	0.009	Blue	0.62	0.05
LAMOST J025710.60-001331.6	02:57:10.63	-00:13:31.1	Galaxy	0.016	Blue	0.99	0.01
2dFGRS TGS522Z138	02:57:45.54	-33:28:55.5	Galaxy	0.335	Red	0.25	0.19
2MASSI J0259103-002239	02:59:10.38	-00:22:39.8	Seyfert 1	0.360	Blue	0.31	0.05
LAMOST J030102.94-004423.6	03:01:02.91	-00:44:23.3	Galaxy	0.009	Blue	0.32	0.04
2SLAQ J030309.82+001337.5	03:03:09.83	00:13:37.8	Galaxy	–	–	–	–
UGC 2517	03:04:12.47	-01:11:33.8	Galaxy	0.013	Blue	0.32	0.13
2SLAQ J030417.77-004931.7	03:04:17.77	-00:49:31.5	Galaxy	–	Blue	0.40	0.08
LEDA 1142424	03:04:34.76	-00:28:30.7	Seyfert 1	0.006	Blue	0.97	0.03
LBQS 0302-0019	03:04:49.85	-00:08:13.4	QSO	3.295	Blue	0.95	0.02
2MASS J03045799+0057131	03:04:57.98	00:57:14.0	Blue	0.012	Blue	0.99	0.01
MCG+00-08-089	03:05:18.24	-00:09:34.1	Galaxy	0.009	Blue	0.74	0.05
LBQS 0303+0110	03:06:12.72	01:21:57.3	QSO	1.335	Blue	0.44	0.05
WISE J030629.21-335332.3	03:06:29.22	-33:53:32.3	AGN	0.780	Blue	0.41	0.07
SDSS J030630.33-000622.9	03:06:30.33	-00:06:22.9	Galaxy	0.106	Red	0.04	0.48
SDSS J030715.63+004352.1	03:07:15.60	00:43:52.6	Galaxy	0.010	Blue	1.00	0.00
2SLAQ J030757.55+000712.0	03:07:57.55	00:07:12.1	QSO	1.343	Blue	0.74	0.04
2SLAQ J031129.69-001701.4	03:11:29.70	-00:17:01.5	QSO	1.357	Blue	0.70	0.05
2QZ J031130.9-315250	03:11:30.92	-31:52:51.1	WD*	–	Blue	0.35	0.08
ESO 417-20	03:12:48.61	-31:29:10.7	GinGroup	0.013	Blue	0.97	0.03
SDSS J031258.36-000453.6	03:12:58.36	-00:04:53.6	Galaxy	0.117	Blue	0.78	0.03
2dFGRS TGS398Z109	03:13:56.08	-31:28:12.6	Galaxy	0.014	Blue	0.19	0.06
2SLAQ J031428.25+004506.6	03:14:28.25	00:45:07.0	Galaxy	–	Blue	1.00	0.00
2dFGRS TGS471Z114	03:16:15.31	-31:12:33.3	Galaxy	0.004	Blue	0.82	0.04
2SLAQ J031618.00-003025.2	03:16:18.01	-00:30:24.9	Galaxy	–	Blue	0.99	0.01
2dFGRS TGS524Z054	03:16:50.66	-33:18:04.0	Galaxy	0.006	Blue	0.96	0.04
2SLAQ J031829.06-000040.3	03:18:29.06	-00:00:40.5	Galaxy	–	Blue	0.66	0.04

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	HAC	P(Blue)	P(Red)
						HDBSCAN	HDBSCAN
[VV2006] J031845.2-001844	03:18:45.17	-00:18:45.3	QSO	3.224	Blue	0.96	0.04
2SLAQ J031937.30-002641.1	03:19:37.29	-00:26:41.0	QSO	1.371	Blue	0.21	0.05
SDSS J032244.90+004442.4	03:22:44.90	00:44:42.3	QSO Candidate	0.304	Blue	0.97	0.03
6dFGS gJ032504.2-365540	03:25:04.15	-36:55:39.9	GinPair	0.006	Blue	0.65	0.08
6dFGS gJ032512.9-362210	03:25:13.07	-36:22:09.7	Galaxy	0.004	Blue	0.46	0.08
[JPB2015] 051.4803133-32.8829964	03:25:55.27	-32:52:58.9	GlCl	–	Blue	0.50	0.05
SDSS J033226.29-011126.2	03:32:26.29	-01:11:26.0	QSO	1.361	Blue	0.30	0.04
ESO 358-16	03:33:09.26	-35:43:07.4	GinCl	0.006	Blue	0.96	0.04
SDSS J033310.10+000849.1	03:33:10.08	00:08:49.2	Seyfert 2	0.327	Red	0.08	0.24
[VV2006] J033458.5-000744	03:34:58.48	-00:07:43.9	QSO	1.357	Blue	0.81	0.02
[VV2006] J033821.6+003106	03:38:21.51	00:31:06.6	QSO	1.349	Blue	0.63	0.05
2XMM J033841.3-353134	03:38:41.37	-35:31:34.2	BLLac	0.360	Blue	0.16	0.03
[VV2006] J033927.5-344707	03:39:27.45	-34:37:07.0	QSO	1.364	Blue	0.55	0.05
CXO J034012.4-353740	03:40:12.39	-35:37:40.1	HMXB	–	Blue	0.11	0.04
SDSS J034019.89+010330.7	03:40:19.89	01:03:30.7	EmG	0.322	Blue	0.44	0.07
2XMM J034050.4-352620	03:40:50.48	-35:26:21.7	AGN	1.366	Blue	0.36	0.05
ESO 358-51	03:41:32.55	-34:53:19.0	GinGroup	0.006	Blue	0.87	0.07
2MASS J03424773+0109331	03:42:47.72	01:09:33.0	Seyfert 1	0.360	Blue	0.10	0.05
2SLAQ J034304.64+002512.1	03:43:04.65	00:25:12.3	Star	–	Blue	0.26	0.06
LCRS B034214.4-381736	03:44:04.68	-38:08:11.9	Galaxy	–	Red	0.00	1.00
[VV2006] J034408.3-003106	03:44:08.25	-00:31:05.8	QSO	1.646	Blue	0.94	0.03
SDSS J034427.73-002740.4	03:44:27.73	-00:27:40.2	Galaxy	0.041	Blue	0.31	0.18
SDSS J034517.02-001549.8	03:45:17.01	-00:15:49.7	QSO	1.335	Blue	1.00	0.00
6dFGS gJ034545.4-362046	03:45:45.38	-36:20:46.1	GinCl	0.004	Blue	0.89	0.02
SDSS J034602.53-000058.7	03:46:02.53	-00:00:58.6	Seyfert 2	0.308	Red	0.09	0.33
2MASX J03472195-3251054	03:47:21.94	-32:51:05.2	GinCl	0.116	Red	0.15	0.39
SDSS J034907.92+010943.3	03:49:07.93	01:09:43.2	LSB G	0.014	Blue	0.96	0.04
MCG+00-10-021	03:49:08.87	01:09:46.3	LSB G	0.014	Blue	0.84	0.05
FASTT 83	03:51:19.36	00:32:16.6	EB*	–	Red	0.08	0.35
LEDA 607287	03:55:02.55	-38:35:40.2	Galaxy	–	Red	0.07	0.47
Gaia DR2 4857261601188886016	03:55:16.01	-37:29:44.7	Candidate WD*	–	–	–	–
[ZJM2003] SA 95-2230	03:55:38.45	00:28:34.9	Star	–	Blue	0.98	0.02
2dFGRS TGS817Z154	03:56:05.58	-49:28:40.7	Galaxy	0.003	Blue	0.98	0.02
SDSSCGB 74387.1	03:56:50.79	-00:14:34.9	Galaxy	–	Red	0.35	0.31
2dFGRS TGS848Z501	03:57:22.10	-37:01:54.1	Galaxy	0.016	Blue	0.97	0.03
6dFGS gJ035732.5-000047	03:57:32.27	-00:00:47.6	Galaxy	0.017	Blue	0.74	0.07
MCG+00-11-002	03:57:58.63	-00:11:25.9	Seyfert	0.014	Red	0.05	0.29
UGC 2913	03:59:03.91	01:21:33.6	Galaxy	0.013	Blue	0.56	0.10
ESO 201-14	04:00:29.38	-49:01:48.4	EmG	0.004	Blue	0.79	0.06
2MASS J04004608-3424277	04:00:46.07	-34:24:27.7	AGN Candidate	–	Blue	0.33	0.08
6dFGS gJ040053.2-351416	04:00:53.13	-35:14:16.2	Galaxy	0.015	Blue	0.83	0.04
QSO B0401-3505	04:03:10.56	-34:56:56.8	QSO	3.251	Blue	0.68	0.05
LCRS B040209.0-382209	04:03:56.75	-38:13:58.5	Galaxy	–	Blue	0.33	0.09
6dFGS gJ040441.2-345756	04:04:41.14	-34:57:55.8	Galaxy	0.008	Blue	0.81	0.07
6dFGS gJ040520.4-364859	04:05:20.40	-36:48:58.8	Galaxy	0.003	Blue	0.25	0.07
[VV96] J041130.5-335331	04:11:30.51	-33:53:31.1	QSO	1.350	Blue	0.77	0.05
2dFGRS TGS894Z351	04:11:58.37	-37:58:44.0	Galaxy	0.013	Blue	0.96	0.04
ESO 420-11	04:12:53.30	-31:18:30.2	GinGroup	0.005	Blue	0.96	0.04
2dFGRS TGS894Z291	04:13:00.00	-38:19:42.5	Galaxy	0.012	Blue	0.98	0.02
Gaia DR2 4872129059981617536	04:20:06.78	-32:51:20.0	Candidate WD*	–	–	–	–
LEDA 685147	04:20:56.84	-32:50:42.8	Galaxy	–	–	–	–
2MASX J04255940-4316225	04:25:59.40	-43:16:22.6	Galaxy	0.078	Blue	0.98	0.02
LEDA 579779	04:26:32.58	-41:01:56.6	Galaxy	–	Blue	0.98	0.02
LEDA 125483	04:26:44.34	-42:15:41.2	Galaxy	0.015	Blue	1.00	0.00
MCG-07-10-009	04:27:42.24	-42:38:20.4	Galaxy	0.016	Blue	0.85	0.06
2MASX J04282877-4314283	04:28:28.71	-43:14:29.1	Galaxy	0.015	Blue	0.69	0.04
6dFGS gJ043139.6-301514	04:31:39.57	-30:15:14.1	CataclyV*	-0.000	Blue	0.16	0.05
LEDA 697927	04:32:44.32	-32:01:20.1	Galaxy	–	Red	0.10	0.85

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group HAC	P(Blue)	P(Red)
						HDBSCAN	HDBSCAN
ESO 251-12	04:33:09.79	-43:46:13.6	Galaxy	0.010	Blue	0.81	0.05
ESO 304-2	04:35:19.80	-42:12:11.8	Galaxy	0.007	Blue	1.00	0.00
LEDA 494343	04:35:41.79	-47:52:43.9	Galaxy	–	Blue	0.85	0.05
[HM2015b] 236 3	04:37:11.17	-46:41:06.8	Galaxy	–	Blue	0.23	0.04
LEDA 692923	04:37:18.21	-32:22:10.3	Galaxy	–	Blue	1.00	0.00
Gaia DR2 4790561549357871744	04:37:35.04	-44:20:28.8	Candidate RRLyr	–	Blue	1.00	0.00
LEDA 681270	04:38:23.53	-33:05:08.6	Galaxy	–	Blue	0.88	0.02
6dFGS gJ043927.4-425912	04:39:27.45	-42:59:11.8	Galaxy	0.362	Blue	0.26	0.04
LEDA 606705	04:45:44.98	-38:38:48.7	Galaxy	–	Blue	1.00	0.00
LEDA 88363	04:48:23.05	-44:52:57.7	Galaxy	–	Blue	0.96	0.04
2dFGRS TGS880Z325	04:50:02.35	-47:28:39.0	Galaxy	0.010	Blue	0.42	0.06
LEDA 512705	04:51:48.86	-46:37:36.8	Galaxy	–	Blue	1.00	0.00
Gaia DR2 4811451372635865984	04:52:31.12	-44:11:04.3	Star	–	Blue	0.30	0.07
LEDA 686311	04:53:19.46	-32:46:32.8	Galaxy	–	Blue	0.85	0.04
[VV98] J045444.5-481300	04:54:43.04	-48:13:20.2	Seyfert 1	0.363	Blue	0.42	0.03
SN 2012at	04:54:52.74	-37:19:15.5	SN	–	Blue	0.35	0.08
2MASX J04550020-3715351	04:55:00.19	-37:15:35.4	GinPair	0.008	Blue	0.87	0.06
LEDA 715392	04:55:26.51	-30:35:28.6	Galaxy	–	Blue	0.12	0.04
ESO 499-24	09:57:01.69	-26:29:28.5	Galaxy	0.015	Blue	0.87	0.06
LEDA 859547	09:57:06.62	-19:07:06.1	Galaxy	–	Blue	0.81	0.08
LEDA 1022680	09:57:23.46	-07:12:51.4	Galaxy	–	Blue	0.11	0.04
LEDA 860244	09:57:53.74	-19:03:45.5	GinGroup	0.014	Blue	1.00	0.00
2MASX J09583711-4704597	09:58:37.10	-47:05:00.4	Galaxy	0.012	Blue	0.65	0.08
CRTS J095840.2-175426	09:58:40.20	-17:54:27.9	RRLyr	–	Blue	0.99	0.01
[CMI2006b] H42-f02-1939	09:59:46.82	-19:28:00.0	Galaxy	0.265	Blue	0.99	0.01
CRTS J095950.7-383024	09:59:50.88	-38:30:22.9	RRLyr	–	Blue	0.96	0.04
LEDA 605183	10:00:05.44	-38:47:28.7	Galaxy	–	Blue	0.58	0.09
NGC 3095	10:00:05.83	-31:33:10.8	GinGroup	0.009	Red	0.05	0.28
LEDA 154528	10:00:49.12	-30:32:41.9	Galaxy	–	Blue	0.76	0.08
SDSS J100059.08+032751.4	10:00:59.07	03:27:51.5	Galaxy	0.007	Blue	0.46	0.05
COSMOS 1949846	10:01:01.83	02:20:18.3	Star	0.040	Red	0.06	0.50
ESO 567-3	10:01:09.07	-19:26:29.7	LSB G	0.012	Blue	0.99	0.01
LEDA 1011555	10:01:34.08	-07:52:55.7	Galaxy	–	Blue	0.92	0.08
LEDA 835683	10:01:53.87	-20:52:58.7	Galaxy	–	Blue	0.97	0.03
RE J1002-19	10:02:11.73	-19:25:37.1	CataclyV*	–	Blue	0.19	0.04
2QZ J100215.7-001056	10:02:15.83	-00:10:55.8	QSO	0.353	Blue	0.38	0.06
VVDS 100108471	10:02:25.38	01:19:36.8	Galaxy	0.123	Blue	0.98	0.02
ESO 262-15	10:02:38.71	-45:29:53.9	GinGroup	0.012	Red	0.09	0.30
[BCP93] F2 H6	10:02:51.82	-26:09:23.9	HII	–	Blue	0.99	0.01
[EBU2007] 7	10:02:54.69	-26:08:59.6	BlueSG*	0.001	Blue	0.14	0.05
[H69] NGC 3109 12	10:02:56.31	-26:08:58.5	HII	–	Blue	0.16	0.05
[PRS2007] HII 44	10:02:59.47	-26:08:46.4	HII	–	Blue	0.25	0.06
DENIS J100302.1-260857	10:03:02.10	-26:08:58.7	AGN Candidate	–	Blue	0.69	0.08
2MASX J10030450-1949377	10:03:04.51	-19:49:38.1	Galaxy	0.012	Blue	0.86	0.05
CRTS J100306.6-333707	10:03:06.57	-33:37:08.6	EB*	–	Blue	0.68	0.08
2dFGRS TGN094Z280	10:03:15.05	-05:54:32.8	Galaxy	0.013	Blue	0.95	0.05
[EBU2007] 3	10:03:17.64	-26:10:01.7	BlueSG*	0.001	Blue	0.16	0.05
[VV96] J100342.1-150808	10:03:41.93	-15:08:08.9	QSO	0.342	Blue	0.27	0.04
2MASX J10035230-3124480	10:03:52.32	-31:24:48.5	Galaxy	0.009	Blue	0.98	0.02
2MASX J10041992-4425311	10:04:19.86	-44:25:32.6	Galaxy	0.012	Blue	0.92	0.08
CRTS J100433.7-364726	10:04:33.77	-36:47:25.8	EB*	–	Blue	0.99	0.01
2QZ J100506.4+001051	10:05:06.44	00:10:51.7	Galaxy	0.330	Blue	0.99	0.01
2MASX J10050765-1951299	10:05:07.68	-19:51:30.2	Galaxy	–	Blue	0.66	0.07
2dFGRS TGN421Z115	10:05:17.31	01:38:21.7	Galaxy	0.004	Blue	0.87	0.05
2MFGC 7816	10:05:28.51	-38:07:30.1	Galaxy	–	Blue	1.00	0.00
[VV2006] J100539.9+040914	10:05:39.88	04:09:14.7	QSO	1.355	Blue	0.36	0.05
CRTS J100548.9-254146	10:05:48.99	-25:41:47.1	RRLyr	–	Blue	1.00	0.00
2MASX J10060024-4251262	10:06:00.27	-42:51:26.2	Galaxy	0.011	Red	0.10	0.45

