








<b>Publication Year</b>	2019
<b>Acceptance in OA @INAF</b>	2021-04-12T10:27:49Z
<b>Title</b>	Deciphering the Large-scale Environment of Radio Galaxies in the Local Universe: Where Are They Born? Where Do They Grow? Where Do They Die?
<b>Authors</b>	MASSARO, Francesco; Álvarez-Crespo, N.; CAPETTI, Alessandro; BALDI, RANIERI DIEGO; PILLITTERI, Ignazio Francesco; et al.
<b>DOI</b>	10.3847/1538-4365/aaf1c7
<b>Handle</b>	<a href="http://hdl.handle.net/20.500.12386/30733">http://hdl.handle.net/20.500.12386/30733</a>
<b>Journal</b>	THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES
<b>Number</b>	240



# Deciphering the Large-scale Environment of Radio Galaxies in the Local Universe: Where Are They Born? Where Do They Grow? Where Do They Die?

F. Massaro<sup>1,2,3</sup> , N. Álvarez-Crespo<sup>1,2,3</sup>, A. Capetti<sup>2</sup> , R. D. Baldi<sup>4</sup> , I. Pillitteri<sup>5</sup> , R. Campana<sup>6</sup> , and A. Paggi<sup>1,2,3</sup> 

<sup>1</sup> Dipartimento di Fisica, Università degli Studi di Torino, via Pietro Giuria 1, I-10125 Torino, Italy; [f.massaro@unito.it](mailto:f.massaro@unito.it)

<sup>2</sup> INAF-Osservatorio Astrofisico di Torino, via Osservatorio 20, I-10025 Pino Torinese, Italy

<sup>3</sup> Istituto Nazionale di Fisica Nucleare, Sezione di Torino, I-10125 Torino, Italy

<sup>4</sup> Department of Physics and Astronomy, University of Southampton, Highfield, Southampton SO17 1BJ, UK

<sup>5</sup> INAF-Osservatorio Astronomico di Palermo G.S. Vaiana, Piazza del Parlamento 1, I-90134 Palermo, Italy

<sup>6</sup> INAF/OAS, via Piero Gobetti 101, I-40129, Bologna, Italy

Received 2018 April 19; revised 2018 October 26; accepted 2018 November 16; published 2019 January 24

## Abstract

The role played by the large-scale environment in the nuclear activity of radio galaxies (RGs) is still not completely understood. Accretion mode, jet power, and galaxy evolution are connected with their large-scale environment on scales from tens to hundreds of kiloparsecs. Here we present a detailed statistical analysis of the large-scale environment for two samples of RGs up to redshifts  $z_{\text{src}} = 0.15$ . The main advantages of our study over studies in the literature are the extremely homogeneous selection criteria of the catalogs adopted to perform our investigation. This is also coupled with the use of several clustering algorithms. We performed a direct search of galaxy-rich environments around RGs by using them as beacons. To perform this study we also developed a new method that does not appear to suffer from a strong  $z_{\text{src}}$  dependence as other algorithms do. We conclude that, despite their radio morphological classification (FR I versus FR II) and/or their optical classification (high- or low-excitation radio galaxy (HERG or LERG)), RGs in the local universe tend to live in galaxy-rich large-scale environments that have similar characteristics and richness. We highlight that the fraction of FR I LERGs inhabiting galaxy-rich environments appears to be larger than that of FR II LERGs. We also found that five out of seven FR II HERGs, with  $z_{\text{src}} \leq 0.11$ , lie in groups/clusters of galaxies. However, we recognize that, despite the high level of completeness of our catalogs, when restricting to the local universe, the low number of HERGs ( $\sim 10\%$  of the total FR IIs investigated) prevents us drawing a strong statistical conclusion about this source class.

*Key words:* galaxies: active – galaxies: clusters: general – galaxies: jets – methods: statistical – radio continuum: galaxies – surveys

## 1. Introduction

In the early 1970s Fanaroff & Riley (1974) proposed to classify extragalactic radio sources that have an extended structure resolved in two or more components at 1.4 GHz. Their scheme is based on the ratio  $R_{\text{FR}}$  of the angular separation between regions of highest surface brightness on opposite sides of the central radio galaxy or quasar, to the total extent of the source measured up to the lowest contour level. Any compact component located on the central galaxy, such as the radio core, was not taken into account. Radio sources having  $R_{\text{FR}} \leq 0.5$  (i.e., edge-darkened) were placed in class I (and named FR Is), while those for which  $R_{\text{FR}} > 0.5$  (i.e., edge-brightened) were placed in class II (and known as FR IIs).

This radio morphological distinction corresponds to a sharp division in luminosities. Radio sources having  $L_{178\text{MHz}}$  lower than  $2 \times 10^{25} \text{ W Hz}^{-1} \text{ s}^{-1}$  appeared to be almost all FR Is while those above this threshold were FR IIs. This luminosity threshold was remarkably close to the dividing line between radio sources with strong and weak cosmological evolution (see, e.g., Longair 1971).

This FR classification scheme was then linked a few decades later to the environment on the megaparsec scale of the extragalactic radio galaxies. It was found that FR Is generally inhabit galaxy-rich environments, being members of groups or galaxy clusters, while FR IIs tend to be more isolated (see, e.g., Zirbel 1997), with a few well known exceptions (see, e.g., Hardcastle & Worrall 2000 for a recent analysis of the X-ray

observations of FR IIs), such as the archetypal Cygnus A (see, e.g., Carilli & Barthel 1996 for a review).

In the last decade a firm link between the properties of optical emission, accretion mode, and host galaxy, including star formation rate, was established for the radio galaxy population (see, e.g., Baldi & Capetti 2008, 2010; Balmaverde et al. 2008; Tasse et al. 2008; Smolčić et al. 2009; Hardcastle et al. 2013; Mingo et al. 2014). An additional classification was developed for radio galaxies in the 1980s. This was based on the properties of their optical emission lines (Hine & Longair 1979), distinguishing between high- and low-excitation radio galaxies (HERGs and LERGs, respectively; see also Laing et al. 1994). Their differences are not simply related to the orientation of the active galaxy with respect to the line of sight but are also related to their accretion modes (i.e., radiatively efficient versus inefficient) (see, e.g., Chiaberge et al. 2002; Hardcastle et al. 2006, 2009; Best & Heckman 2012, and references therein). In addition, HERGs appear to have, almost exclusively, an FR II radio morphology, while LERGs can be FR I or FR II (see, e.g., Hine & Longair 1979; Laing et al. 1994). Hence accretion mode does not directly determine radio morphological class (Heckman & Best 2014).