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	HAC	P(Blue)	P(Red)
					Group	HDBSCAN	HDBSCAN
2MASX J10061715-0634276	10:06:17.20	-06:34:27.7	Galaxy	0.011	Blue	0.95	0.05
6dFGS gJ100622.2-264958	10:06:22.20	-26:49:57.2	Galaxy	0.015	Blue	0.99	0.01
6dFGS gJ100631.4-320236	10:06:31.39	-32:02:35.9	GinGroup	0.008	Red	0.24	0.16
NGC 3125	10:06:33.37	-29:56:07.8	HII G	0.004	Blue	0.55	0.10
1RXH J100633.9-295612	10:06:33.90	-29:56:11.6	X	–	Blue	0.68	0.07
2MASX J10071106-1904039	10:07:11.07	-19:04:04.6	Galaxy	0.012	Blue	0.82	0.06
LEDA 699275	10:07:27.24	-31:55:26.9	Galaxy	–	Blue	0.93	0.07
CRTS J100733.7-301921	10:07:33.84	-30:19:19.4	EB*	–	Blue	0.81	0.03
RX J1007.5-2017	10:07:34.65	-20:17:32.4	CataclyV*	–	Blue	0.29	0.07
CRTS J100734.9-383117	10:07:35.06	-38:31:17.1	EB*	–	Red	0.07	0.17
CRTS J100751.1-184918	10:07:51.14	-18:49:19.1	EB*	–	Red	0.03	0.44
2MASX J10081071-3331017	10:08:10.76	-33:31:02.2	Galaxy	0.010	Blue	0.33	0.14
2MASX J10082199-1448362	10:08:22.01	-14:48:36.1	Galaxy	0.008	Blue	0.73	0.06
LEDA 768685	10:08:30.09	-26:21:33.0	Galaxy	–	Blue	0.95	0.05
2MASX J10091380-4300089	10:09:13.81	-43:00:09.0	GinGroup	0.015	Red	0.11	0.21
LEDA 648630	10:09:50.26	-35:27:43.3	Galaxy	–	Blue	0.97	0.03
LEDA 3094360	10:09:58.73	-20:30:59.5	Galaxy	–	Blue	0.87	0.06
ESO 435-50	10:10:50.41	-30:25:24.4	Galaxy	0.009	Blue	0.70	0.07
LEDA 654529	10:10:51.81	-35:00:28.1	Galaxy	–	Blue	1.00	0.00
NGC 3146	10:11:09.90	-20:52:14.0	EmG	0.013	Blue	0.74	0.07
LEDA 729120	10:11:13.49	-29:27:27.9	Galaxy	–	Blue	0.39	0.08
ATO J152.8181-17.6910	10:11:16.35	-17:41:27.6	V*	–	Red	0.00	1.00
CRTS J101200.8-365725	10:12:00.81	-36:57:25.2	EB*	–	Red	0.32	0.34
LEDA 691325	10:12:03.36	-32:28:06.8	Galaxy	–	Blue	1.00	0.00
Gaia DR2 5407412036686860672	10:12:47.58	-47:33:51.1	Star	–	Blue	0.64	0.11
LEDA 655538	10:12:59.65	-34:56:06.6	Galaxy	–	Blue	0.96	0.01
2MASX J10134201-3451194	10:13:41.91	-34:51:18.3	EmG	0.015	Blue	0.74	0.08
LEDA 658182	10:13:54.08	-34:44:23.0	Galaxy	–	Blue	0.85	0.04
LEDA 713928	10:14:25.56	-30:42:30.1	Galaxy	–	Blue	0.69	0.04
2MASX J10142679-2329036	10:14:26.81	-23:29:04.9	Galaxy	0.012	Blue	0.98	0.02
ESO 263-21	10:14:41.74	-44:51:14.1	EmG	0.004	Blue	0.83	0.03
IC 2559	10:14:45.36	-34:03:33.0	EmG	0.010	Red	0.16	0.24
ESO 263-22	10:14:48.13	-43:31:49.5	Galaxy	0.010	Blue	0.71	0.06
ESO 263-23	10:14:57.32	-43:37:09.2	Galaxy	0.010	Red	0.20	0.14
ESO 567-32	10:15:44.54	-20:17:44.0	EmG	0.012	Red	0.04	0.21
V* KO Vel	10:15:58.31	-47:58:09.1	CataclyV*	-0.000	Blue	0.20	0.05
IC 2560	10:16:18.68	-33:33:49.8	Seyfert 2	0.010	Red	0.19	0.19
ESO 567-39	10:17:13.15	-21:04:00.3	EmG	0.012	Blue	0.74	0.05
LEDA 702814	10:18:05.73	-31:38:49.0	Galaxy	–	Blue	0.61	0.07
CRTS SSS120320 J101854-400644	10:18:53.51	-40:06:43.7	Candidate CV*	–	Blue	0.18	0.06
ESO 375-7	10:19:01.23	-37:40:19.2	Galaxy	0.016	Blue	1.00	0.00
CTS 1011	10:19:21.17	-22:08:33.4	HII G	0.012	Blue	0.26	0.07
CTS 1011	10:19:21.28	-22:08:35.9	HII G	0.012	Blue	0.36	0.09
NGC 3208	10:19:41.31	-25:48:52.9	EmG	0.010	Blue	0.65	0.07
6dFGS gJ102028.5-232845	10:20:28.52	-23:28:45.3	Galaxy	0.012	Blue	1.00	0.00
LEDA 800754	10:20:32.72	-23:26:54.0	Galaxy	–	Blue	0.43	0.09
CRTS SSS120215 J102042-335002	10:20:42.16	-33:50:02.4	Candidate CV*	–	Blue	0.62	0.08
Gaia DR2 5668001579559758720	10:20:43.31	-20:47:54.6	Star	–	Blue	0.34	0.04
ESO 500-30	10:20:48.90	-23:27:57.1	EmG	0.012	Blue	0.39	0.12
6dFGS gJ102109.3-325140	10:21:09.27	-32:51:39.9	Galaxy	0.010	Blue	0.95	0.05
6dFGS gJ102121.0-213628	10:21:21.03	-21:36:27.7	Galaxy	0.011	Blue	0.98	0.01
LEDA 592969	10:22:02.22	-39:52:45.9	Galaxy	–	Blue	0.89	0.06
2MASS J10223994-3029305	10:22:39.94	-30:29:30.6	AGN Candidate	0.317	Blue	0.27	0.05
ESO 263-30	10:22:59.54	-42:49:38.9	Galaxy	0.009	Blue	0.83	0.04
ESO 317-19	10:23:02.34	-39:09:59.8	GinGroup	0.010	Blue	0.84	0.07
ESO 375-18	10:23:40.27	-35:49:33.5	EmG	0.015	Red	0.21	0.26
ESO 263-32	10:24:21.47	-43:55:01.6	Galaxy	–	Blue	0.85	0.03
CRTS J102424.0-164933	10:24:24.03	-16:49:33.3	EB*	–	Red	0.30	0.25

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
						HAC	HDBSCAN
ESO 500-34	10:24:31.43	-23:33:09.6	Seyfert 2	0.012	Red	0.03	0.32
CRTS J102513.4-354014	10:25:13.46	-35:40:16.7	EB*	–	Blue	0.98	0.02
6dFGS gJ102607.5-243321	10:26:07.41	-24:33:20.2	Galaxy	0.013	Blue	0.44	0.05
ESO 436-21	10:26:21.70	-29:11:57.8	GinCl	–	Blue	0.99	0.01
NAME OT J102706-434341	10:27:05.83	-43:43:41.3	CataclyV*	–	Blue	0.76	0.05
ESO 500-42	10:27:20.27	-23:48:19.6	Galaxy	0.012	Blue	0.95	0.01
SN 2001db	10:27:50.37	-43:54:20.8	SN	–	Red	0.39	0.36
[LWZ2002] 5	10:27:51.28	-43:53:58.5	X	–	Blue	0.44	0.07
NGC 3256	10:27:51.29	-43:54:14.0	IG	0.009	Red	0.20	0.28
[LDT2000] R02	10:27:51.73	-43:54:13.6	X	–	Blue	0.94	0.06
[EF2003] B3	10:27:52.88	-43:54:11.5	HII	0.010	Blue	0.25	0.04
CRTS J102839.9-295911	10:28:40.13	-29:59:12.6	EB*	–	Blue	0.38	0.20
ESO 436-26	10:28:42.96	-31:02:17.7	AGN	0.014	Red	0.34	0.33
CRTS CSS140309 J102844-161303	10:28:43.86	-16:13:03.3	CataclyV*	–	Blue	0.08	0.05
ESO 317-34	10:29:00.71	-40:04:57.9	GinGroup	0.009	Blue	0.96	0.04
IC 2582	10:29:11.07	-30:20:32.7	EmG	0.014	Blue	0.77	0.05
LEDA 636268	10:30:30.59	-36:28:47.1	Galaxy	–	Blue	0.74	0.08
LEDA 83158	10:30:57.69	-34:42:28.5	GinGroup	–	Blue	0.75	0.05
ESO 317-39	10:31:00.18	-40:10:42.5	Galaxy	0.015	Red	0.07	0.19
ESO 436-32	10:31:29.90	-32:42:47.1	EmG	0.013	Blue	0.52	0.12
NGC 3281	10:31:52.11	-34:51:13.0	Seyfert 2	0.011	Red	0.03	0.22
LEDA 571751	10:31:57.37	-41:48:41.1	Galaxy	–	Blue	0.63	0.04
ATO J158.2117-27.8636	10:32:50.82	-27:51:49.2	Candidate EB*	–	Red	0.28	0.43
[BM98] 2	10:32:59.22	-27:32:36.9	GinCl	0.016	Blue	0.97	0.03
6dFGS gJ103317.5-430444	10:33:17.40	-43:04:43.1	Galaxy	0.010	Blue	0.63	0.07
ESO 375-64	10:34:00.75	-35:16:57.3	GinGroup	0.009	Blue	0.62	0.07
LEDA 754029	10:34:26.74	-27:30:04.0	GinCl	0.012	Blue	0.96	0.04
ESO 436-42	10:34:38.75	-28:35:00.1	EmG	0.012	Blue	0.73	0.04
ESO 568-18	10:34:54.59	-20:32:55.6	EmG	0.012	Blue	0.99	0.01
2MASX J10345852-4054438	10:34:58.52	-40:54:43.3	Galaxy	0.016	Blue	0.73	0.06
6dFGS gJ103502.9-293024	10:35:02.88	-29:30:23.8	GinCl	0.012	Blue	0.96	0.04
ESO 437-3	10:35:07.72	-27:59:28.7	EmG	0.008	Blue	0.68	0.04
ESO 375-69	10:35:18.72	-36:52:42.5	EmG	0.011	Blue	0.96	0.04
ESO 501-22	10:35:21.68	-27:41:44.5	GinCl	0.010	Blue	0.95	0.05
6dFGS gJ103530.3-182048	10:35:30.33	-18:20:47.6	Galaxy	0.344	Blue	0.31	0.05
LEDA 712419	10:35:31.70	-30:50:00.0	Galaxy	–	Blue	0.64	0.03
LEDA 535830	10:35:34.16	-44:34:41.1	Galaxy	–	Blue	0.48	0.05
LEDA 784823	10:36:02.66	-24:54:24.1	Galaxy	–	Blue	0.82	0.04
LEDA 743415	10:36:06.94	-28:17:45.0	Galaxy	–	Blue	0.83	0.06
ESO 501-32	10:36:22.11	-25:22:35.4	EmG	0.013	Blue	0.82	0.06
[CZ2003] 1060C-393 25	10:36:30.34	-27:54:04.0	GinCl	0.008	Blue	0.52	0.04
6dFGS gJ103645.4-281005	10:36:45.48	-28:10:02.7	GinCl	0.012	Blue	0.92	0.03
LEDA 769967	10:36:54.86	-26:14:26.0	HII G	0.012	Blue	0.09	0.03
6dFGS gJ103656.1-265414	10:36:56.08	-26:54:13.6	Galaxy	0.096	Blue	0.99	0.01
LEDA 742546	10:37:01.84	-28:22:01.7	GinCl	–	Blue	0.86	0.06
6dFGS gJ103704.4-312157	10:37:04.45	-31:21:57.3	Galaxy	0.010	Blue	0.82	0.04
NGC 3314	10:37:12.87	-27:41:02.2	EmG	0.009	Blue	0.37	0.07
6dFGS gJ103719.9-281420	10:37:19.89	-28:14:19.9	GinCl	0.012	Blue	1.00	0.00
LEDA 753354	10:37:22.21	-27:32:41.9	Galaxy	–	Blue	0.67	0.08
WISE J103754.92-242544.5	10:37:54.92	-24:25:44.6	MIR	–	Red	0.09	0.30
ESO 501-61	10:38:05.84	-25:05:40.1	IG	0.012	Blue	0.97	0.02
LEDA 747838	10:38:11.92	-27:56:14.2	GinCl	0.010	Blue	0.95	0.05
[WLH83] 1036-378A	10:38:14.37	-38:05:25.5	HII	–	Blue	0.94	0.04
LEDA 740766	10:38:28.68	-28:30:55.0	GinCl	–	Blue	0.96	0.04
2MASX J10383034-2332546	10:38:30.34	-23:32:54.7	Galaxy	–	Blue	0.89	0.05
ESO 501-65	10:38:33.42	-27:44:13.8	EmG	0.015	Blue	0.87	0.06
WPVS 78	10:38:41.50	-25:35:32.2	EmG	0.010	Blue	0.29	0.07
6dFGS gJ103857.2-200242	10:38:57.24	-20:02:41.8	Galaxy	0.007	Blue	0.91	0.02

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
						HAC	HDBSCAN
LEDA 838980	10:39:13.02	-20:38:12.5	Galaxy	-	Blue	0.60	0.04
MCG-04-25-054	10:39:26.01	-23:45:16.8	EmG	0.013	Blue	0.61	0.06
2MASS J10395999-4701261	10:39:59.97	-47:01:26.3	CataclyV*	-	Blue	0.34	0.08
CRTS J104001.0-260913	10:40:01.01	-26:09:15.9	EB*	-	Red	0.39	0.32
ESO 437-37	10:40:31.01	-29:16:10.5	IG	0.012	Blue	0.88	0.03
ESO 568-20	10:40:58.70	-21:47:04.3	EmG	0.012	Red	0.11	0.36
6dFGS gJ104102.1-304740	10:41:02.12	-30:47:40.1	Galaxy	0.011	Blue	0.78	0.05
CRTS J104104.0-341120	10:41:03.86	-34:11:23.4	RRLyr	-	Blue	1.00	0.00
ESO 568-21	10:41:15.17	-21:01:22.9	Seyfert 1	0.012	Blue	0.95	0.03
ESO 437-42	10:41:27.71	-31:46:49.1	Galaxy	0.009	Blue	0.82	0.06
LEDA 3081775	10:41:35.15	-37:28:09.5	Galaxy	-	Blue	0.81	0.06
6dFGS gJ104139.4-274638	10:41:39.43	-27:46:38.2	Galaxy	0.014	Blue	0.62	0.07
ESO 568-22	10:42:06.62	-22:06:20.1	IG	0.007	Blue	0.74	0.05
LEDA 31904	10:42:19.50	-36:19:13.7	Galaxy	-	Blue	0.96	0.04
6dFGS gJ104238.0-235609	10:42:37.99	-23:56:08.4	Galaxy	0.003	Blue	0.90	0.03
ESO 437-50	10:43:31.00	-30:46:20.0	EmG	0.013	Blue	0.74	0.03
ATO J160.9042-19.0551	10:43:37.04	-19:03:18.5	Candidate EB*	-	Red	0.04	0.49
6dFGS gJ104409.7-204909	10:44:09.71	-20:49:09.5		0.013	Blue	0.76	0.05
ESO 569-2	10:45:00.21	-22:09:08.2	IG	0.010	Blue	1.00	0.00
6dFGS gJ104534.8-241702	10:45:34.75	-24:17:01.3	Galaxy	0.012	Blue	0.80	0.07
6dFGS gJ104617.1-282524	10:46:17.11	-28:25:23.6	EmG	0.012	Blue	0.72	0.05
LEDA 718607	10:46:30.26	-30:19:17.8	Galaxy	-	Blue	0.85	0.04
ESO 376-20	10:46:38.45	-36:21:11.9	EmG	0.014	Blue	0.69	0.07
ESO 501-96	10:46:47.54	-23:19:39.8	Galaxy	0.011	Blue	0.94	0.06
Gaia DR2 5391507429181636352	10:47:23.91	-41:59:49.3	Star	-	Blue	0.33	0.04
EC 10453-2041	10:47:44.36	-20:57:48.8	EmG	0.012	Blue	0.56	0.09
2MASX J10475221-2004542	10:47:52.13	-20:04:53.5	HII G	0.013	Blue	1.00	0.00
[KRB2015] A	10:48:23.47	-25:09:43.6	Radio	-	Blue	0.38	0.13
2MASX J10482527-2151000	10:48:25.30	-21:51:00.5	Galaxy	0.015	Blue	0.76	0.05
SN 2018aqi	10:48:25.45	-25:09:36.1	SN	0.012	Blue	0.12	0.05
LEDA 688498	10:48:42.32	-32:38:37.4	Galaxy	-	Blue	1.00	0.00
LEDA 738826	10:49:46.88	-28:40:37.1	Galaxy	-	Blue	0.50	0.08
CRTS J104955.6-320059	10:49:55.81	-32:01:00.9	EB*	-	Red	0.35	0.37
2MASX J10503963-1832342	10:50:39.64	-18:32:34.4	Galaxy	0.014	Blue	0.25	0.14
LEDA 844461	10:51:00.37	-20:14:21.3	Galaxy	-	Blue	1.00	0.00
6dFGS gJ105101.9-282017	10:51:01.81	-28:20:16.5	Galaxy	0.011	Blue	0.86	0.03
LEDA 851789	10:51:27.40	-19:41:37.0	Galaxy	-	Blue	0.99	0.01
6dFGS gJ105149.2-215323	10:51:49.07	-21:53:17.5	Galaxy	0.010	Blue	0.68	0.05
6dFGS gJ105233.0-230900	10:52:33.04	-23:08:59.6	Galaxy	0.318	Blue	0.18	0.05
CRTS J105319.4-441618	10:53:19.58	-44:16:20.9	RRLyr	-	Blue	0.88	0.07
MASTER OT J105440.86-391319.0	10:54:40.84	-39:13:19.0	Candidate SN*	-	Blue	0.45	0.05
6dFGS gJ105521.6-232527	10:55:21.63	-23:25:27.3		0.012	Blue	0.92	0.05
2MASX J10563839-2047119	10:56:38.39	-20:47:12.2	Galaxy	0.012	Blue	0.86	0.04
LEDA 849870	10:56:48.51	-19:50:00.4	Galaxy	-	Blue	0.86	0.08
ESO 376-28	10:57:04.32	-33:09:20.3	Galaxy	0.013	Red	0.26	0.21
ESO 264-52	10:57:13.91	-47:40:11.3	Galaxy	0.016	Blue	0.61	0.15
LEDA 648093	10:57:36.52	-35:30:15.8	Galaxy	-	Blue	1.00	0.00
2MASX J10584423-1909304	10:58:44.25	-19:09:31.1	Galaxy	0.012	Blue	1.00	0.00
EC 10566-3120	10:58:59.03	-31:36:34.1	CataclyV*	-	Blue	0.49	0.08
2MASX J10590982-2759589	10:59:09.86	-27:59:59.3		0.005	Blue	0.93	0.04
Gaia DR2 5386613537284200960	11:01:51.26	-46:53:04.5	Candidate RRLyr	-	Blue	1.00	0.00
Gaia DR2 3537117430403448320	11:01:57.97	-23:47:27.3		PM*	-	Blue	0.33
V* TU Crt	11:03:36.57	-21:37:45.9	CataclyV*	-	Blue	0.35	0.06
NGC 3513	11:03:46.16	-23:14:42.1	GinPair	0.004	Blue	0.78	0.04
FAUST 2807	11:03:59.06	-18:46:36.1	UV	-	Blue	1.00	0.00
ESO 570-5	11:07:13.10	-19:49:07.2	IG	0.012	Blue	0.87	0.05
NGC 3529	11:07:19.14	-19:33:19.5	EmG	0.013	Blue	0.98	0.02
NGC 3565	11:07:47.84	-20:01:20.2	IG	0.013	Blue	0.86	0.04