As occurs for FR Is and FR IIs, LERGs are preferentially low-luminosity radio sources, mostly lying at low redshifts  $z_{\text{src}}$ ,<sup>7</sup> while HERGs dominate the high-luminosity radio sky,

<sup>7</sup> Here we adopt the symbol  $z_{\text{src}}$  to indicate the source redshift rather than the usual  $z$  to distinguish it from the redshift of a possible nearby galaxy group or cluster, labelled as  $z_{\text{cl}}$ .

being at higher  $z_{\text{src}}$ . This appears clear even when considering extragalactic sources selected from radio surveys with high flux limits and large beams, which, as shown recently, are also not representative of the whole radio galaxy population (see, e.g., Capetti et al. 2017a). Therefore it is crucial to consider both radio and optical classifications for the radio galaxy population while investigating their large-scale environments.

Using the tenth-nearest-neighbor estimator in the  $z_{\text{src}}$  range between 0.02 and 0.10, it has been found that radio-loud active galaxies are preferentially located in galaxy groups and in galaxy clusters of poor-to-moderate richness, consistent with previous results (see, e.g., Prestage & Peacock 1988; Hill & Lilly 1991). In particular, the flux ratio of absorption line to emission line changes dramatically with the environment, with essentially all radio-loud active galaxies in rich environments showing no emission lines (see, e.g., Best 2004). Thus considerable care must be taken in selecting samples of radio-loud active galaxies from their optical emission-line properties (LERGs versus HERGs), since, when investigating how environmental properties are related to their optical spectra, selection criteria should not be related to their optical properties.

Recently, Gendre et al. (2013) showed that at a given radio luminosity  $L_R$  at 1.4 GHz, the FR morphological dichotomy is consistent with both accretion modes even when restricting to only rich or only poor environments. This could imply that radio morphology is independent of the accretion mode and depends on the power of the jet and its interactions with the larger-scale environment. Thus, FR Is lie in higher-density environments than FR IIs. This picture is therefore consistent with FR Is having jets disrupted by a denser surrounding medium (Bicknell 1994; Laing & Bridle 2008). Gendre et al. (2013) also claimed that accretion modes could be linked to the large-scale environment, with HERGs living almost exclusively in low-density environments and LERGs inhabiting a wider range of galaxy densities, independently of their radio morphology.

Using X-ray observations, Ineson et al. (2013) performed a systematic study of cluster environments of radio galaxies at  $z_{\text{src}} \sim 0.5$ . They found tentative evidence for a correlation between radio luminosity and cluster X-ray luminosity, possibly driven by the LERG subpopulation. Then, at  $z_{\text{src}} \sim 0.1$ , Ineson et al. (2015) claimed a stronger link between radio luminosity and richness and between radio luminosity and central density for LERGs, but not for HERGs, although there are fewer HERGs at low  $z_{\text{src}}$ . No differences in LERGs were found between the two analyses.

In contrast with the results of both Best (2004) and Ineson et al. (2015), Belsole et al. (2007) found no link between radio luminosity and galaxy density at higher  $z_{\text{src}}$ ; however, their sample could be biased toward a selection of HERGs.

More recently, Miraghaei & Best (2017) compared FR I LERGs with FR II LERGs at fixed stellar mass and radio luminosity, showing that the former typically reside in richer environments and are hosted by smaller galaxies with higher mass surface density. This picture is again consistent with jet disruption effects, a possible driver of the FR dichotomy.

Finally, adopting the fifth nearest neighbor density  $\Sigma_5$ , as in the analysis of Best (2004) Ching et al. (2017) confirmed previous results with a larger sample. LERGs and HERGs exist in different large-scale environments depending on their radio luminosity, with LERGs of high radio luminosity more likely

to be in galaxy groups. In contrast, the environments of HERGs and low-luminosity LERGs are indistinguishable from those of a radio-quiet control sample.

Comparing claims and results from different analyses, carried out with different techniques and on different samples, requires extreme caution. Methods to estimate the cluster richness, or procedures to associate a source with a galaxy group or cluster, or differences in the region sizes selected for galaxy counts, could introduce biases. In addition, the possible evolution of the environments with  $z_{\text{src}}$ , changes in HERG and LERG populations with  $z_{\text{src}}$ , a lack of powerful sources in our local universe, and the Malmquist bias in flux-limited catalogs could also affect analyses and comparisons. Nevertheless, analyses based on ill-defined small groups of sources and, as recently shown, conclusions based on samples selected from radio surveys with a high flux limit and large beam, such as the Third Cambridge catalog (3C; see, e.g., Edge et al. 1959; Bennett 1962; Spinrad et al. 1985), could also be strongly affected by selection biases (Capetti et al. 2017a, 2017b).

To shed light on the role played by the large-scale environment in the nuclear activity of radio galaxies, here we present a detailed study of the large-scale environment of radio galaxies using well-defined and statistically homogeneous catalogs of FR I (LERGs) and FR II (LERGs and HERGs) radio galaxies at  $z_{\text{src}} \leq 0.15$ . We highlight differences and advantages of the analysis carried out here in comparison with literature studies.

The paper is organized as follows. In Section 2 we present all samples and catalogs used to carry out our analysis, while in Section 3 we outline definitions adopted to perform our investigation. Then in Section 4 we describe step-by-step the clustering procedure used. In Section 5 we discuss results obtained, then Section 6 is devoted to our summary and conclusions. A comparison with literature claims is presented in Section 7. Finally, in Section 8 we discuss future perspectives and possible developments of our analysis achievable with dedicated X-ray observations. Technical details on clustering algorithms used here are fully described in the Appendix.

We adopt cgs units for numerical results and we also assume a flat cosmology with  $H_0 = 69.6 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.286$ , and  $\Omega_\Lambda = 0.714$  (Bennett et al. 2014), unless otherwise stated. Thus,  $1''$  corresponds to 0.408 kpc at  $z_{\text{src}} = 0.02$  and 2.634 kpc at  $z_{\text{src}} = 0.15$ .

## 2. Sample Selection

Several source samples and catalogs have been used to carry out our analysis: (i) two catalogs of radio galaxies, extremely homogeneous and carefully selected on the basis of multi-frequency observations; (ii) a catalog of random positions in the Sloan Digital Sky Survey (SDSS) footprint; (iii) a sample of quiescent elliptical galaxies; and (iv) two catalogs of groups and clusters of galaxies, again based on the SDSS observations. Here we describe them briefly with particular attention to their selection criteria.

### 2.1. Radio Galaxies

We recently created two catalogs of FR I and FR II radio galaxies (i.e., FRICAT and FRIICAT respectively; Capetti et al. 2017a, 2017b) that combine observations available in the SDSS Data Release 9 (Ahn et al. 2012), the National Radio

Astronomy Observatory (NRAO) Very Large Array (VLA) Sky Survey (Condon et al. 1998), and the Faint Images of the Radio Sky at Twenty cm (FIRST) survey (White et al. 1997). All sources in these catalogs have optical spectra that allowed us to obtain their  $z_{\text{src}}$  and determine their LERG versus HERG classification, precisely and unambiguously.

Radio galaxy catalogs were selected starting from the original sample of Best & Heckman (2012). We first considered only those sources classified as active galactic nuclei, and we then performed a cut at  $z_{\text{src}}$  lower than 0.15. This led to the selection of 3356 sources out of the original 18,286. Subsequently we visually inspected all FIRST images for each individual source, selecting only those having radio emission beyond 30 kpc, measured from the position of the optical host galaxy. Radio contours of surface brightness were constructed at the level of 0.45 mJy/beam, thus matching the FIRST sensitivity and taking into account the cosmological dimming of the surface brightness. The total number of radio sources selected decreases to 743. Then we performed a final classification distinguishing between FR Is and FR IIs.

For the present analysis we restricted our radio galaxy catalogs to those sources lying in the central part of the SDSS footprint (see, e.g., Ahn et al. 2012), the same area as covered by the main catalog of groups and clusters of galaxies adopted in our analysis (see subsequent sections). In this way, the FRICAT, which includes 219 radio galaxies, all optically classified as LERGs and spanning a redshift range between 0.02 and 0.15, was reduced to 195 sources, while for the FRIICAT the number of sources decreased from 129 to 115 with  $z_{\text{src}}$  between 0.045 and 0.15. In the FRIICAT there are 14 radio galaxies classified as HERGs while all the others are LERGs.

At a given [O III] luminosity, sources listed in the FRICAT show radio luminosities spanning about two orders of magnitude and extending to much lower ratios between radio and line power than the FR Is listed in the 3C catalog (see Capetti et al. 2017a for additional details). On the other hand, the majority of the FR IIs listed in the FRIICAT have a radio luminosity up to two orders of magnitude lower than the threshold one between FR Is and FR IIs of the 3C catalog (see, e.g., Capetti et al. 2017b for more details). For both catalogs the relation between the morphological classification and radio luminosity disappears when considering low-power radio sources.

## 2.2. Mock Sources

In our analysis we used a catalog of mock sources (labeled as MOCK hereinafter) to estimate the efficiency of our procedures and to estimate their uncertainties. This has been created by shifting the position of all FRICAT and FRIICAT radio galaxies by a random radius between  $2^\circ$  and  $3^\circ$  in a random direction of the sky, so as to obtain 5000 fake sources/positions. The range of values for the random shift was chosen to be larger than the maximum angular separation corresponding to 2 Mpc in the radio galaxy catalogs (i.e.,  $1^\circ.1$ ) and smaller than  $3^\circ$  to preserve the sky distribution of sources in the SDSS footprint. Similar procedures have already been successfully adopted in previous analyses with optical and infrared catalogs (see, e.g., Massaro et al. 2011, 2014; D’Abrusco et al. 2014).

We then removed from the MOCK sample all sources having a radio counterpart within  $5''$ . To preserve a redshift

distribution of the MOCK catalog similar to that of radio galaxies, we verified that the source ratio between the two catalogs, per bin of  $z_{\text{src}}$  equal to 0.01, is at least 15. Finally we highlight that to create the MOCK catalog we kept all optical magnitudes associated with  $z_{\text{src}}$  so as to also keep the luminosity distribution similar to that of radio galaxies.

The final MOCK sample lists 4056 sources, more than an order of magnitude larger than the total number of radio galaxies considered (i.e., 310).

## 2.3. Quiescent Elliptical Galaxies

We also built a catalog of *quiescent elliptical galaxies* (hereinafter ELL). This allows us to investigate optical colors of sources in the large-scale environment of radio galaxies. This catalog will be used only to search for elliptical galaxies surrounding our radio galaxies and to estimate their local source density, and not for a comparison with the radio galaxy catalogs.

1. We first considered all 667,944 sources listed in the Galaxy Zoo<sup>8</sup> data release 1 (Lintott et al. 2008).
2. We then selected those having a single counterpart in the SDSS data release 9 within  $5''$ . We considered only galaxies with SDSS flags *spType* and *spClass* equal to GALAXY and *subclass* NULL. We chose only those objects having an elliptical classification based on at least 45 votes, according to the Galaxy Zoo analysis, and with spectroscopic  $z_{\text{src}}$  smaller than 0.15, as for radio galaxies.
3. We included only elliptical galaxies with clean photometry (i.e., SDSS flags *q\_mode* = 1 and *Q* > 2) and classified as *galaxies* (i.e., SDSS flag *cl* equal to 3).
4. We excluded sources for which the Galaxy Zoo classification is uncertain.
5. We did not include galaxies having an uncertain estimate of  $z_{\text{src}}$ .
6. Sources with a radio counterpart within  $5''$  were also excluded to avoid a possible contamination from a radio galaxy.

## 2.4. Catalogs of Groups and Clusters of Galaxies

Several catalogs of galaxy clusters and groups are available for the SDSS footprint. We selected the one created by Tempel et al. (2012, hereinafter T12) to carry out our analysis since it has the largest number of cluster/group detections with spectroscopic redshifts. This catalog of groups and clusters was created using a modified version of the friends-of-friends (FoF) algorithm (Huchra & Geller 1982; Tago et al. 2010). Its distribution of redshift  $z_{\text{cl}}$  spans a range between 0.009 and 0.20, peaking around 0.08 and thus becoming less efficient at larger  $z_{\text{cl}}$  values.

We considered only groups and clusters with an estimate of spectroscopic redshift listed in the T12 catalog, for a total of 77,858 sources, for which the galaxy density, indicated by the parameter  $N_{\text{gal}}$ , was also computed.

Then, we also considered a second catalog of galaxy groups and clusters built using a new Gaussian mixture brightest cluster galaxy (GMBCG) algorithm (Hao et al. 2010). This was created using the *red sequence* (Visvanathan & Sandage 1977; Gladders et al. 1998) combined with the search for a brightest

<sup>8</sup> <https://www.galaxyzoo.org>



cluster galaxy (BCG). The GMBCG catalog was chosen since it is more efficient than the T12 at  $z_{\text{cl}}$  larger than 0.08. The GMBCG catalog, including only 1296 entries with spectroscopic  $z_{\text{cl}}$  below 0.15 out of 55,424 clusters/groups, allowed us to verify the number of BCG candidates in our radio galaxy and quiescent elliptical catalogs.

### 3. Cosmological Neighbors and Candidate Elliptical Galaxies

For each radio galaxy we downloaded a table listing all optical sources detected in the SDSS DR9 with clean photometry (i.e., SDSS flags  $q_{\text{mode}} = 1$  and  $Q = 3$  and  $mode = 1$ ), lying within a 2 Mpc radius, computed at  $z_{\text{src}}$  of the central source. A radius of 2 Mpc was chosen to be slightly larger than the typical size of massive galaxy clusters (i.e.,  $R_{200} \sim 1.4$  Mpc; Rines et al. 2013). We then defined two types of sources in their environment.