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
					HAC	HDBSCAN	HDBSCAN
ESO 570-10	11:10:50.57	-21:58:28.3	Galaxy	0.012	Blue	0.88	0.09
LEDA 821419	11:10:57.23	-21:56:53.3	Galaxy	–	Blue	0.90	0.03
6dFGS gJ111351.0-212655	11:13:50.97	-21:26:54.8	Galaxy	0.012	Blue	0.89	0.07
NGC 3597	11:14:42.00	-23:43:39.9	EmG	0.012	Blue	0.68	0.09
[VV96] J111644.8-171127	11:16:43.58	-17:11:41.5	QSO	0.375	Blue	0.25	0.05
LEDA 861413	11:17:15.01	-18:58:24.4	Galaxy	–	–	–	–
LEDA 809402	11:17:35.06	-22:45:06.2	Galaxy	–	Red	0.15	0.33
CRTS J112256.0-242841	11:22:56.09	-24:28:40.0	EB*	–	Blue	1.00	0.00
LEDA 786212	11:29:34.18	-24:46:39.1	Galaxy	–	Blue	0.47	0.08
[VV2010c] J113128.4-195903	11:31:28.46	-19:59:02.8	AGN	0.363	Blue	0.36	0.04
CRTS SSS110509 J113219-213943	11:32:19.01	-21:39:42.9	Candidate CV*	–	–	–	–
2dFGRS TGN444Z198	11:34:24.52	01:09:15.7	Galaxy	0.017	Blue	0.88	0.04
MGC 22410	11:36:12.70	00:04:54.9	Star	–	Blue	1.00	0.00
UGC 6578	11:36:36.73	00:49:02.1	EmG	0.004	Blue	0.36	0.09
GAMA 6821	11:36:36.79	00:48:55.8	Galaxy	0.004	Blue	0.05	0.02
V* RZ Leo	11:37:22.18	01:48:58.9	CataclyV*	-0.000	Blue	0.10	0.04
Gaia DR2 3541998025080414336	11:37:49.97	-20:07:37.1	Candidate WD*	–	Blue	0.23	0.06
2dFGRS TGN238Z266	11:38:54.33	-01:38:34.1	Galaxy	0.006	Blue	0.50	0.09
CRTS J113855.5-211148	11:38:55.60	-21:11:47.7	RRLyr	–	Blue	0.97	0.03
SDSS J113901.39+012017.8	11:39:01.39	01:20:17.7	Galaxy	0.005	Blue	0.23	0.07
LBQS 1136-0109	11:39:04.35	-01:26:25.0	QSO	1.375	Blue	0.83	0.04
6dFGS gJ114135.0-181141	11:41:35.04	-18:11:40.5	Galaxy	0.012	Blue	0.98	0.02
2dFGRS TGN238Z191	11:41:45.67	-01:54:04.8	HII G	0.006	Blue	0.88	0.06
ESO 571-16	11:42:09.14	-18:10:08.7	Galaxy	0.012	Red	0.03	0.24
SDSS J114212.38+002002.5	11:42:12.33	00:20:03.4	PartofG	0.019	Blue	0.98	0.02
2QZ J114214.5-023154	11:42:14.64	-02:31:53.3	Galaxy	0.319	Blue	0.88	0.02
CRTS J114238.0-202722	11:42:37.96	-20:27:21.8	RRLyr	–	Blue	0.98	0.01
2QZ J114250.9+013057	11:42:50.95	01:30:58.2	Seyfert 1	0.361	Blue	0.46	0.05
LAMOST J114253.57+000947.7	11:42:53.59	00:09:47.4	Galaxy	0.018	Blue	0.63	0.08
SDSS J114329.34-020319.7	11:43:29.34	-02:03:19.5	QSO	3.304	–	–	–
SDSSCGB 59619.2	11:43:46.11	-01:16:34.0	Galaxy	–	–	–	–
GAMA 396970	11:43:47.41	01:30:53.9	Galaxy	0.102	Blue	0.59	0.06
SDSS J114408.82+012420.5	11:44:08.82	01:24:20.7	RRLyr	0.001	Blue	0.98	0.02
2QZ J114450.8+014324	11:44:50.95	01:43:24.8	EmG	0.333	Blue	0.53	0.08
Gaia DR2 3544179185567992320	11:44:55.76	-17:56:39.4	Candidate WD*	–	Blue	0.27	0.07
2dFGRS TGN310Z256	11:45:08.04	-00:59:18.2	Galaxy	0.004	Blue	0.48	0.10
SDSS J114511.70-005402.6	11:45:11.72	-00:54:02.5	Galaxy	0.204	Red	0.06	0.46
2MASX J11451524-2044471	11:45:15.26	-20:44:47.5	Galaxy	0.012	Blue	0.71	0.08
Z 12-78	11:45:26.30	00:00:14.8	Galaxy	0.013	Blue	0.96	0.04
[SHM2017] J176.40613-17.33394	11:45:37.48	-17:20:02.2	RRLyr	–	Blue	1.00	0.00
SDSS J114600.44+001037.4	11:46:00.45	00:10:37.0	Galaxy	0.311	Red	0.15	0.39
2dFGRS TGN310Z211	11:46:07.72	-00:27:28.7	Galaxy	0.013	Blue	0.97	0.03
SDSS J114643.10+011118.6	11:46:43.12	01:11:18.8	QSO	3.220	Blue	0.81	0.05
2QZ J114711.4-002706	11:47:11.47	-00:27:05.8	EmG	0.312	Blue	0.97	0.03
2MASX J11481815-0138230	11:48:18.21	-01:38:23.8	Seyfert 1	0.013	Blue	0.81	0.06
SDSS J114818.33-013830.8	11:48:18.35	-01:38:30.5	Galaxy	0.013	Blue	0.52	0.09
[VV2006] J114939.6+014624	11:49:39.60	01:46:25.5	QSO	1.362	Blue	0.34	0.04
2dFGRS TGN378Z115	11:50:23.78	-00:31:41.9	HII G	0.013	Blue	0.33	0.07
[P78] ACO 1392 C	11:50:36.30	-00:34:06.6	GinCl	0.006	Blue	0.36	0.09
SDSS J115036.42-003402.0	11:50:36.39	-00:34:02.6	Galaxy	0.006	Blue	0.34	0.08
[VV2006] J115049.2-005149	11:50:49.29	-00:51:49.1	QSO	1.354	Blue	0.21	0.05
LEDA 807513	11:51:13.02	-22:53:25.2	Galaxy	–	Red	0.18	0.29
SDSS J115129.42-000333.8	11:51:29.45	-00:03:33.6	Galaxy	0.326	Red	0.11	0.23
2dFGRS TGN242Z154	11:51:32.96	-02:22:21.9	Galaxy	0.004	Blue	0.15	0.05
LEDA 37102	11:51:33.35	-02:22:21.7	Seyfert 1	0.003	Blue	0.04	0.02
SDSS J115216.86+012327.2	11:52:16.88	01:23:27.5	Galaxy	0.304	Red	0.07	0.37
2QZ J115217.3-025303	11:52:17.34	-02:53:02.7	EmG	0.320	Blue	0.96	0.02
ESO 572-7	11:52:27.62	-20:06:14.1	GinGroup	0.005	Blue	0.79	0.03

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
						HAC	HDBSCAN
Mrk 1307	11:52:37.30	-02:28:09.4	Seyfert 1	0.004	Blue	0.14	0.05
SDSS J115237.67-022806.3	11:52:37.67	-02:28:06.8	HII G	0.004	Blue	0.16	0.05
2dFGRS TGN311Z206	11:52:47.52	-00:40:07.7	Seyfert 1	0.005	Blue	0.15	0.05
2dFGRS TGN176Z274	11:53:14.07	-03:24:32.6	Galaxy	0.004	Blue	0.62	0.07
2dFGRS TGN243Z103	11:53:28.66	-03:13:48.9	Galaxy	0.005	Blue	1.00	0.00
[VV2006] J115345.5-024320	11:53:45.44	-02:43:20.4	QSO	1.347	Blue	0.52	0.03
UM 465B	11:54:12.31	00:08:12.4	GinPair	–	Blue	0.88	0.08
SDSS J115456.54+001106.0	11:54:56.59	00:11:05.5	Galaxy	0.004	–	–	–
SDSS J115511.13+002905.1	11:55:11.19	00:29:05.5	Galaxy	0.011	Blue	0.18	0.06
SDSS J115511.67+002925.0	11:55:11.70	00:29:25.0	Galaxy	0.011	Blue	0.25	0.06
2dFGRS TGN243Z202	11:57:11.86	-02:41:12.9	Galaxy	0.005	Blue	0.35	0.08
SDSS J115712.38-024111.2	11:57:12.29	-02:41:11.3	Galaxy	0.005	Blue	0.33	0.08
ESO 572-25	11:57:28.04	-19:37:26.5	Galaxy	0.006	Blue	0.54	0.09
2QZ J115737.0-020138	11:57:37.09	-02:01:37.2	Galaxy	0.328	Blue	0.86	0.05
[VV2006] J115748.0+014320	11:57:48.02	01:43:20.9	QSO	1.364	Blue	0.56	0.05
[VV2006] J115754.2-013815	11:57:54.26	-01:38:16.0	QSO	4.380	–	–	–
LEDA 839904	11:57:56.69	-20:33:56.4	Galaxy	–	Blue	0.24	0.07
2MASX J11580803-1753363	11:58:08.00	-17:53:36.2	Galaxy	0.008	Blue	0.78	0.05
6dFGS gJ115823.8-193103	11:58:23.80	-19:31:03.2	Galaxy	0.005	Blue	0.55	0.08
ESO 572-34	11:58:58.18	-19:01:47.7	EmG	0.004	Blue	0.34	0.08
GAMA 137854	11:59:23.49	-01:43:22.3	Galaxy	0.304	Blue	0.36	0.11
SN 1996W	11:59:28.93	-19:15:22.8	SN	–	Blue	0.86	0.05
LEDA 836770	12:00:19.81	-20:48:07.5	Galaxy	–	Blue	0.79	0.05
2MASX J12002013-0106229	12:00:20.20	-01:06:23.8	Galaxy	0.005	Blue	0.91	0.04
SDSS J120021.76-024331.0	12:00:21.77	-02:43:30.9	QSO	3.248	Blue	0.82	0.04
[BKD2008] WR 14	12:00:26.30	-01:06:07.0	PartofG	0.005	Blue	0.52	0.08
[VV2006] J120038.3+011246	12:00:38.29	01:12:46.5	QSO	1.358	Blue	0.41	0.05
QSO B1158-1842	12:00:44.95	-18:59:44.5	QSO	2.453	Blue	0.33	0.05
2dFGRS TGN244Z048	12:00:47.47	-03:25:12.1	Galaxy	0.005	Blue	1.00	0.00
UGC 6998	12:00:47.72	-00:01:24.3	Galaxy	0.006	Blue	0.96	0.04
UGC 7000	12:01:10.85	-01:17:50.2	GinPair	0.005	Blue	1.00	0.00
QSO B1158+007	12:01:23.26	00:28:28.5	QSO	1.369	Blue	0.49	0.05
LEDA 802182	12:01:30.48	-23:19:06.8	Galaxy	–	Blue	0.74	0.02
[CEB2007] Cluster 2	12:01:50.41	-18:52:12.4	Cl*	–	Blue	0.55	0.08
CXOU J120150.4-185221	12:01:50.41	-18:52:19.8	HMXB	–	Blue	0.56	0.07
[ZBF2015] Arp244 82	12:01:50.49	-18:52:02.5	HII	–	Blue	0.93	0.07
[NU2000] 9 3	12:01:51.13	-18:52:28.8	Radio	–	Blue	0.86	0.03
[NU2000] 13 5	12:01:51.24	-18:51:45.3	Radio	–	Blue	0.62	0.04
[MLT2008] S2-2	12:01:51.90	-18:52:28.2	Cl*	–	Blue	0.80	0.04
[ZBF2015] Arp244 123	12:01:52.28	-18:52:19.4	HII	–	Blue	0.97	0.03
[ZBF2015] Arp244 80	12:01:52.96	-18:52:03.5	HII	–	Red	0.15	0.40
[ZBF2015] Arp244 5	12:01:52.98	-18:52:08.7	HII	–	Blue	0.98	0.02
[ZBF2015] Arp244 14	12:01:53.52	-18:51:44.2	HII	–	Blue	0.90	0.04
[WZ2002] 1	12:01:53.57	-18:53:09.0	Radio	–	Red	0.27	0.34
[BEK2006] Complex 6	12:01:54.54	-18:52:07.5	Cl*	–	Blue	0.55	0.04
[WS95] 89	12:01:54.54	-18:53:03.8	PartofG	–	Blue	0.36	0.05
[WBC2014] 180.48062-18.88025	12:01:55.35	-18:52:48.7	MolCld	0.005	Blue	0.20	0.04
[ZBF2015] Arp244 6	12:01:55.54	-18:52:22.9	HII	–	Blue	0.83	0.02
[ZFB2014] GMC 98	12:01:55.68	-18:52:14.0	MolCld	–	Blue	0.71	0.02
[WZ2002] 9	12:01:55.70	-18:52:42.8	Radio	–	Blue	0.21	0.04
[ZBF2015] Arp244 10	12:01:56.29	-18:52:38.8	HII	–	Blue	1.00	0.00
CRTS J120206.7-230305	12:02:06.75	-23:03:06.0	EB*	–	Blue	0.96	0.03
SDSS J120250.38+001931.6	12:02:50.39	00:19:31.5	Galaxy	0.333	Red	0.03	0.23
SDSS J120515.80-024222.6	12:05:15.80	-02:42:22.6	WD*	0.000	Blue	0.21	0.05
LEDA 913203	12:06:37.73	-15:17:17.2	Galaxy	–	Blue	0.97	0.03
6dFGS gJ120650.7-141256	12:06:50.66	-14:12:55.9	Galaxy	0.013	Blue	0.80	0.04
[VV2006] J120700.4+011155	12:07:00.41	01:11:56.4	QSO	1.520	Blue	0.73	0.07
SDSS J120920.53-002855.3	12:09:20.55	-00:28:55.3	QSO	3.237	Blue	0.82	0.05