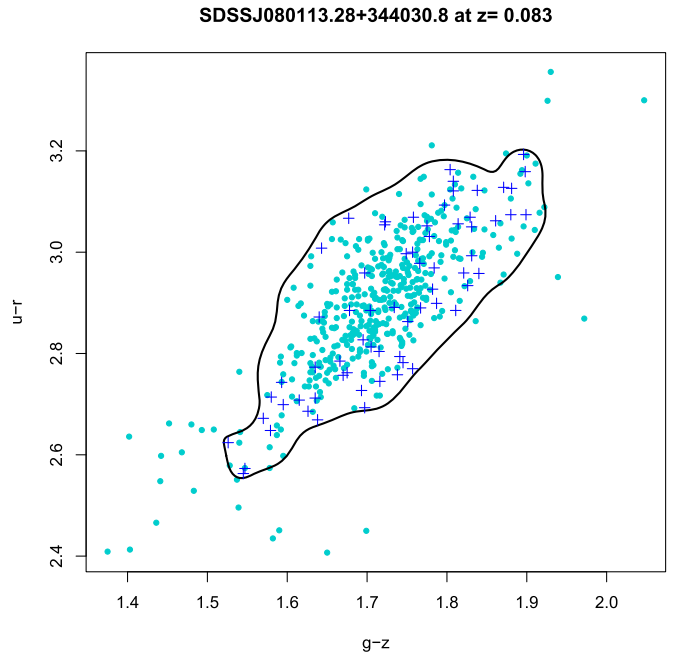
*A. Cosmological neighbors:* all optical sources lying within the 2 Mpc radius computed at  $z_{\text{src}}$  of the central object with all the SDSS magnitude flags indicating a galaxy-type object (i.e.,  $uc = rc = gc = ic = zc = 3$ ), and having a spectroscopic redshift  $z$  with  $\Delta z = |z_{\text{src}} - z| \leq 0.005$  (i.e.,  $\sim 1500$  km s $^{-1}$ ). This choice of  $\Delta z$  corresponds to the maximum velocity dispersion in groups and clusters of galaxies (see, e.g., Moore et al. 1993; Eke 2004; Berlind et al. 2006).

*B. Candidate elliptical galaxies:* all optical sources lying within the 2 Mpc distance from the central radio galaxy, estimated at  $z_{\text{src}}$ , and having  $u - r$  and  $g - z$  colors consistent with those of the ELL sample within  $\Delta z < 0.005$ . This color-color selection is based on the isodensity contours computed adopting kernel density estimation (KDE: see, e.g., Richards et al. 2004; D’Abrusco et al. 2009; Massaro et al. 2013a), at a 90% level of confidence. Sources selected as candidate elliptical galaxies do not necessarily have spectroscopic redshifts, they have only the same colors as elliptical galaxies at  $z_{\text{src}}$  of the radio galaxy. We only perform this selection for elliptical-type galaxies, in the large-scale environment of our radio galaxies, because their fraction in galaxy groups or clusters is much larger than that of spirals (see, e.g., Biviano 2000).

In Figure 1 we show the color-color plot ( $u - r$  versus  $g - z$ ) for optical sources surrounding SDSS J080113.28+344030.8. Sources in the ELL sample within  $\Delta z < 0.005$  centered at  $z_{\text{src}}$  of SDSS J080113.28+344030.8 are reported as cyan circles together with their KDE isodensity contours (black) while blue crosses are the candidate elliptical galaxies in the 2 Mpc field of the central radio galaxy.

In Figure 2 we show the FR I radio galaxy SDSS J101114.38+191425.7 (central black circle in both panels) where all the SDSS sources lying within 2 Mpc (gray background circles in both panels), computed at  $z_{\text{src}}$  of the central object, are shown together with (i) SDSS sources with spectroscopic  $z$  (orange circles in the left panel), (ii) cosmological neighbors (red circles in the right panel) and (iii) candidate elliptical galaxies (blue crosses in both panels). In the same figure we also show the location of the closest galaxy cluster/group in the T12 catalog, labelled with its  $z_{\text{cl}}$  (green circle in the right panel).

Given our color-color selection of candidate elliptical galaxies, based on four SDSS magnitudes, we built color-magnitude plots to verify that selected cosmological neighbors and candidate elliptical galaxies also belong to a well known feature of galaxy clusters: the red sequence. This is just an



**Figure 1.**  $u - r$  vs.  $g - z$  color-color plot with cyan circles representing quiescent elliptical galaxies in the ELL sample within  $\Delta z < 0.005$  centered at  $z_{\text{src}}$  of SDSS J080113.28+344030.8. The black line is their 90% contour level computed with KDE. Blue crosses are the candidate elliptical galaxies, selected among those optical sources lying within the angular separation corresponding to 2 Mpc around the central radio galaxy.

additional check to verify the presence of a galaxy-rich large-scale environment around the radio galaxies investigated, since galaxies that are members of groups and clusters tend to be redder than background and foreground galaxies in the same field. In Figure 3 we show the plot built with the  $r$  and  $i$  magnitudes (i.e., those used in the GMBCG) for the FR I radio galaxy SDSS J080113.28+344030.8. It is clear that both cosmological neighbors (red circles) and candidate elliptical galaxies (blue crosses) belong to the red sequence. This color code for both cosmological neighbors and candidate elliptical galaxies will be maintained for the rest of the figures reported in the paper.

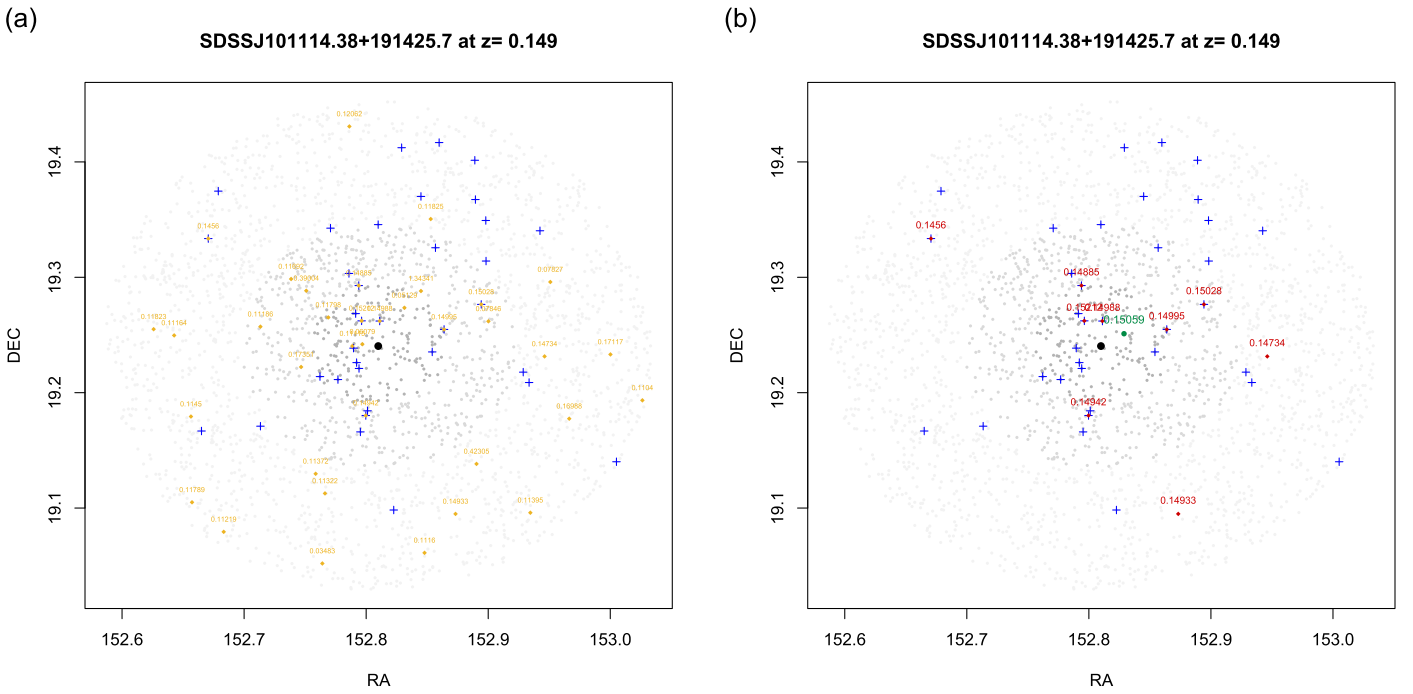
Finally, we note that the whole analysis reported above was performed not only for both the radio galaxy catalogs but also for the MOCK catalog, adopting exactly the same criteria and thresholds, to quantify the “noise” of our procedures, as described in the following.

## 4. The Step-by-step Clustering Analysis

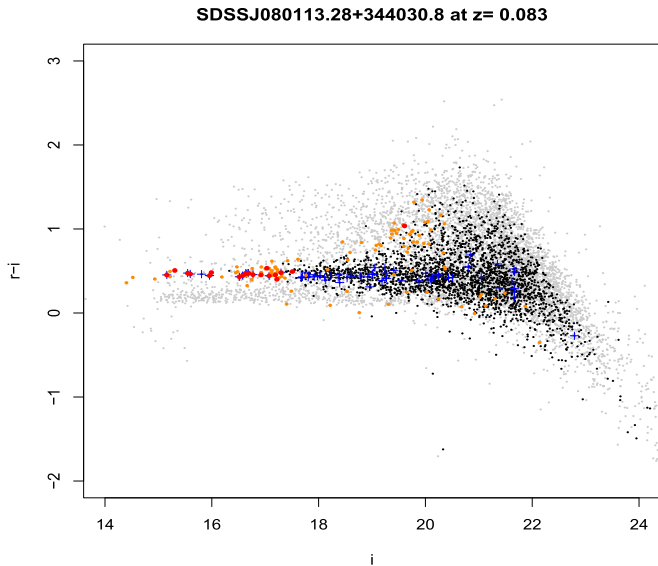
### 4.1. Step 1: Positional Cross-matches with Catalogs of Groups and Clusters of Galaxies

The first step to test whether a radio galaxy lives in a galaxy-rich large-scale environment was performed by searching for groups and/or clusters of galaxies listed in the T12 catalog and within a 2 Mpc radius and having  $\Delta z = |z_{\text{src}} - z_{\text{cl}}| \leq 0.005$ , computed using only spectroscopic redshifts. The same analysis was then carried out for the MOCK catalog, where the  $z_{\text{src}}$  value corresponds to that of the fake source listed therein.

Figure 4 shows one of the results of the cross-matching analysis, plotting the projected distance  $d_{\text{proj}}$  between each radio galaxy and the closest galaxy group or cluster as a function of  $\Delta z$ . The same is shown for the MOCK catalog.

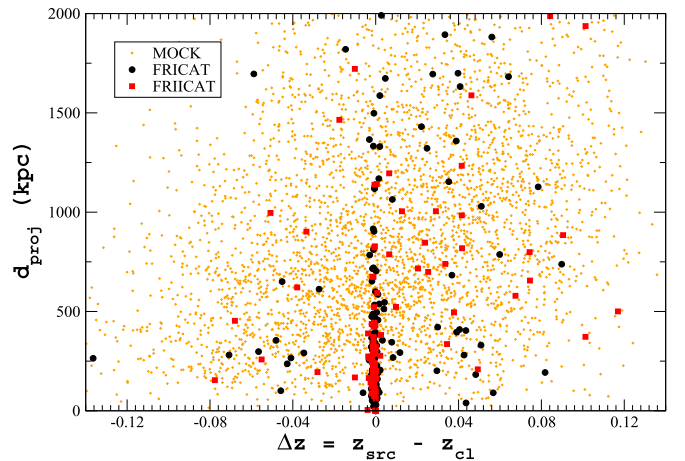


**Figure 2.** (a) The position of all SDSS sources within 2 Mpc computed at  $z_{\text{src}} = 0.149$  for the radio galaxy SDSS J101114.38+191425.7. Different intensities of gray indicate those lying within 500 kpc, 1 Mpc, and 2 Mpc, respectively. All SDSS objects with a spectroscopic  $z$  are shown as orange circles and their  $z$  value is also reported close to their location. Blue crosses in both panels mark candidate elliptical galaxies, i.e., SDSS sources in the field with optical colors similar to *quiescent elliptical galaxies* at  $z_{\text{src}} = 0.149$  and within  $\Delta z$  of 0.005. (b) Cosmological neighbors are shown as red circles, while the green point marks the location of the closest group or cluster of galaxies, again within  $\Delta z$  of 0.005, listed in the T12 catalog of galaxy clusters/groups.



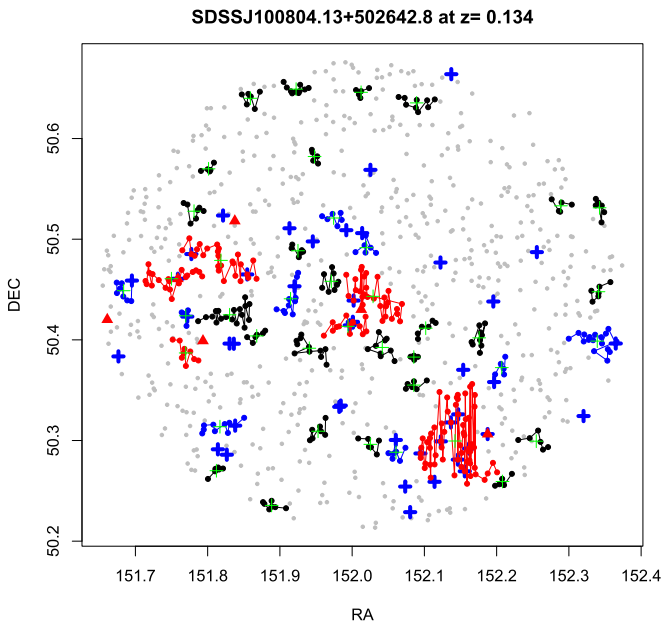
**Figure 3.** Color–magnitude plot using the SDSS  $r$  and  $i$  magnitudes for a radio galaxy in our sample. Background/foreground SDSS sources within 2 Mpc from the central source are marked with black circles while cosmological neighbors are shown in red and candidate elliptical galaxies as blue crosses. Generic SDSS sources with spectroscopic  $z$  are shown as orange circles. It is quite evident how both cosmological neighbors and candidate elliptical galaxies follow the “red sequence.”