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group HAC	P(Blue)	P(Red)
						HDBSCAN	HDBSCAN
SDSS J121026.38-000513.2	12:10:26.41	-00:05:13.2	Galaxy	0.310	Blue	0.69	0.04
SDSS J121043.55-003907.2	12:10:43.59	-00:39:08.5	Seyfert 1	0.331	Red	0.02	0.04
2QZ J121101.0+012024	12:11:01.05	01:20:25.0	EmG	0.333	Blue	0.95	0.05
Mrk 1313	12:12:14.73	00:04:20.6	Seyfert 1	0.008	Blue	0.72	0.08
2dFGRS TGN246Z007	12:12:15.90	-00:33:53.2	Galaxy	0.008	Blue	0.32	0.04
2MASS J12125978+0149231	12:12:59.78	01:49:23.2	EB*	–	Blue	0.38	0.14
SDSS J121304.91-003901.2	12:13:04.93	-00:39:01.2	Galaxy	0.189	Red	0.04	0.41
2dFGRS TGN247Z167	12:13:38.79	-01:17:36.3	Galaxy	0.008	Blue	0.84	0.04
6dFGS gJ121348.2-143140	12:13:48.16	-14:31:39.8	Galaxy	0.330	Blue	0.39	0.04
SDSS J121435.24-015924.4	12:14:35.26	-01:59:24.4	QSO	3.233	Blue	0.79	0.09
2QZ J121446.6-013339	12:14:46.69	-01:33:38.0	Galaxy	0.331	Red	0.00	1.00
[VV2006] J121515.2-013542	12:15:15.23	-01:35:40.8	QSO	1.350	Blue	0.59	0.05
2QZ J121539.4-022149	12:15:39.47	-02:21:47.2	Galaxy	0.319	Blue	0.80	0.06
2QZ J121607.5-022559	12:16:07.54	-02:25:57.6	Galaxy	0.324	Blue	1.00	0.00
SDSS J121759.99+002558.1	12:18:00.05	00:25:57.7	Galaxy	0.003	–	–	–
2dFGRS TGN181Z079	12:18:07.07	-03:06:28.8	Galaxy	0.001	Red	0.05	0.47
LEDA 927634	12:18:19.06	-14:12:19.9	Galaxy	–	Blue	0.97	0.03
QSO B1216+0216	12:18:55.80	02:00:02.1	QSO	0.327	Blue	0.18	0.06
[VV2006] J121942.5-001821	12:19:42.47	-00:18:21.4	QSO	1.337	Blue	0.49	0.05
2dFGRS TGN385Z034	12:19:53.13	01:46:24.0	HII G	0.007	Blue	0.89	0.06
SDSS J122003.72+010632.0	12:20:03.73	01:06:32.4	Galaxy	0.315	Red	0.04	0.13
2dFGRS TGN385Z025	12:20:11.53	01:57:31.1	LSB G	0.007	Blue	0.33	0.08
2dFGRS TGN181Z173	12:20:28.80	-01:50:21.0	Galaxy	0.008	Blue	0.82	0.06
LEDA 1143004	12:20:30.39	-00:27:03.0	Galaxy	0.007	–	–	–
[VV2006] J122130.9+010727	12:21:30.97	01:07:28.1	QSO	1.370	Blue	0.68	0.05
Gaia DR2 3521773745637847552	12:21:34.41	-14:57:50.5	Star	–	Blue	0.05	0.02
LEDA 3294456	12:21:55.83	-01:35:36.0	Galaxy	0.006	Blue	0.55	0.09
Gaia DR2 3521681421020417408	12:22:39.34	-15:29:12.1	Star	–	Blue	0.26	0.06
SDSS J122322.39-000801.6	12:23:22.39	-00:08:01.7	Galaxy	0.318	Blue	0.70	0.03
MCG+00-32-004	12:24:12.47	00:34:01.0	Galaxy	0.007	Blue	0.96	0.04
2SLAQ J122421.12+002354.1	12:24:21.13	00:23:54.4	QSO	0.334	Blue	0.33	0.07
NGC 4385	12:25:42.74	00:34:21.9	AGN	0.007	Blue	0.62	0.04
2QZ J122547.3-012007	12:25:47.38	-01:20:05.7	Galaxy	0.317	Blue	0.73	0.03
SHOC 373a	12:26:22.64	-01:15:17.3	HII G	0.007	Blue	0.21	0.06
SHOC 373b	12:26:22.73	-01:15:12.3	HII G	0.007	Blue	0.12	0.05
[VV2006] J122625.7+011604	12:26:25.67	01:16:04.6	QSO	2.478	Blue	0.36	0.06
2SLAQ J122641.43-002005.1	12:26:41.45	-00:20:05.1	Seyfert 1	0.353	Blue	0.37	0.04
MCG+00-32-013	12:27:04.56	-00:54:22.0	GinPair	0.007	Blue	0.87	0.03
[VV2006] J122707.1+010811	12:27:07.13	01:08:11.3	QSO	2.189	Blue	0.10	0.03
NAME DW 1225+0152	12:27:46.07	01:36:01.5	LSB G	0.004	Blue	0.36	0.08
2dFGRS TGN387Z059	12:28:15.92	01:49:43.7	Galaxy	0.003	Blue	1.00	0.00
2QZ J122851.2-022630	12:28:51.34	-02:26:29.2	EmG	0.331	Blue	0.93	0.07
2dFGRS TGN250Z094	12:29:14.65	-01:21:55.2	Galaxy	0.007	Blue	0.51	0.09
2dFGRS TGN250Z087	12:29:46.33	-01:17:42.0	LSB G	0.008	Blue	0.44	0.09
2dFGRS TGN321Z099	12:29:58.88	00:01:37.9	RadioG	0.008	Red	0.05	0.41
2dFGRS TGN388Z078	12:30:54.31	00:57:50.5	Galaxy	0.008	Blue	0.41	0.05
[BKD2008] WR 29	12:31:48.01	-02:58:13.0	PartofG	0.008	Blue	0.82	0.03
2QZ J123202.6+003124	12:32:02.70	00:31:24.7	EmG	0.329	Blue	0.51	0.05
2dFGRS TGN251Z016	12:32:23.64	-01:44:24.3	HII G	0.007	Blue	0.79	0.04
GALEX 2414740977348515009	12:32:36.17	-03:18:39.4	Blue	–	Blue	0.56	0.05
MGC 34804	12:32:41.58	00:03:26.4	Star	–	Blue	0.48	0.08
[DCD2013] CSS J123702.3-151643	12:37:02.41	-15:16:43.5	RRLyr	–	Blue	0.93	0.01
Gaia DR2 3527007524064861312	12:39:19.45	-14:47:31.1	Star	–	–	–	–
2MASX J12442692-1252359	12:44:26.92	-12:52:35.9	Galaxy	0.018	Blue	1.00	0.00
LEDA 924051	12:50:47.04	-14:29:01.5	Galaxy	–	Blue	1.00	0.00
IC 3834	12:51:32.37	-14:13:16.2	Seyfert 1	0.015	Blue	0.35	0.14
LEDA 932206	12:58:59.15	-13:51:42.3	Galaxy	–	Blue	0.54	0.09
CRTS SSS120721 J125901-133442	12:59:00.82	-13:34:42.0	Candidate CV*	–	Blue	0.30	0.06

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
						HAC	HDBSCAN
6dFGS gJ125904.7-144623	12:59:04.67	-14:46:23.5	Galaxy	0.016	Blue	0.76	0.05
2MASX J12593269-1514196	12:59:32.75	-15:14:19.3	Galaxy	–	Blue	0.19	0.06
LEDA 914340	13:00:03.04	-15:12:17.2	Galaxy	–	Blue	0.98	0.02
NGC 4887	13:00:39.30	-14:39:59.3	GinPair	0.009	Blue	0.97	0.03
[VV96] J130243.5-135553	13:02:43.59	-13:55:52.8	QSO	1.391	Blue	0.32	0.05
LEDA 45114	13:03:33.44	-14:19:23.1	EmObj	0.008	Blue	0.74	0.08
2MASX J13044961-1311288	13:04:49.65	-13:11:28.2	Galaxy	0.010	Blue	0.71	0.07
LCRS B130214.7-120615	13:04:52.39	-12:22:18.7	Galaxy	–	Blue	1.00	0.00
LEDA 949391	13:05:58.52	-12:40:08.5	Galaxy	–	Blue	0.59	0.08
LEDA 105081	13:10:08.77	-12:12:20.4	Galaxy	0.013	Blue	0.91	0.02
UGCA 332	13:11:58.29	-12:03:51.4	EmG	0.007	Blue	0.90	0.04
LEDA 976320	13:12:28.41	-10:35:24.4	Galaxy	–	Blue	0.88	0.08
LCRS B131057.8-121222	13:13:36.14	-12:28:15.2	EmObj	0.013	Blue	0.29	0.08
LEDA 126038	13:15:07.97	-12:31:05.1	Galaxy	0.013	Blue	1.00	0.00
LEDA 981336	13:17:40.89	-10:10:59.5	Galaxy	–	Blue	1.00	0.00
SDSS J131742.35-002015.8	13:17:42.37	-00:20:15.7	Galaxy	0.356	–	–	–
6dFGS gJ131743.9-010002	13:17:43.96	-01:00:01.1	AGN	0.004	Blue	0.05	0.02
2MASX J13192221-1509232	13:19:22.29	-15:09:23.6	Galaxy	0.009	Blue	0.95	0.05
MCG-02-34-029	13:19:42.72	-11:28:28.5	GinGroup	0.009	Blue	0.97	0.03
2SLAQ J131957.59-003446.7	13:19:57.60	-00:34:46.6	Star	–	Blue	0.96	0.04
QSO B1317-122	13:19:59.20	-12:29:16.8	QSO	0.329	Blue	0.28	0.06
NGC 5088	13:20:20.33	-12:34:18.1	Galaxy	0.005	Blue	0.87	0.04
SDSS J132023.46-004730.9	13:20:23.47	-00:47:30.8	QSO	3.255	–	–	–
6dFGS gJ132134.7-151056	13:21:34.68	-15:10:55.5	Galaxy	0.009	Blue	1.00	0.00
6dFGS gJ132137.8-145120	13:21:37.83	-14:51:19.7	Galaxy	0.009	Blue	0.98	0.01
2dFGRS TGN263Z056	13:22:17.11	-00:32:54.4	EmG	0.018	–	–	–
[SHM2017] J200.93368-12.05326	13:23:44.09	-12:03:11.8	RRLyr	–	Blue	0.96	0.03
LEDA 46982	13:25:48.67	-11:36:37.8	BlueCompG	0.004	Blue	0.04	0.02
LEDA 991902	13:26:05.10	-09:22:12.6	Galaxy	–	Blue	0.55	0.09
ATO J201.6140-13.6964	13:26:27.37	-13:41:47.0	V*	–	Red	0.01	0.49
BPS CS 22889-0007	13:31:59.47	-09:53:02.6	RRLyr	0.001	Blue	1.00	0.00
NVSS J133618-072251	13:36:18.64	-07:22:51.8	Radio	–	Blue	0.43	0.05
LCRS B133356.3-061328	13:36:33.05	-06:28:45.2	Galaxy	–	Red	0.04	0.90
GALEX 2697385722761974216	13:39:09.19	-08:19:40.8	Blue	–	Blue	0.64	0.05
LEDA 1025584	13:41:04.99	-07:01:05.8	Galaxy	–	Blue	0.63	0.03
[DCD2013] CSS J134330.9-151858	13:43:31.01	-15:18:58.9	RRLyr	–	Blue	1.00	0.00
V* HS Vir	13:43:38.44	-08:14:03.7	CataclyV*	–	Blue	0.89	0.07
Gaia 18dwd	13:46:39.20	-09:38:36.0	Transient	–	Blue	0.85	0.05
GALEX 2697315366902694354	13:47:49.82	-04:10:10.6	Blue	–	Blue	0.58	0.05
LEDA 1033986	13:48:51.50	-06:20:26.3	Galaxy	–	Blue	0.24	0.05
2dFGRS TGN202Z201	13:49:42.20	-02:11:59.1	Galaxy	0.011	Blue	0.78	0.03
GALEX 2699039396840082228	13:50:33.33	-12:16:42.9	Blue	–	Blue	0.55	0.07
6dFGS gJ135123.7-060412	13:51:23.70	-06:04:11.7	Galaxy	0.010	Blue	0.98	0.02
2dFGRS TGN202Z136	13:51:35.65	-02:33:15.0	Galaxy	0.015	Blue	0.97	0.03
SDSSCGB 287.4	13:52:03.90	-02:07:22.3	Galaxy	–	Blue	0.49	0.06
SDSSCGB 287.2	13:52:04.24	-02:07:48.9	Galaxy	0.015	Blue	0.42	0.05
LEDA 126156	13:54:11.29	-03:26:27.2	Galaxy	0.014	Blue	0.48	0.05
LEDA 126156	13:54:11.40	-03:26:27.0	Galaxy	0.014	Blue	0.92	0.05
QSO B1352-104	13:54:46.53	-10:41:02.6	QSO	0.330	Blue	0.25	0.04
VV 99b	13:55:33.98	-05:58:17.0	GinPair	–	Blue	0.87	0.07
2dFGRS TGN141Z158	13:55:37.68	-04:11:43.8	Galaxy	0.012	Blue	0.20	0.06
VV 100a	13:55:45.46	-06:00:15.9	Galaxy	–	Blue	0.46	0.05
VV 100d	13:55:46.68	-06:00:40.8	Galaxy	–	Blue	0.34	0.09
[VV2006] J135602.8-022624	13:56:02.79	-02:26:23.3	QSO	1.373	–	–	–
2dFGRS TGN203Z232	13:56:53.48	-02:38:52.2	Galaxy	0.013	Blue	0.59	0.06
2dFGRS TGN142Z266	13:58:08.62	-04:08:43.6	Galaxy	0.015	Blue	0.77	0.02
SDSSCGB 16922.1	13:58:41.49	-01:31:15.7	Galaxy	–	Red	0.24	0.42
LEDA 1020082	14:02:45.08	-07:22:25.8	Galaxy	–	Blue	0.53	0.08

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group HAC	P(Blue)	P(Red)
						HDBSCAN	HDBSCAN
2MASS J14265388+0525172	14:26:53.89	05:25:17.4	QSO	0.323	Blue	0.28	0.07
UGC 9252	14:27:10.82	05:07:59.3	Galaxy	0.005	Blue	0.99	0.01
[LAM2019] J1428+0500 B	14:28:55.39	05:00:21.9	Possible lensImage	–	Blue	0.39	0.06
GALEX 2429518413625830432	14:28:55.46	05:00:19.9	Blue	–	Blue	0.51	0.05
DES J142943.42+052122.7	14:29:43.44	05:21:22.9	GinCl	–	Red	0.20	0.28
SDSS J142958.66+044611.0	14:29:58.65	04:46:11.3	BCIG	0.456	Red	0.05	0.11
LEDA 1290447	14:41:28.04	05:51:52.3	Galaxy	0.027	Blue	1.00	0.00
SDSS J145344.51+045645.8	14:53:44.52	04:56:46.0	QSO	3.328	Blue	0.73	0.04
SDSSCGB 43444.3	14:55:33.70	04:46:43.2	AGN	0.334	Red	0.21	0.27
CRTS J145640.1-211601	14:56:40.15	-21:16:01.1	EB*	–	Red	0.31	0.29
ATO J299.9677+00.4034	19:59:52.25	00:24:12.7	Candidate EB*	–	Red	0.17	0.40
SDSS J200143.74+004918.4	20:01:43.73	00:49:18.4	QSO	–	Blue	0.23	0.04
SDSS J200432.38+001041.3	20:04:32.39	00:10:41.4	low-mass*	–	–	–	–
[SHM2017] J302.70083-00.21773	20:10:48.20	-00:13:03.9	RRLyr	–	Blue	1.00	0.00
ATO J304.2000+00.9027	20:16:48.00	00:54:09.9	Candidate EB*	–	Blue	0.45	0.11
2MASS J20223549-0040099	20:22:35.51	-00:40:09.9	RRLyr	–	Blue	0.94	0.02
2MASS J20223779-0002504	20:22:37.80	-00:02:50.5	RRLyr	–	Blue	0.97	0.03
UGC 11566	20:28:12.02	00:17:18.2	Galaxy	0.006	Blue	0.86	0.06
SDSS J202906.80+005453.5	20:29:06.81	00:54:53.6	QSO	–	Blue	0.97	0.03
2SLAQ J204340.03+002853.4	20:43:40.04	00:28:53.6	Seyfert 1	0.317	Blue	0.32	0.05
SDSS J204626.10+002337.7	20:46:26.11	00:23:37.8	QSO	0.332	Red	0.12	0.25
2SLAQ J204720.76+000007.7	20:47:20.76	00:00:07.7	CataclyV*	0.001	Blue	0.43	0.04
2SLAQ J204910.96+001557.2	20:49:10.95	00:15:57.5	Seyfert 1	0.363	Blue	0.25	0.05
[VV2006] J204956.6-001201	20:49:56.62	-00:12:01.7	QSO	0.369	Blue	0.28	0.06
[VV2006] J205316.7+005920	20:53:16.77	00:59:21.1	QSO	4.299	–	–	–
2SLAQ J205352.03-001601.5	20:53:52.04	-00:16:01.5	QSO	0.363	Blue	0.42	0.07
2SLAQ J205614.55-004050.9	20:56:14.55	-00:40:50.6	Star	–	–	–	–
2SLAQ J205712.69+001211.3	20:57:12.69	00:12:11.4	QSO	0.335	Blue	0.31	0.07
CRTS J205732.3+004358	20:57:32.38	00:43:57.5	EB*	–	Red	0.21	0.37
SDSS J205740.76+005418.5	20:57:40.75	00:54:19.0	QSO	0.332	Red	0.08	0.43
Gaia DR2 6794425304909258752	20:58:06.45	-30:08:18.1	Candidate WD*	–	Blue	0.26	0.05
LEDA 687146	20:58:24.54	-32:43:22.5	Galaxy	–	Blue	1.00	0.00
2MASX J20584976-4420243	20:58:49.69	-44:20:24.7	Galaxy	0.018	Blue	0.98	0.02
6dFGS gJ205957.5-213935	20:59:57.53	-21:39:34.9	Galaxy	-0.001	Blue	0.74	0.06
SDSS J210014.12+004446.0	21:00:14.11	00:44:45.9	CataclyV*	0.000	Blue	0.35	0.06
SDSSCGB 52599.6	21:01:55.95	-00:31:24.9	Galaxy	–	Red	0.09	0.37
LEDA 598660	21:01:56.43	-39:23:40.3	Galaxy	–	Blue	0.72	0.03
QSO B2059-330	21:02:41.71	-32:52:44.1	QSO	3.280	Blue	0.49	0.05
LEDA 528866	21:03:02.23	-45:14:41.1	Galaxy	–	Blue	0.86	0.04
Gaia DR2 6808104805812408064	21:03:56.66	-21:47:27.1	Star	–	Blue	0.91	0.04
ESO 286-33	21:04:08.52	-43:32:03.2	IG	0.017	Blue	0.70	0.05
ESO 286-35	21:04:11.17	-43:35:33.8	GinGroup	0.017	Blue	0.35	0.10
[GPM2009] J2104-0035 1	21:04:55.31	-00:35:21.8	EmG	0.005	Blue	0.42	0.09
LEDA 520361	21:05:20.69	-45:59:19.3	Galaxy	–	Blue	0.65	0.07
ESO 286-44	21:05:38.68	-42:46:52.4	Galaxy	0.008	Blue	0.42	0.09
PN G006.0-41.9	21:05:53.57	-37:08:40.4	PN	–	Blue	0.02	0.01
EC 21035-4032	21:06:48.02	-40:20:03.7	Star	–	Blue	0.24	0.07
2MASX J21071198-4733258	21:07:11.98	-47:33:25.2	GinPair	–	Blue	0.91	0.02
2MASX J21071385-4733258	21:07:13.86	-47:33:25.3	GinPair	0.015	Blue	1.00	0.00
[GPM2009] J2112-0016 1	21:12:00.92	-00:16:49.2	EmG	0.012	Blue	0.97	0.03
2MASS J21122459-4128534	21:12:24.59	-41:28:53.3	AGN	0.349	Blue	0.39	0.03
CRTS J211328.1+000332	21:13:28.17	00:03:32.6	EB*	–	Blue	1.00	0.00
ESO 402-20	21:15:19.13	-33:13:34.6	Galaxy	0.018	Blue	0.93	0.07
1RXS J211805.2-341343	21:18:04.28	-34:13:43.3	CataclyV*	–	Blue	0.80	0.03
AT20G J212302-291504	21:23:02.82	-29:15:04.0	Radio(cm)	–	Blue	0.31	0.05
CRTS CSS120613 J212655-012054	21:26:54.54	-01:20:54.1	Candidate CV*	–	Blue	0.20	0.04
LEDA 710984	21:27:08.26	-30:57:08.4	Galaxy	–	Blue	1.00	0.00
2SLAQ J212954.20-010323.3	21:29:54.22	-01:03:23.3	EmG	0.335	Blue	0.71	0.04