More than 70% of the total number of FR Is and more than 55% of all FR IIs lie in galaxy-rich large-scale environments, being within 2 Mpc and within  $\Delta z \leq 0.005$  from a galaxy group/cluster. We also noticed that a large fraction of radio galaxies lie in a  $\Delta z$  range even smaller than the adopted threshold.



**Figure 4.** Projected distance  $d_{\text{proj}}$  as a function of  $\Delta z$  (i.e., the redshift difference between  $z_{\text{src}}$  of the radio galaxy or the MOCK source and  $z_{\text{cl}}$  of the positionally closest galaxy group/cluster in the T12 catalog. FR Is are marked with black circles while FR IIs are shown as red squares. MOCK sources are orange diamonds.

In this cross-matching analysis we initially considered as members of a group/cluster of galaxies only those radio galaxies for which the galaxy density  $N_{\text{gal}}$  in the T12 catalog is larger than 3 (i.e.,  $N_{\text{gal}} > 3$ ). However, we immediately noticed that clustering algorithms, such as the FoF and those described in the Appendix, are not able to find large-scale structures on scales of hundreds of kiloparsecs unless they are extremely rich. This is mainly due to the high number of optical sources in the background and/or in the foreground. Unfortunately, clustering algorithms could find separate groups/clusters that lie at the same  $z_{\text{cl}}$  and are indeed linked/related/connected to each other.



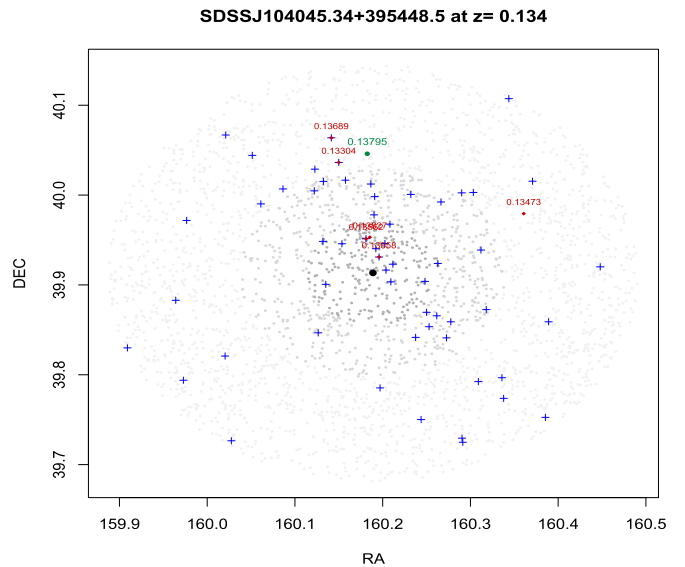
**Figure 5.** All source clusters found and identified adopting the density-based spatial clustering of applications with noise (DBSCAN) algorithm (see Appendix for more details) for a radio galaxy in our sample. Those clusters including at least a candidate elliptical galaxy are highlighted in blue, and those marked in red have at least one cosmological neighbor as a member, while clusters in black lack both. Being within  $\Delta z \leq 0.005$ , all the red clusters could belong to the same structure, but they are indeed found separately by the algorithm. Gray circles in the background mark the position of all SDSS optical sources classified as galaxies in the field of a 2 Mpc distance from the central radio galaxy.

An example of this problem is shown in Figure 5. Here results of one of the clustering algorithms adopted in our analysis, the density-based spatial clustering of applications with noise (DBSCAN, see Appendix for additional details), are reported for the radio galaxy SDSS J100804.13+502642.8. All source clusters identified by the DBSCAN that include at least a *candidate elliptical galaxy* are highlighted in blue, while those marked in red have at least one *cosmological neighbor* as a member. Source clusters found marked in black lack both *candidate elliptical galaxy* and *cosmological neighbor*. All red clusters could belong to the same cosmological structure, being within  $\Delta z \leq 0.005$  from the central radio galaxy, but are found and identified by the DBSCAN algorithm as separate source clusters.

Since clustering algorithms could potentially split a source cluster into smaller groups, the simple cross-matching analysis with the T12 catalog could be biased and the following criterion has been finally adopted to carry it out. We considered sources lying in galaxy-rich large-scale environments to be those having *more than one* galaxy group/cluster within 2 Mpc and with  $\Delta z \leq 0.005$  listed in the T12 catalog. Since the minimum value of  $N_{\text{gal}}$  reported therein is 2, having at least two galaxy groups/clusters within  $\Delta z \leq 0.005$  corresponds to the threshold previously adopted.

Finally, we performed cross-matches between radio galaxy catalogs with the GMBCG catalog. The main advantages here are that this catalog of galaxy clusters is based on a procedure more efficient than the T12 one at larger  $z_{\text{cl}}$ , and that it permits us also to search for FRIs and FR IIs that could be BCG candidates.

Cross-matches with the GMBCG were computed adopting the same procedure used for the T12. We assumed that a radio



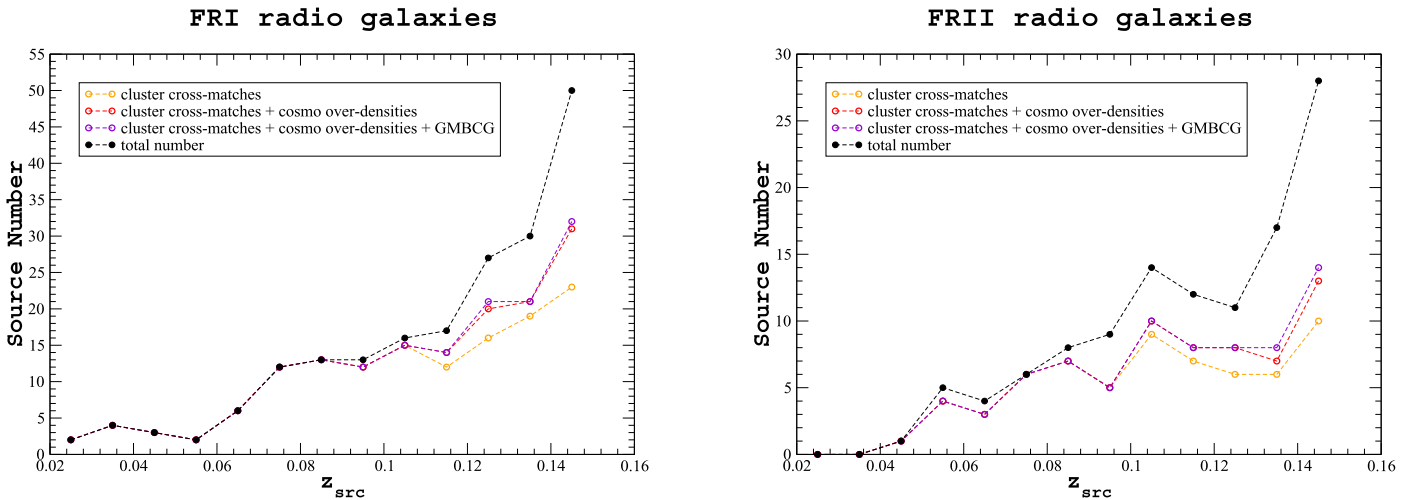
**Figure 6.** FR I radio galaxy SDSS J104045.34+395448.5 at  $z = 0.134$ , for which the T12 galaxy groups and clusters indicate a single cluster within 2 Mpc with an environmental density  $N_{\text{gal}} = 2$  and located more than 1 Mpc from the central source. In this case the number of cosmological neighbors is at least five, three of which lie within 500 kpc. Cosmological neighbors are shown as red circles, with their spectroscopic redshifts reported, while the green point marks the location of the closest group or cluster of galaxies within  $\Delta z < 0.005$  and blue crosses mark the locations of candidate elliptical galaxies. SDSS J104045.34+395448.5 lies in the center of the field, marked with a black circle. Gray circles mark all SDSS sources within a 2 Mpc distance from the central radio galaxy.

galaxy is associated with a group/cluster of galaxies that includes a BCG candidate when the redshift difference  $\Delta z$  computed between that of the central source  $z_{\text{src}}$  and  $z_{\text{cl}}$  reported in the GMBCG is less than 0.005. However, for these cross-matches  $\Delta z$  was computed using spectroscopic redshifts of both radio galaxy and GMBCG catalogs, when available, and with photometric estimates for GMBCG only in all other cases. The average uncertainty, at redshifts lower than 0.15, in the estimates of the photometric redshifts reported in the GMBCG is of the order of 0.023, evaluated as the mean difference between the spectroscopic and photometric values available for more than 1200 sources listed in the catalog. We noticed a posteriori that cross-matches with the GMBCG catalog provide only a negligible improvement over other methods, but, as previously stated, it was useful to identify radio galaxies that are potential BCG candidates.

#### 4.2. Step 2: Cosmological Overdensities

Cross-matching analysis has another problem in addition to the previous one: some source clusters are split in the T12 catalog. The efficiency of the algorithm used to create the galaxy group/cluster catalogs typically decreases at higher  $z_{\text{src}}$ , thus potentially biasing our analysis.

There are radio galaxies, for example SDSS J104045.34+395448.5, classified as FR I and shown in Figure 6, that have a large number of cosmological neighbors within 2 Mpc and even within 500 kpc but are associated with a T12 group/cluster with  $N_{\text{gal}} = 2$ . This galaxy density is too small to be considered a galaxy-rich large-scale environment according to our thresholds. However, sources such as SDSS J104045.34+395448.5 certainly lie in galaxy-rich large-scale environments, and thus an



**Figure 7.** Left: number of FR I radio galaxies in redshift bins of 0.01. Filled black circles represent the total number of sources per bin of  $z_{\text{src}}$ , while empty circles mark those sources having (i) a cross-match with a cluster/group of galaxies in the T12 cluster catalog (orange circles) plus (ii) a cosmological overdensity (red circles) plus (iii) an association with a cluster hosting a BCG candidate in the GMBCG catalog (magenta circles). Right: same as left panel but for the FR II radio galaxies.

additional criterion and/or method must be used to recover similar cases.

We then carried out the following Monte Carlo procedure to estimate the *cosmological overdensity*.

We also indicated as sources lying in galaxy-rich large-scale environments those whose number of cosmological neighbors within 500 kpc was within 5% of that measured for fake sources belonging to the MOCK catalog, in a  $z_{\text{src}}$  bin of 0.01. For redshifts larger than 0.1 we additionally required to have more than two cosmological neighbors within 1 Mpc. For example, in the  $z_{\text{src}}$  range between 0.13 and 0.14, as for SDSS J104045.34+395448.5 shown in Figure 6, there are 493 sources in the MOCK catalog, but only 18 (i.e., less than 4% in this redshift bin) show a larger number of cosmological neighbors than this radio galaxy. Thus we claimed that SDSS J104045.34+395448.5 also lies in a galaxy-rich large-scale environment even if it cannot be found by the cross-matching analysis.

The threshold of 5% on the fraction of cosmological neighbors around each radio galaxy was set arbitrarily. However, the results of our analysis do not change if we consider a more conservative value. Changing this threshold can only decrease the fraction of sources claimed to be in galaxy-rich large-scale environments but will preserve our main results. Estimating cosmological overdensity was indeed necessary at redshifts larger than  $\sim 0.1$ , where the efficiency of the T12 cluster catalog strongly decreases (Tempel et al. 2012). In addition, the cosmological overdensity method appears to be redshift-independent with respect to cluster cross-matches, as discussed in detail in the following sections.

#### 4.3. Threshold Summary

In summary, we claim that a source (i.e., radio galaxy or MOCK) lies in a galaxy-rich large-scale environment when at least one of the following statements is verified.

1. There is a group/cluster of the T12 catalog within 2 Mpc, with  $N_{\text{gal}}$  larger than 3 and with  $\Delta z \leq 0.005$ , or more than one group/cluster of galaxies with the same constraints but also having  $N_{\text{gal}} = 2$ .