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group HAC	P(Blue)	P(Red)
						HDBSCAN	HDBSCAN
LBQS 2128-4555	21:31:29.53	-45:41:50.5	QSO	0.623	Blue	0.37	0.05
2SLAQ J213242.28-010309.0	21:32:42.29	-01:03:09.2	Galaxy	–	Blue	0.89	0.05
2SLAQ J213245.24+000146.4	21:32:45.26	00:01:46.8	Seyfert 1	0.234	Blue	0.17	0.05
2MASS J21333817+0126291	21:33:38.14	01:26:29.0	QSO	1.004	Blue	0.95	0.05
SDSS J213455.08+001056.9	21:34:55.09	00:10:56.8	QSO	3.289	Blue	0.78	0.03
WISEA J213649.75-012852.2	21:36:49.75	-01:28:52.2	QSO	3.280	Blue	1.00	0.00
QSO B2134-453	21:38:07.49	-45:08:18.0	QSO	4.360	–	–	–
2MASS J21381896+0112224	21:38:18.96	01:12:22.5	Seyfert 1	0.344	Blue	0.27	0.07
CRTS J213937.6-023913	21:39:37.58	-02:39:13.0	Candidate CV*	–	Blue	0.22	0.05
2SLAQ J214106.46+004733.3	21:41:06.44	00:47:33.5		2.452	Blue	0.28	0.05
SDSSCGB 15831.2	21:41:42.73	00:45:34.9	Galaxy	–	Blue	0.96	0.01
SDSS J214155.04-011734.3	21:41:55.04	-01:17:34.2	QSO	3.286	Blue	0.99	0.01
SN 2017hxv	21:44:22.94	-29:54:59.0	SN	0.019	Red	0.07	0.13
2SLAQ J214455.94+002305.8	21:44:55.92	00:23:06.1	EmG	0.330	Blue	0.83	0.05
6dFGS gJ214540.0-291937	21:45:40.01	-29:19:36.9	Galaxy	0.341	Red	0.09	0.14
2SLAQ J214830.60-004752.6	21:48:30.61	-00:47:52.5	EmG	0.332	–	–	–
SDSS J215002.69+011343.8	21:50:02.70	01:13:43.8	QSO	3.267	Blue	0.97	0.03
2SLAQ J215010.52-001000.6	21:50:10.53	-00:10:00.6	QSO	0.335	Blue	0.30	0.07
2dFGRS TGS406Z223	21:53:05.55	-31:28:17.9	Galaxy	0.019	Blue	0.55	0.10
2MASX J21541799+0056318	21:54:18.00	00:56:31.9	Galaxy	0.010	Blue	1.00	0.00
LEDA 214792	21:56:13.83	-01:09:42.8	Galaxy	–	Blue	0.85	0.04
SDSSCGB 16345.1	21:56:19.81	-01:10:03.7	Galaxy	0.016	Blue	0.48	0.07
2dFGRS TGS059Z257	21:57:20.89	-25:08:02.4	Galaxy	0.009	Blue	0.96	0.04
LAMOST J215805.07-030748.5	21:58:05.06	-03:07:48.2	Galaxy	0.016	Blue	0.93	0.05
SDSS J215824.23-004413.7	21:58:24.28	-00:44:13.7	HII G	0.016	Blue	0.94	0.03
SDSS J215902.90-003318.4	21:59:02.89	-00:33:18.0	Galaxy	–	Blue	0.71	0.08
LEDA 214793	21:59:03.11	-01:57:18.3	Galaxy	–	Blue	0.67	0.05
2dFGRS TGS114Z230	22:02:07.08	-26:26:38.0	Galaxy	0.309	Blue	0.36	0.07
PB 5049	22:03:15.14	01:17:21.0	Star	–	Blue	0.03	0.02
2dFGRS TGS115Z105	22:04:53.68	-25:03:05.2	Galaxy	0.009	Blue	0.96	0.04
2SLAQ J220529.34-003110.6	22:05:29.34	-00:31:10.7	QSO	2.454	Blue	0.41	0.03
NGC 7204	22:06:55.31	-31:03:10.6	IG	0.009	Red	0.14	0.19
2dFGRS TGS251Z159	22:07:34.55	-28:39:29.3	Galaxy	0.018	Blue	0.36	0.09
NGC 7208	22:08:24.43	-29:03:03.6	GinGroup	0.009	Blue	0.96	0.04
2dFGRS TGS333Z140	22:08:51.96	-30:38:58.8		0.008	–	–	–
[VV2006] J220852.0-010603	22:08:51.97	-01:06:03.7	QSO	0.351	Blue	0.30	0.06
MCG-05-52-033a	22:08:55.90	-27:13:22.0	GinPair	0.009	Blue	0.78	0.06
2dFGRS TGS061Z180	22:09:19.05	-24:07:12.4	QSO	0.320	Red	0.09	0.11
2dFGRS TGS116Z088	22:09:22.91	-25:25:04.6	Galaxy	0.008	Blue	0.59	0.08
2QZ J220948.6-301357	22:09:48.63	-30:13:55.8	WD*	–	Blue	0.13	0.05
SDSSCGB 41857.2	22:09:51.35	01:09:00.0	Galaxy	0.149	Red	0.20	0.25
SDSS J220954.57-012717.6	22:09:54.57	-01:27:17.6	QSO	3.296	Blue	0.58	0.03
2QZ J221000.7-311400	22:10:00.75	-31:14:00.0	EmG	0.328	Blue	0.64	0.05
2dFGRS TGS116Z082	22:10:03.74	-25:20:07.3	Galaxy	0.009	Blue	0.99	0.01
2QZ J221005.7-275439	22:10:05.76	-27:54:38.7	Galaxy	0.330	Blue	0.92	0.04
6dFGS gJ221058.1-250431	22:10:58.01	-25:04:31.1	Galaxy	0.017	Blue	0.81	0.03
2QZ J221058.3-273930	22:10:58.33	-27:39:29.4	Galaxy	0.313	Blue	0.49	0.05
LSQ 12dwl	22:12:41.57	00:30:43.1	SN	0.014	Blue	0.33	0.07
[VV2006] J221335.7-282542	22:13:35.65	-28:25:41.7	QSO	2.469	Blue	0.25	0.05
2dFGRS TGS175Z009	22:14:02.86	-27:32:21.4	Galaxy	0.014	Blue	1.00	0.00
NGC 7229	22:14:03.23	-29:22:57.8	EmG	0.014	Blue	0.76	0.05
2dFGRS TGS254Z244	22:14:05.10	-29:22:52.9	Galaxy	0.015	Blue	0.95	0.05
ESO 467-25	22:14:24.23	-29:58:51.5	EmG	0.015	Blue	0.75	0.03
2dFGRS TGS254Z138	22:14:41.93	-28:26:39.6	Galaxy	0.012	Blue	1.00	0.00
2dFGRS TGS334Z074	22:14:47.16	-29:41:12.4	Galaxy	0.006	Blue	0.76	0.01
2QZ J221517.1-285358	22:15:17.10	-28:53:57.5	EmG	0.312	Blue	0.73	0.07
[VV2006] J221532.6-281805	22:15:32.58	-28:18:03.9	QSO	1.330	Blue	0.68	0.04
2QZ J221630.0-290054	22:16:30.07	-29:00:53.3	EmG	0.330	Blue	0.10	0.06

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
						HAC	HDBSCAN
6dFGS gJ221706.5-303447	22:17:06.52	-30:34:46.1	Galaxy	0.337	Red	0.20	0.16
[VV2006] J221722.5+010436	22:17:22.44	01:04:36.3	QSO	1.403	Blue	0.06	0.03
2dFGRS TGS176Z011	22:17:41.15	-27:21:54.6	Galaxy	0.009	Blue	0.73	0.04
SDSS J221813.90+001625.1	22:18:13.91	00:16:25.3	Galaxy	0.333	–	–	–
LEDA 191887	22:18:15.03	01:15:16.9	Galaxy	0.013	Blue	0.79	0.05
SDSS J221817.26+003623.6	22:18:17.26	00:36:23.7	AGN	0.331	Red	0.05	0.12
2QZ J221819.4-271544	22:18:19.39	-27:15:44.2	Seyfert 1	0.355	Blue	0.27	0.06
2SLAQ J221846.76-011119.0	22:18:46.74	-01:11:18.8	Galaxy	–	Blue	0.79	0.03
2QZ J221925.9-305108	22:19:25.92	-30:51:07.7	EmG	0.307	Blue	0.66	0.04
SDSSCGB 28259.2	22:19:44.75	-00:14:40.0	Galaxy	–	Red	0.11	0.26
2QZ J221945.1-293414	22:19:45.08	-29:34:13.4	EmG	0.343	Blue	0.96	0.04
[DD2013] W4+2-1 115196	22:19:53.86	00:29:04.8	Galaxy	0.087	Blue	0.38	0.07
2SLAQ J222021.37+004040.2	22:20:21.36	00:40:40.5	Star	–	Blue	0.34	0.04
2QZ J222113.6-280421	22:21:13.62	-28:04:20.9	Seyfert 1	0.332	Red	0.16	0.12
2dFGRS TGS337Z130	22:22:54.92	-30:42:28.4	Galaxy	0.014	Blue	0.97	0.03
6dFGS gJ222313.7-285844	22:23:13.70	-28:58:44.6	Galaxy	0.006	Blue	0.28	0.07
2QZ J222336.0-283140	22:23:36.03	-28:31:39.6	EmG	0.333	Blue	0.52	0.04
2SLAQ J222403.36-005724.2	22:24:03.35	-00:57:24.2	QSO	0.313	Blue	0.35	0.07
2QZ J222416.2-292421	22:24:16.26	-29:24:21.7	Candidate CV*	–	Blue	0.36	0.07
2dFGRS TGS338Z083	22:27:38.90	-31:08:10.3	Galaxy	0.015	Blue	0.80	0.06
2SLAQ J222825.11-002217.4	22:28:25.12	-00:22:17.2	Galaxy	–	Blue	0.30	0.13
LEDA 711478	22:28:47.80	-30:54:43.2	Galaxy	–	Blue	0.25	0.07
2dFGRS TGS337Z266	22:28:53.67	-30:58:51.4	Galaxy	0.013	Blue	0.87	0.05
SDSS J222923.00-020042.7	22:29:23.00	-02:00:42.4	QSO	3.294	–	–	–
2SLAQ J222956.53+003126.5	22:29:56.54	00:31:26.5	QSO	1.340	Blue	0.27	0.04
2dFGRS TGS338Z165	22:30:01.84	-29:35:52.7	Galaxy	0.010	Blue	0.85	0.02
NAME Kinman Dwarf	22:30:36.83	-00:06:35.8	BlueCompG	0.006	Blue	0.35	0.09
LEDA 1149494	22:31:06.00	-00:11:43.9	Galaxy	–	Blue	1.00	0.00
2QZ J223114.0-312005	22:31:13.95	-31:20:04.4	Star	–	Blue	0.48	0.05
[VV2006] J223251.7-303250	22:32:51.74	-30:32:49.6	QSO	0.350	Blue	0.19	0.04
2QZ J223342.5-301936	22:33:42.56	-30:19:35.4	Galaxy	0.324	Blue	0.49	0.04
FASTT 1560	22:34:39.93	00:41:27.5	CataclyV*	0.001	Blue	0.33	0.05
2dFGRS TGS414Z208	22:34:56.56	-31:08:44.1	Galaxy	0.009	Blue	1.00	0.00
SDSS J223508.41-005359.3	22:35:08.42	-00:53:59.4	Seyfert 2	0.328	Red	0.23	0.37
2QZ J223532.2-294634	22:35:32.24	-29:46:33.0	EmG	0.325	Blue	0.86	0.03
2SLAQ J223543.05-005436.5	22:35:43.05	-00:54:36.6	Galaxy	–	Blue	0.68	0.03
[EKS96] NGC 7314 28	22:35:44.87	-26:02:27.4	HII	–	Blue	0.86	0.06
[EKS96] NGC 7314 43	22:35:45.40	-26:02:10.8	HII	–	Blue	0.65	0.02
[EKS96] NGC 7314 71	22:35:46.29	-26:04:29.1	HII	–	Blue	0.49	0.05
[EKS96] NGC 7314 130	22:35:48.23	-26:01:24.0	HII	–	Blue	0.93	0.07
[VV2006] J223633.5+002652	22:36:33.54	00:26:52.8	QSO	1.354	Blue	0.43	0.05
SDSS J223649.60+005413.5	22:36:49.60	00:54:13.8	QSO	3.313	–	–	–
2SLAQ J223723.84-010120.3	22:37:23.88	-01:01:19.2	Galaxy	–	Blue	0.96	0.03
PHL 354	22:38:23.25	-00:57:08.2	QSO	0.361	Blue	0.21	0.04
2SLAQ J223844.30-005655.3	22:38:44.29	-00:56:55.3	QSO	1.357	Blue	0.53	0.08
LEDA 131686	22:41:19.49	-39:58:23.3	Galaxy	–	Blue	0.73	0.05
2QZ J224149.5-301945	22:41:49.50	-30:19:44.5	Galaxy	0.322	Blue	0.40	0.05
[GDB2008] 503	22:43:16.38	-39:51:39.9	Galaxy	0.007	Blue	0.50	0.03
[GDB2008] 504	22:43:19.40	-39:52:44.4	Galaxy	0.007	Blue	0.28	0.08
[GDB2008] 509	22:43:25.11	-39:55:20.0	Galaxy	0.327	Red	0.00	1.00
SDSS J224352.11-002259.7	22:43:52.21	-00:22:59.9	Galaxy	0.010	Blue	0.96	0.04
2SLAQ J224531.20-004509.4	22:45:31.20	-00:45:09.3	QSO	1.368	Blue	0.19	0.04
SDSS J224539.94-002419.7	22:45:39.94	-00:24:19.6	QSO	3.280	Blue	0.95	0.05
2MASS J2245608+0002182	22:49:56.08	00:02:18.4	QSO	3.307	Blue	0.83	0.07
2SLAQ J225012.91-003959.0	22:50:12.92	-00:39:58.9	Galaxy	–	Blue	0.48	0.05
SDSS J225149.74-002811.7	22:51:49.75	-00:28:11.4	QSO	3.228	Blue	0.96	0.04
2QZ J225157.1-292451	22:51:57.10	-29:24:50.8	EmG	0.318	Blue	0.24	0.04
2SLAQ J225257.45+002731.5	22:52:57.44	00:27:31.6	Star	–	Blue	0.35	0.07

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	Group	P(Blue)	P(Red)
					HAC	HDBSCAN	HDBSCAN
2QZ J225352.9-300944	22:53:52.96	-30:09:43.7	Seyfert 1	0.326	Blue	0.16	0.06
[VV2006] J225411.2-312712	22:54:11.15	-31:27:11.3	QSO	1.360	Blue	0.70	0.05
SDSS J225411.96-004949.5	22:54:11.96	-00:49:49.4	QSO	3.297	Blue	0.30	0.04
2QZ J225503.9-301914	22:55:03.89	-30:19:13.3	EmG	0.326	Blue	0.99	0.01
2QZ J225908.1-312717	22:59:08.12	-31:27:16.7	Star	–	Blue	0.18	0.04
2SLAQ J230030.09-003005.9	23:00:30.09	-00:30:05.8	Galaxy	–	Blue	0.63	0.08
2SLAQ J230201.20+003047.2	23:02:01.20	00:30:47.3	QSO	1.344	Blue	0.66	0.05
[VV2006] J230235.5-285630	23:02:35.44	-28:56:29.7	QSO	0.368	Blue	0.37	0.05
2SLAQ J230316.40-001211.5	23:03:16.41	-00:12:11.4	QSO	1.516	Blue	0.55	0.08
V* HY Psc	23:03:51.63	01:06:51.4	CataclyV*	-0.000	Blue	0.41	0.05
SDSS J230428.31+005701.2	23:04:28.34	00:57:01.2	QSO	0.317	Blue	0.34	0.07
2SLAQ J230444.16-010251.7	23:04:44.16	-01:02:51.5	QSO	1.377	Blue	0.28	0.05
LEDA 1122038	23:08:10.78	-01:17:58.5	Galaxy	–	Blue	0.75	0.07
SDSS J230855.49+003705.6	23:08:55.49	00:37:05.7	QSO	1.784	Blue	1.00	0.00
ESO 469-15	23:08:55.60	-30:51:28.2	GinGroup	0.005	Blue	0.45	0.10
[VV2006] J230914.4-305913	23:09:14.31	-30:59:12.5	QSO	1.380	Blue	0.59	0.05
2MASS J23094616+0000496	23:09:46.16	00:00:49.0	Seyfert 1	0.352	Blue	0.41	0.07
[GPM2009] J2310-0109 1	23:10:41.99	-01:09:48.0	EmG	0.013	Blue	0.77	0.07
[VV2006] J231135.1-312644	23:11:35.12	-31:26:44.1	QSO	1.350	Blue	0.59	0.05
2dFGRS TGS422Z155	23:12:08.96	-31:04:13.3	Galaxy	0.165	Blue	0.25	0.14
2SLAQ J231231.36-011137.5	23:12:31.36	-01:11:37.3	QSO	1.360	Blue	0.73	0.05
SDSS J231259.07+010805.6	23:12:59.06	01:08:05.9	QSO	3.295	Blue	0.85	0.04
[VV2006] J231311.9-004538	23:13:11.91	-00:45:38.0	QSO	1.364	Blue	0.55	0.05
SDSS J231351.87-011031.9	23:13:51.86	-01:10:30.8	HII G	0.012	Blue	0.54	0.05
2MASX J23145046+0123280	23:14:50.52	01:23:26.7	LSB G	0.016	Blue	1.00	0.00
[VV2006] J231519.4-303857	23:15:19.39	-30:38:57.2	QSO	1.356	Blue	0.81	0.05
MASTER OT J231531.76-304848.3	23:15:31.78	-30:48:48.7	Transient	–	Blue	0.44	0.06
[VV2006] J231652.0+005125	23:16:52.04	00:51:25.9	QSO	3.229	Blue	0.81	0.03
3XMM J231742.5+000535	23:17:42.61	00:05:35.3	Seyfert 1	0.321	Blue	0.35	0.04
[VV2006] J231942.8-302629	23:19:42.76	-30:26:29.5	QSO	2.473	Blue	0.53	0.05
LEDA 71137	23:20:35.22	-00:52:50.8	Galaxy	0.015	Blue	0.56	0.08
2QZ J232126.5-310730	23:21:26.51	-31:07:29.5	Galaxy	0.309	Blue	1.00	0.00
[SIG2010] 389821	23:23:31.32	01:08:06.0	RRLyr	–	Blue	0.84	0.02
GALEX 2417063145906373262	23:24:20.34	-00:06:25.0	HII	–	Blue	0.06	0.03
[GPM2009] J2324-0006	23:24:21.37	-00:06:29.4	HII G	0.009	Blue	0.14	0.05
2SLAQ J232457.75+002153.2	23:24:57.75	00:21:53.4	QSO	0.345	Blue	0.12	0.05
2SLAQ J232524.40+004612.0	23:25:24.43	00:46:12.2	Galaxy	0.016	Blue	0.89	0.07
2MASS J23255145-0140232	23:25:51.48	-01:40:23.8	CataclyV*	–	Blue	0.20	0.06
[VV2006c] J232555.5-003710	23:25:55.51	-00:37:10.7	Seyfert 1	0.332	–	–	–
SDSS J232743.68-020055.8	23:27:43.70	-02:00:55.7	BlueCompG	0.018	Blue	0.08	0.03
6dFGS gJ232744.4-020047	23:27:44.38	-02:00:46.8	Galaxy	0.018	Blue	0.06	0.03
LEDA 1127711	23:28:12.30	-01:03:44.8	HII G	0.009	Blue	0.50	0.04
GD 1662	23:29:00.44	-29:46:46.0	CataclyV*	0.000	Blue	0.56	0.08
SDSS J233104.38-004237.2	23:31:04.40	-00:42:37.1	QSO	1.353	Blue	0.32	0.06
2QZ J233254.8-305844	23:32:54.78	-30:58:43.8	EmG	0.329	Blue	0.58	0.08
SDSS J233256.68+011122.9	23:32:56.68	01:11:23.1	BCIG	0.382	–	–	–
SDSS J233300.21-002030.5	23:33:00.22	-00:20:30.5	QSO	3.328	Blue	0.95	0.05
[VV2006] J233438.5+002341	23:34:38.55	00:23:41.9	QSO	1.385	Blue	0.28	0.05
2MASX J23352102+0110271	23:35:20.98	01:10:27.4	Galaxy	0.085	Red	0.24	0.20
2SLAQ J233522.69-000635.2	23:35:22.69	-00:06:35.2	QSO	1.373	Blue	0.53	0.07
LEDA 135900	23:36:46.97	00:37:23.8	LSB G	0.009	Blue	0.99	0.01
[VV2006] J233722.0+002239	23:37:22.02	00:22:39.2	QSO	1.377	Blue	0.29	0.04
2XMM J233731.7+002559	23:37:31.79	00:25:59.9	AGN	0.314	Blue	0.43	0.09
DES J233747.57+001742.6	23:37:47.57	00:17:42.8	GinCl	–	–	–	–
[VV2006] J233853.1-295457	23:38:53.09	-29:54:57.2	QSO	1.349	Blue	0.41	0.05
[VV2006] J234329.1-300200	23:43:29.16	-30:02:00.1	QSO	1.358	Blue	0.54	0.07
2SLAQ J234440.53-001205.8	23:44:40.53	-00:12:06.1	CataclyV*	–	Blue	0.25	0.04
LEDA 1109937	23:48:23.99	-01:47:31.1	Galaxy	–	Blue	1.00	0.00

Table B1: –continued

Id Object	RA	Dec	Type	Redshift	HAC	P(Blue)	P(Red)
						HDBSCAN	HDBSCAN
2dFGRS TGS356Z227	23:50:01.55	-30:11:07.1	Galaxy	0.010	Blue	0.58	0.07
2SLAQ J235115.66-000000.0	23:51:15.66	-00:00:00.0	Galaxy	–	Blue	0.15	0.03
[VV2006] J235546.2-002342	23:55:46.14	-00:23:42.8	QSO	3.245	Blue	0.93	0.01
[VV2006] J235718.4+004350	23:57:18.37	00:43:50.5	QSO	4.366	Red	0.03	0.06
SDSS J235805.25-012153.9	23:58:05.25	-01:21:53.9	QSO	1.368	Blue	0.64	0.04

## APPENDIX C: SDSS SPECTRA

Table C1: Espectra from SDSS DR16.

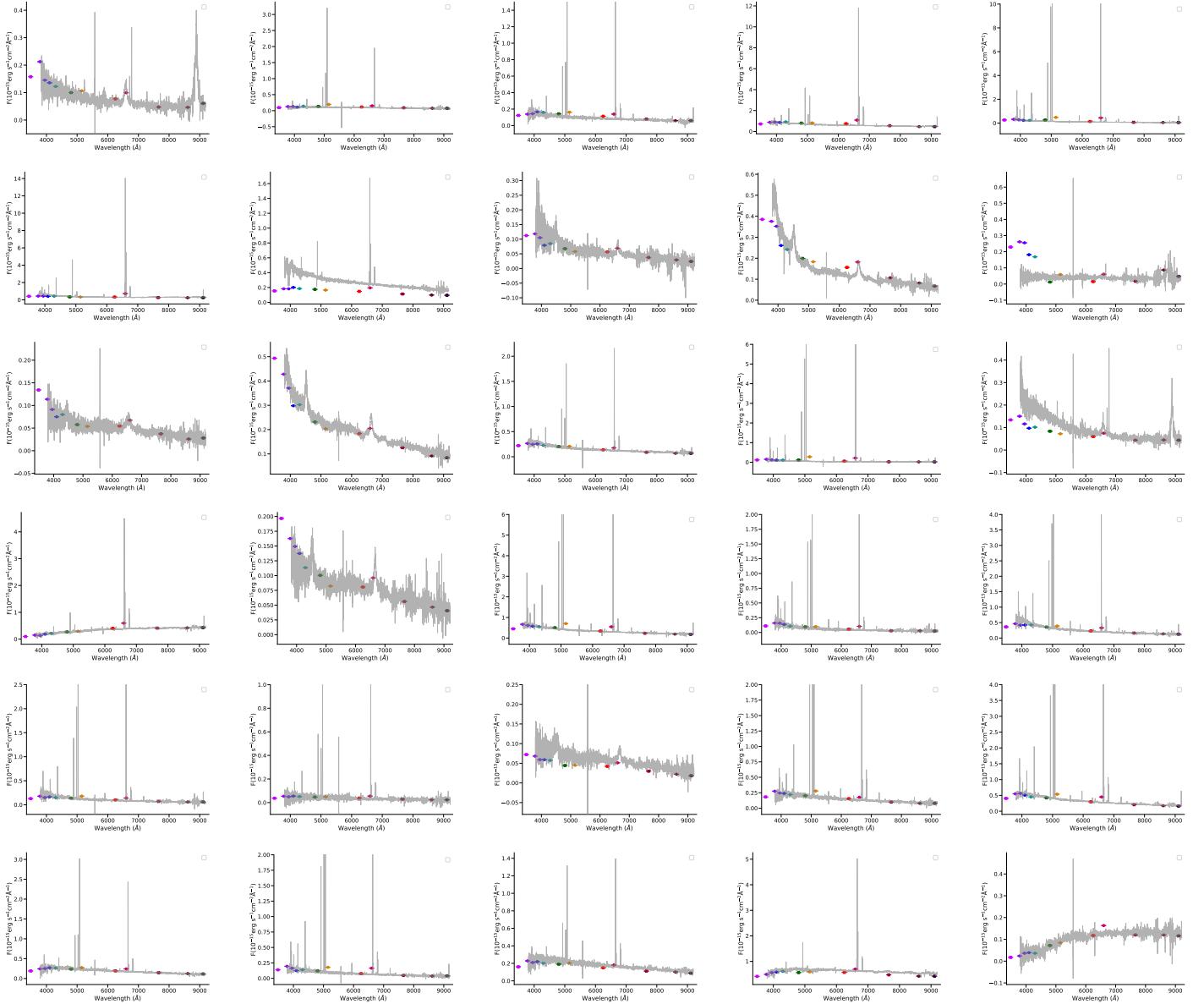


Table C1: –continued

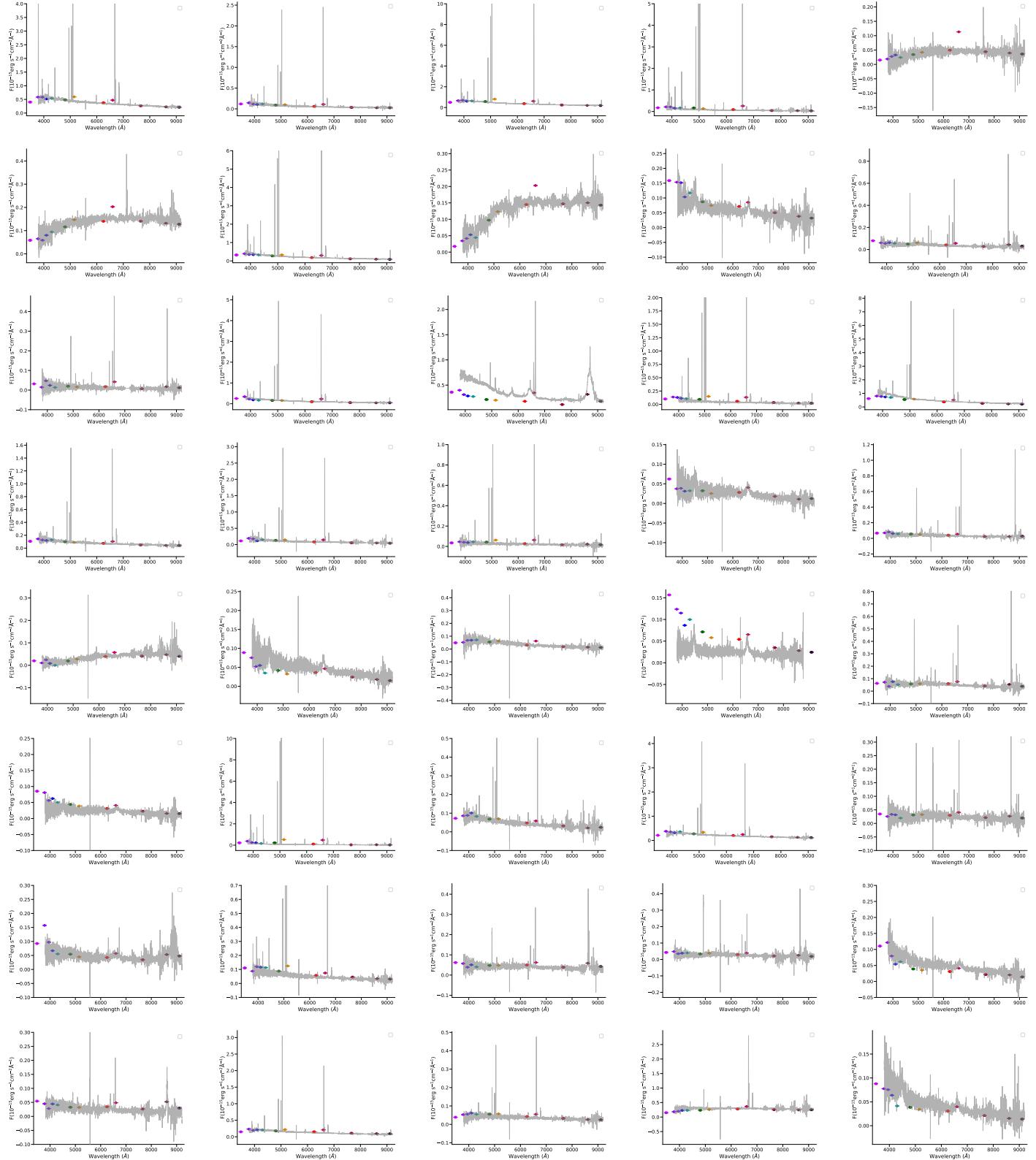


Table C1: –continued

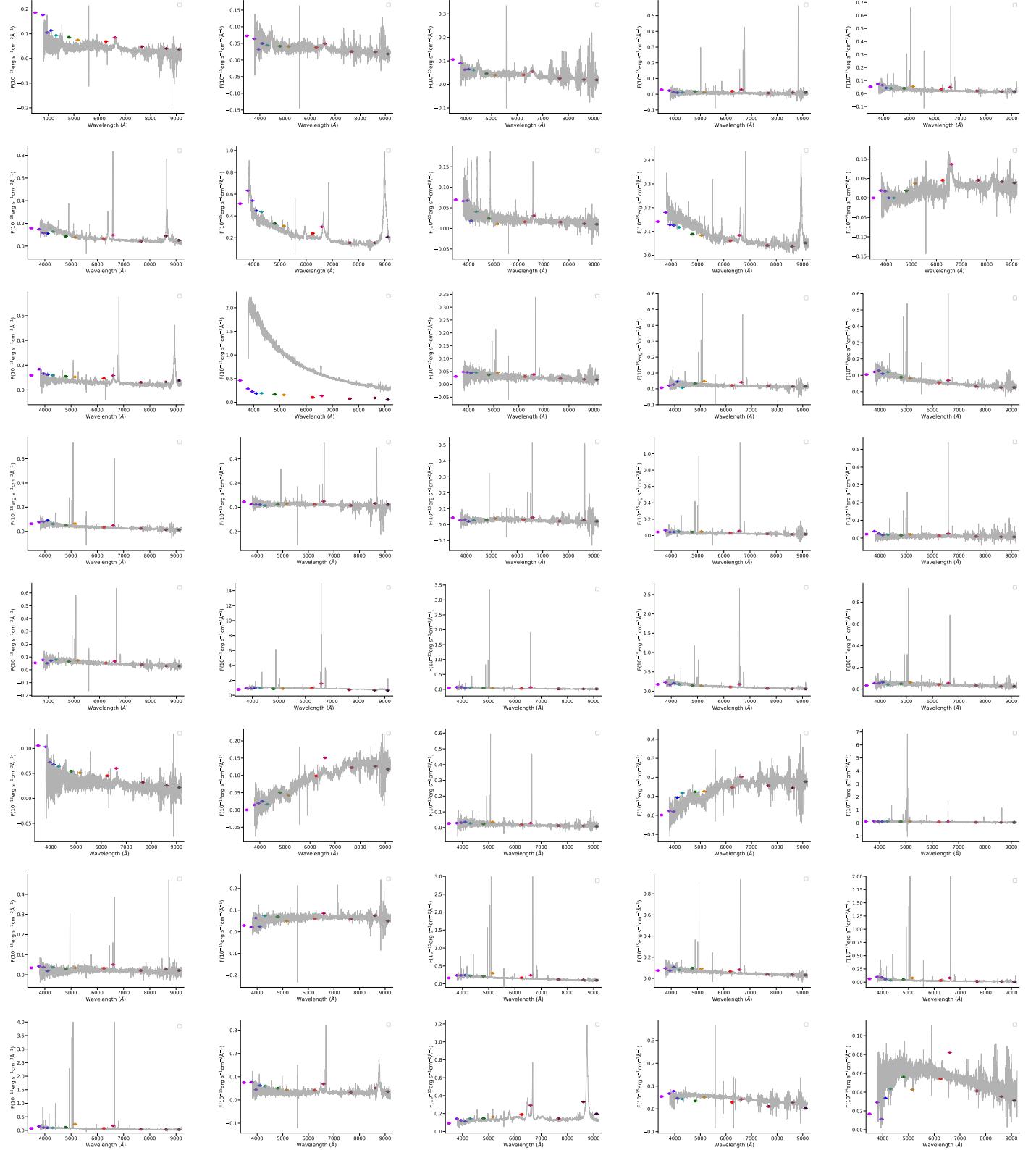


Table C1: –continued

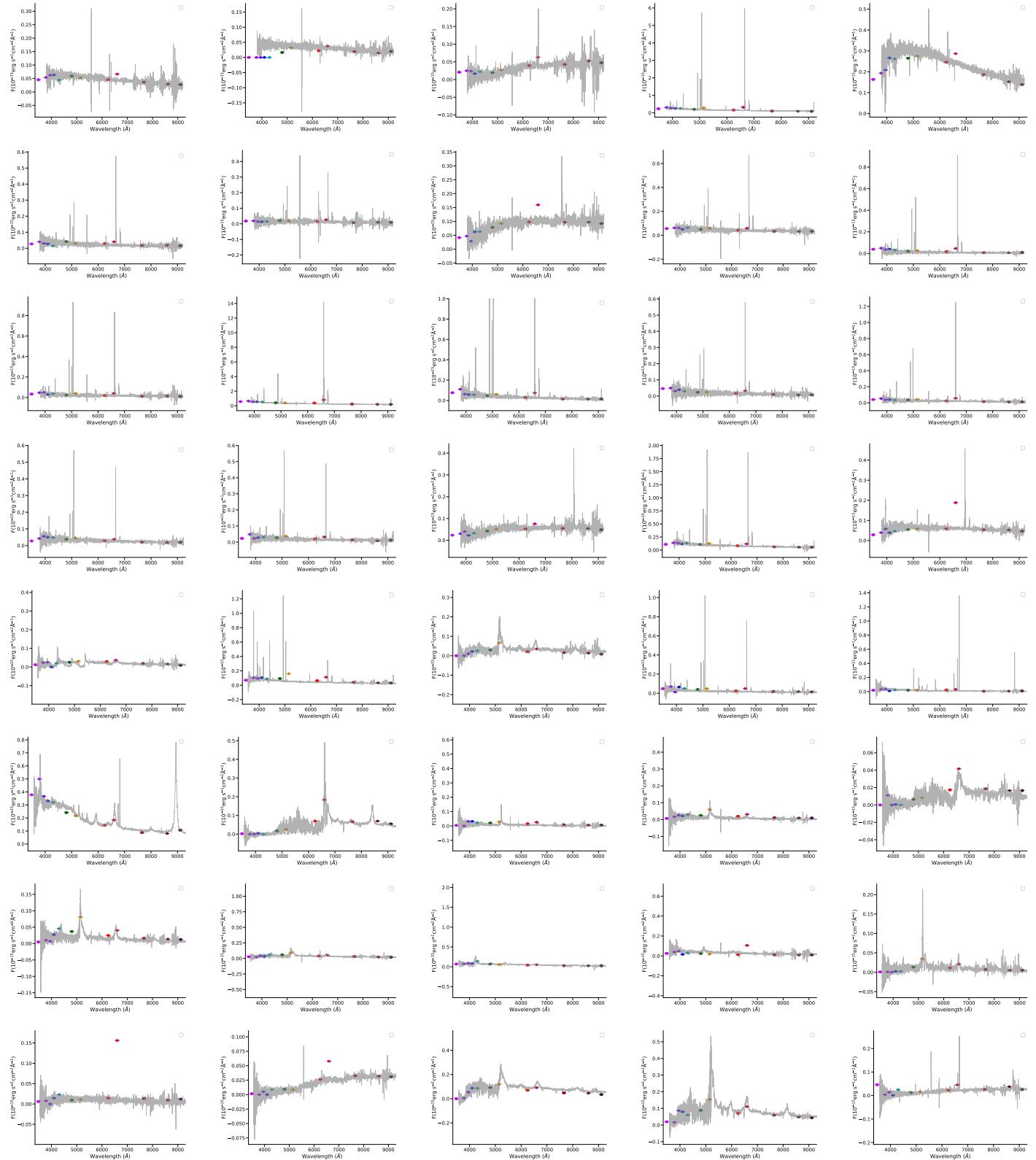


Table C1: –continued

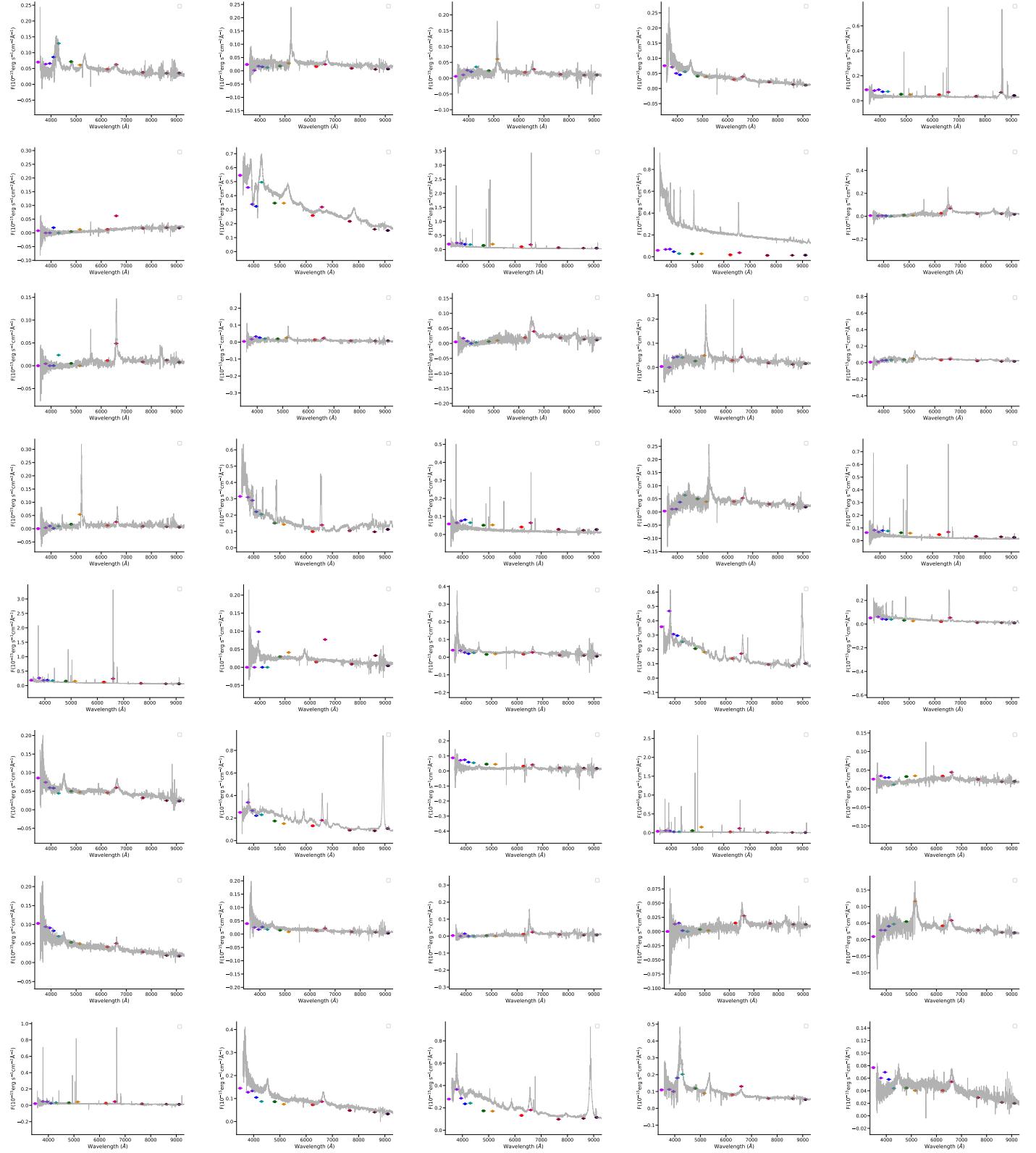


Table C1: –continued

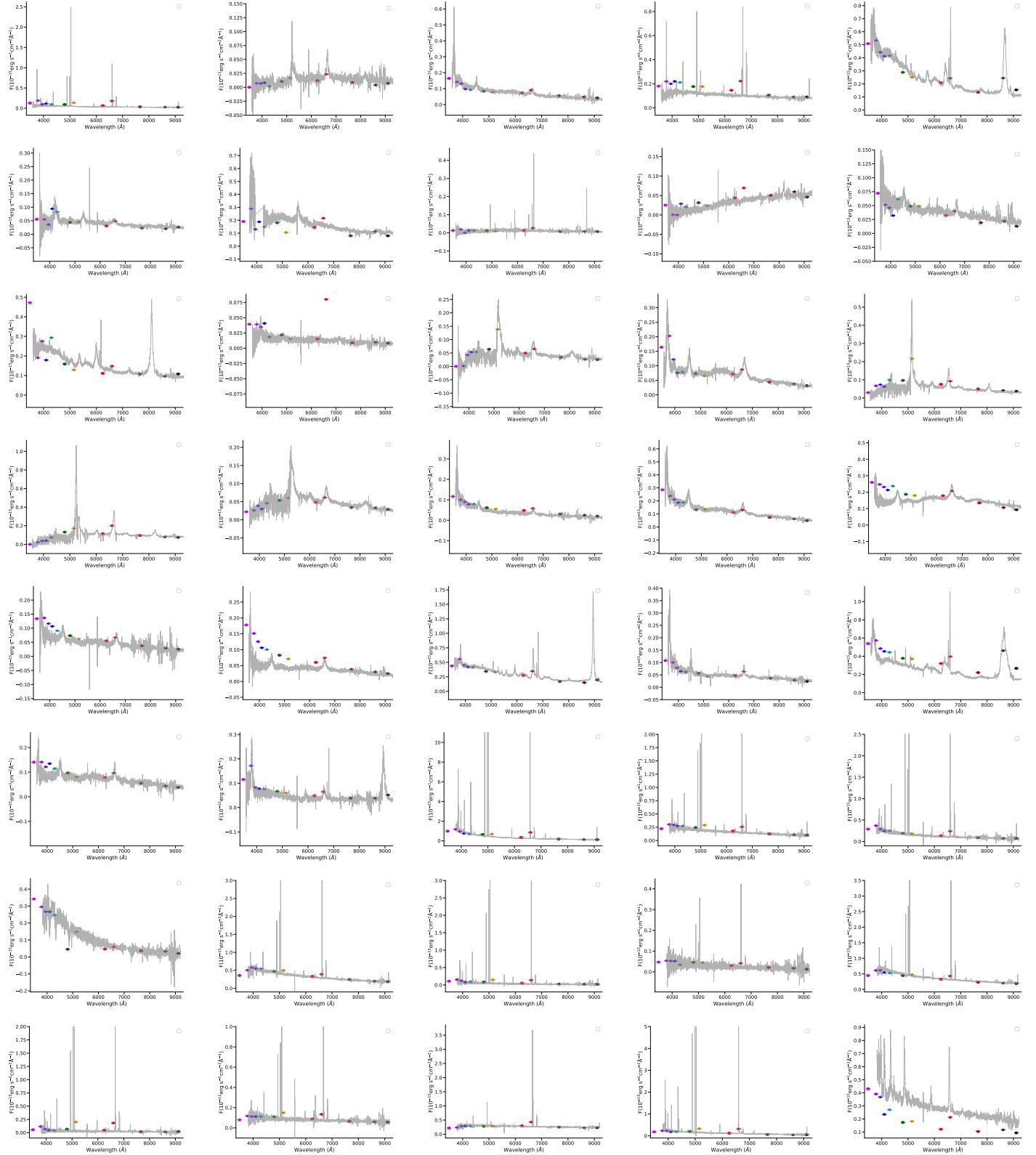


Table C1: –continued

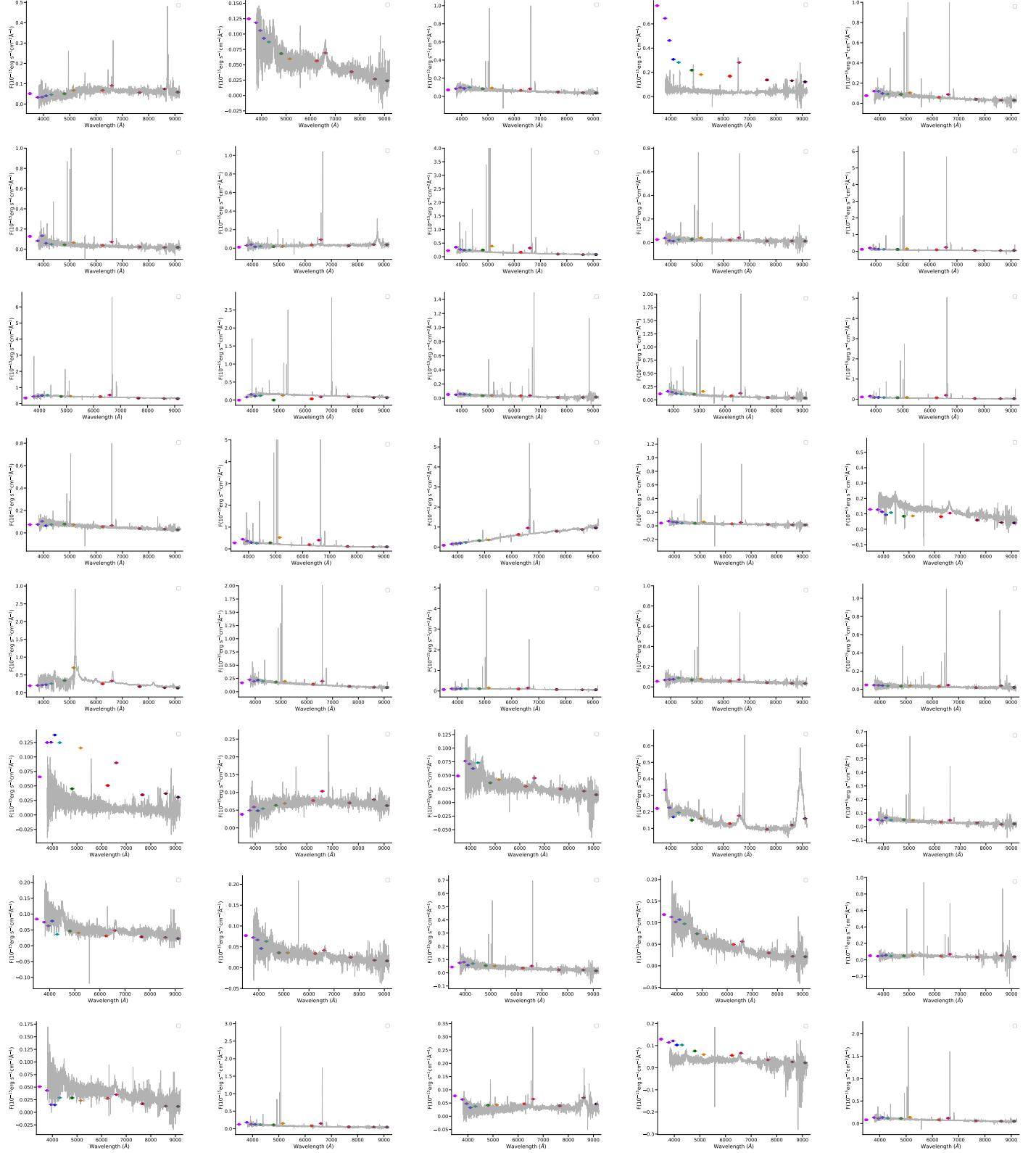


Table C1: –continued

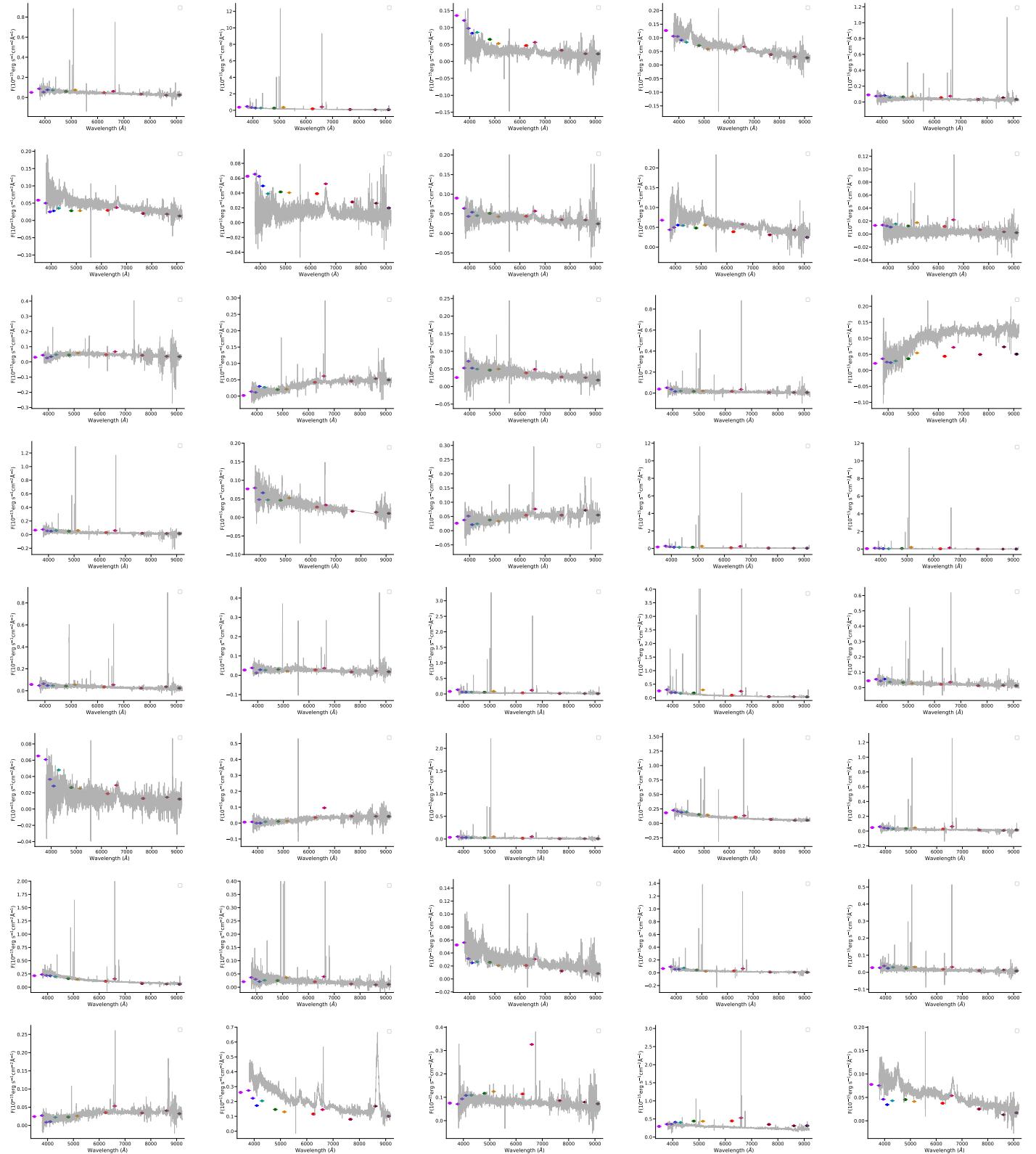


Table C1: –continued

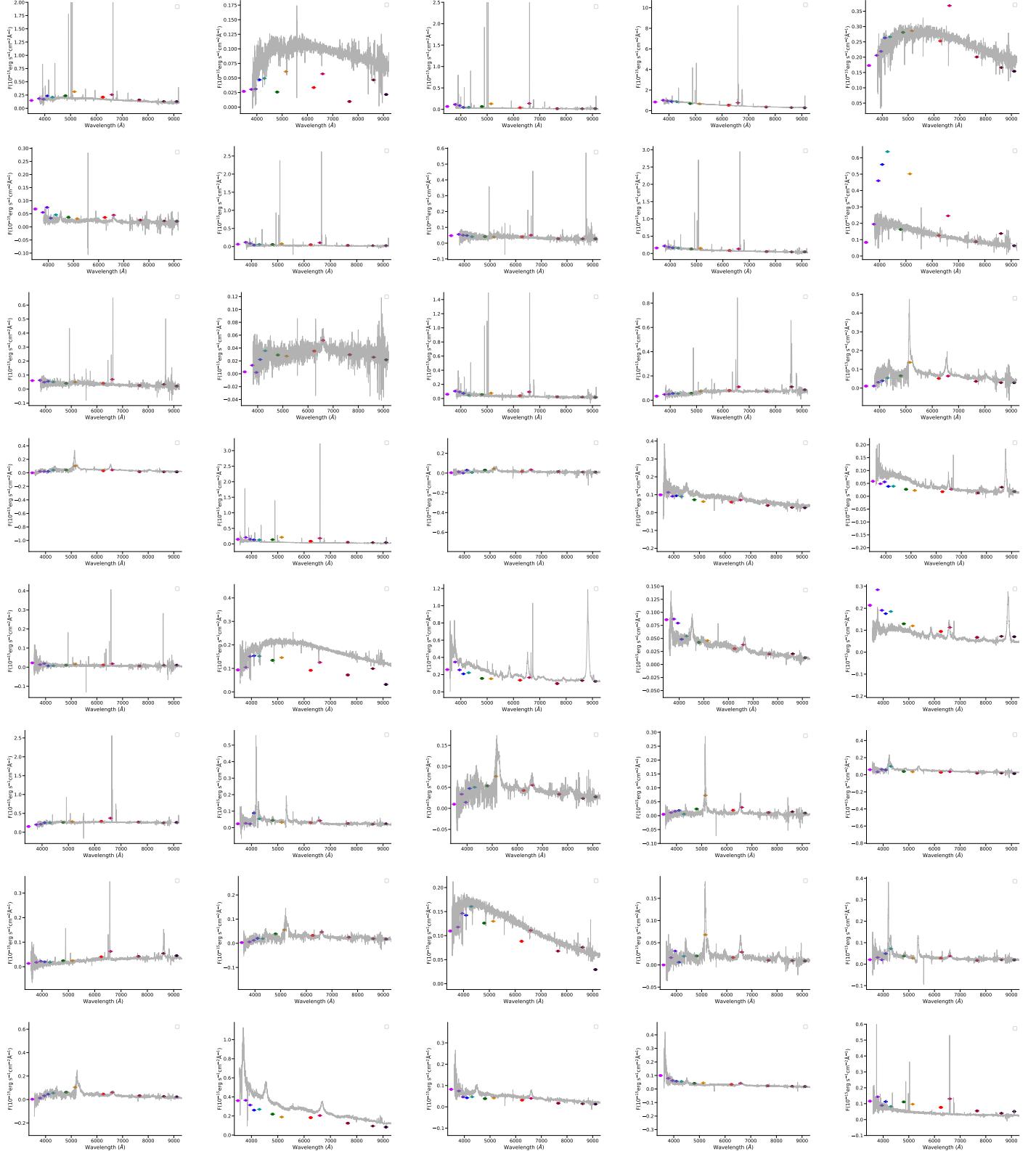
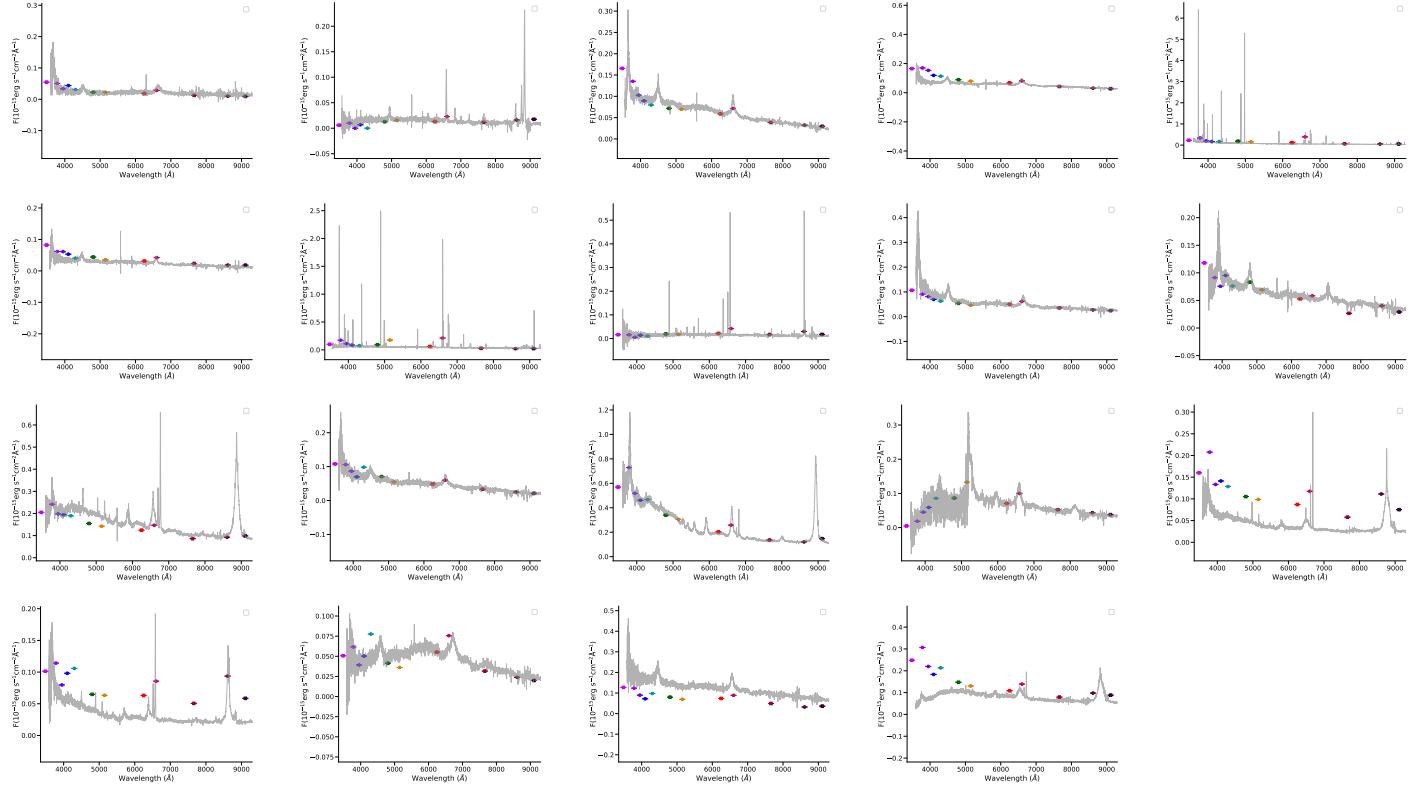


Table C1: –continued



## APPENDIX D: LAMOST SPECTRA

Table D1: Espectra from LAMOST DR6

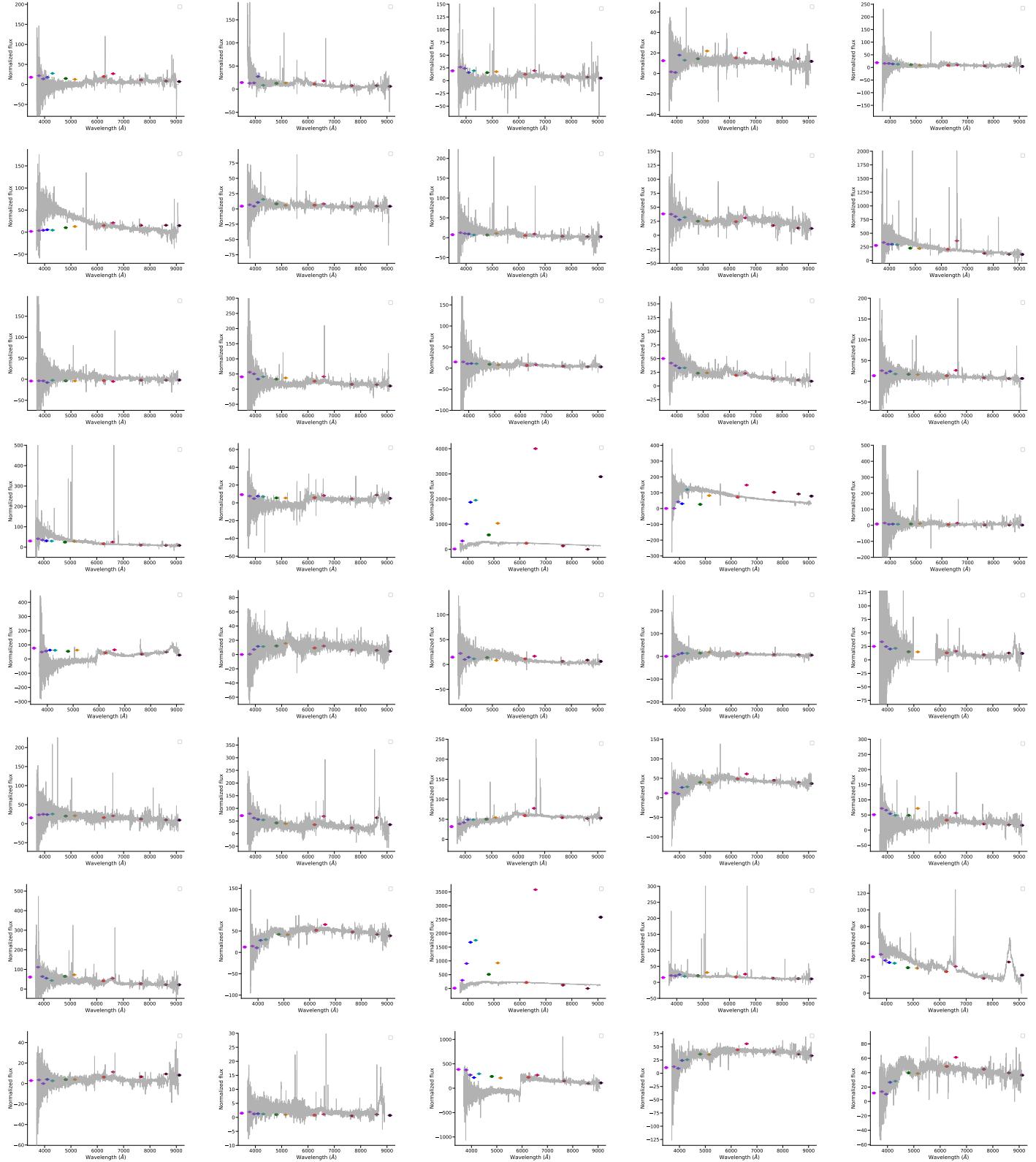


Table D1: –continued

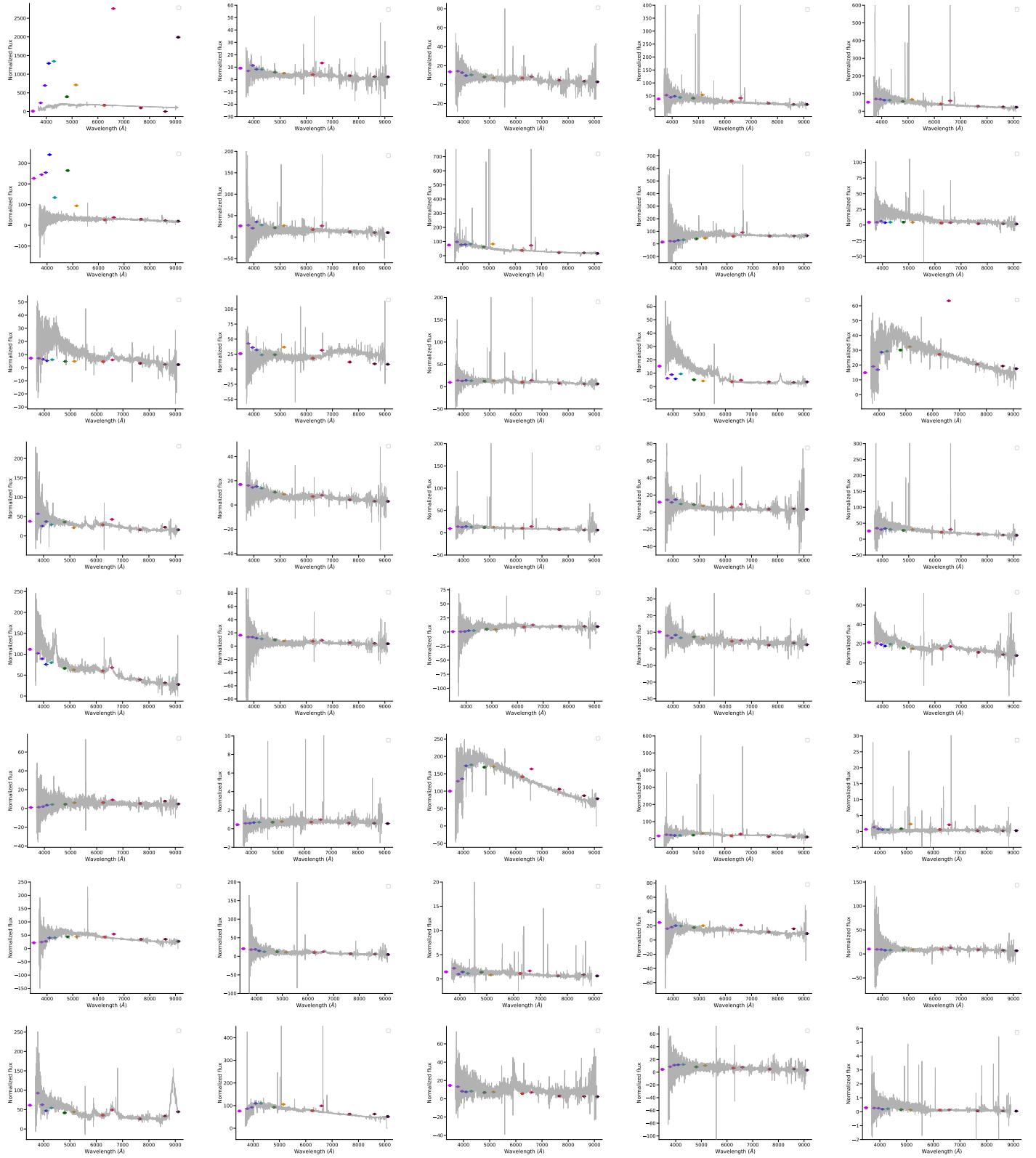
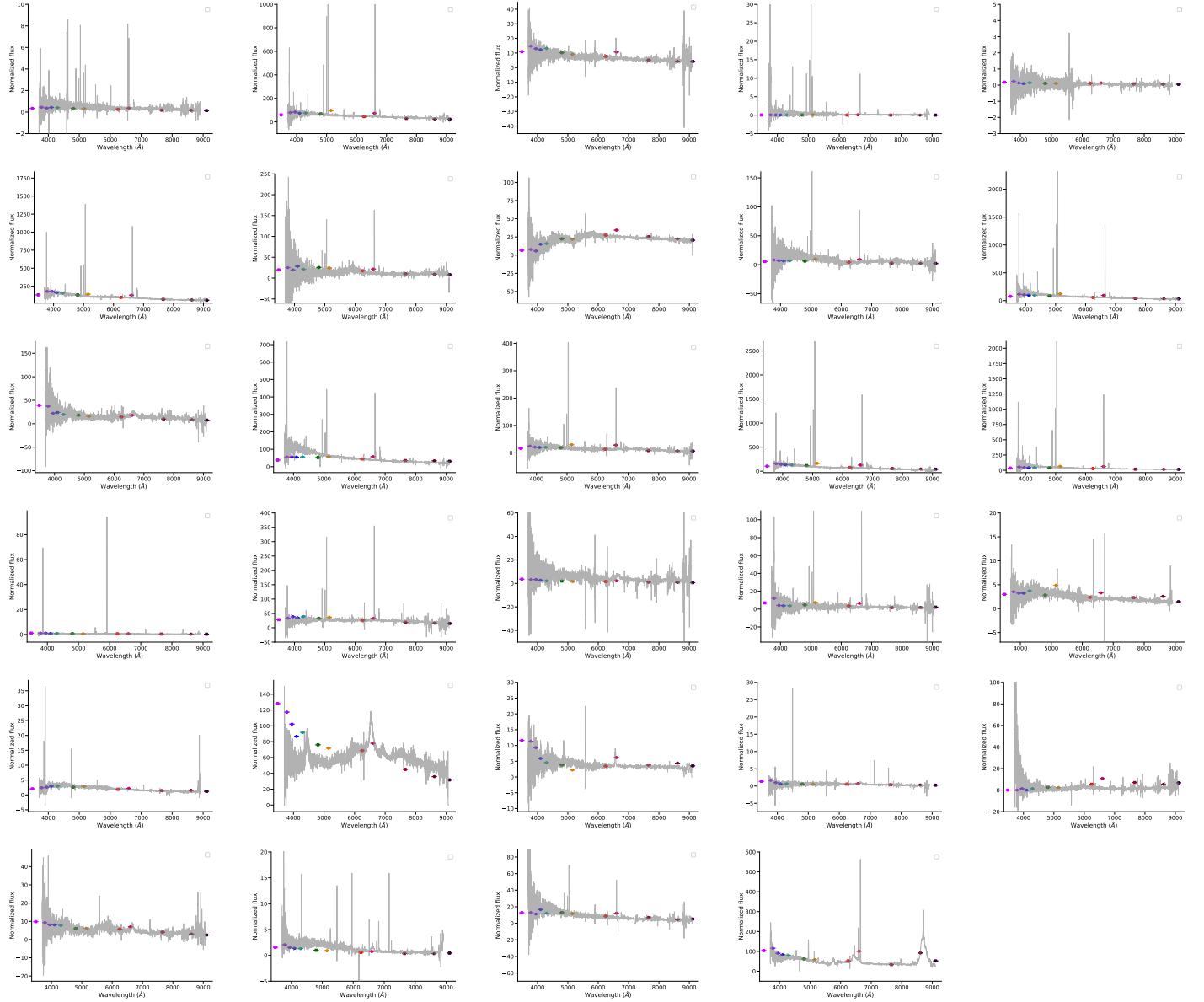


Table D1: –continued



This paper has been typeset from a TeX/LaTeX file prepared by the author.