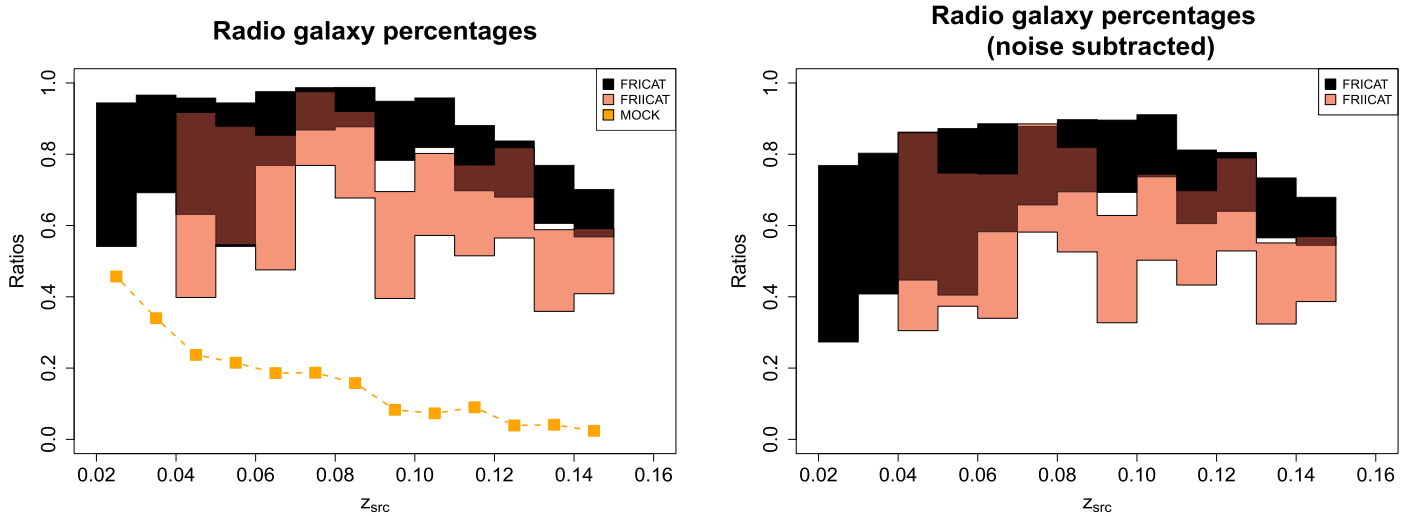
2. The redshift difference  $\Delta z$  computed between that of the central source and  $z_{\text{cl}}$  reported in the GMBCG is less than 0.005.
3. The number of cosmological neighbors is more than expected in random positions of the sky within a 5% threshold for the redshift bin that the source belongs to.

The first two criteria are related to the cluster cross-matching analysis performed with the T12 and GMBCG catalogs, while the third one helps us to recover split clusters, at the same redshift, potentially due to the FoF algorithm in the T12 catalog. It is worth highlighting that all the above criteria are equivalent in identifying galaxy-rich large-scale environments. Thus, as shown in the following, below  $z_{\text{cl}} = 0.08$ , where the T12 catalog has the higher efficiency in detecting galaxy clusters and groups, radio galaxies are found in galaxy-rich large-scale environments adopting either the first or the second criterion with only a few exceptions.

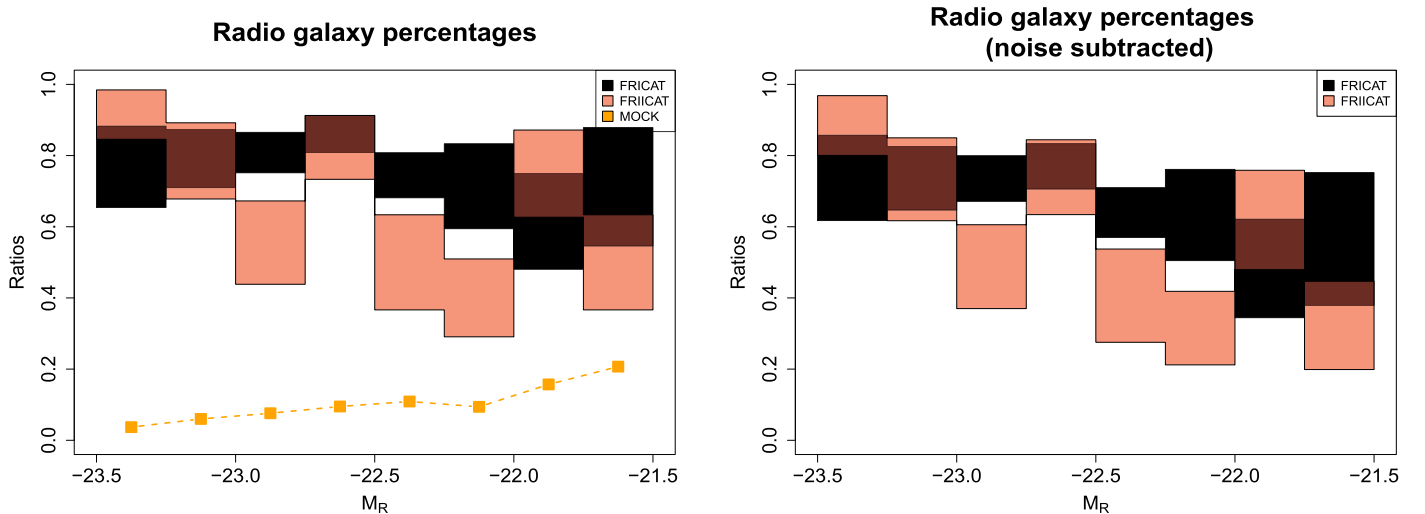
The most important characteristic of using the cosmological overdensities is that by selecting our thresholds on the basis of the MOCK catalog it is adaptive and it does not appear to have  $z_{\text{src}}$  dependence, being affected only by the SDSS spectroscopic completeness. Nevertheless, it is worth mentioning that the cosmological overdensity allows us to mitigate the bias due to the large number of galaxy clusters found at low redshift in the T12 catalog (see next sections for more details).

Finally, to prove that our results are independent of the clustering algorithms chosen to carry out the analysis, as they are of the thresholds of  $\Delta z$  and the 2 Mpc radius, we tried three additional methods, namely DBSCAN, Voronoi tessellation, and minimum spanning tree (MST). These are all clustering algorithms already used in large optical and infrared surveys to search for galaxy groups and clusters as spatial overdensities, as alternatives to the FoF method (Huchra & Geller 1982). For all these three methods we considered a radio galaxy to be in a galaxy-rich large-scale environment when the number of cosmological neighbors belonging to one cluster (i.e., a region of high density of optical sources), found by applying the algorithm, is larger than the top 5% of those detected in the MOCK sample by adopting the same method. We also run all three algorithms considering the number of





**Figure 8.** Left: ratio between the number of radio galaxies (FR Is in black and FR IIs in red) in comparison with MOCK sources (orange) living in galaxy-rich large-scale environments to their total number as a function of  $z_{\text{src}}$ . These fractions were estimated adopting (i) cross-matches with the T12 cluster catalog plus (ii) cosmological overdensities and (iii) associations with the GMBGC catalog. The rise of the fraction for the MOCK sources at low redshifts is due to the T12 cross-matches. Both classes of radio galaxies appear to follow the same trend of lying in a galaxy-rich large-scale environment. Confidence intervals for the ratios of radio galaxies in each redshift bin are estimated as described in Section 4. Right: same as left panel but with noise subtracted. We remark that there are no FR II radio galaxies in the first two redshift bins.



**Figure 9.** Same as Figure 8 where the ratios are expressed as a function of the absolute magnitude in the  $R$  band,  $M_R$ . The left panel reports the comparison with the results obtained in the MOCK catalog while the right panel refers to the ratios with noise subtracted as described in Section 4.

candidate elliptical galaxies instead of that of cosmological neighbors (see the Appendix for more details).

#### 4.4. Noise and Uncertainties

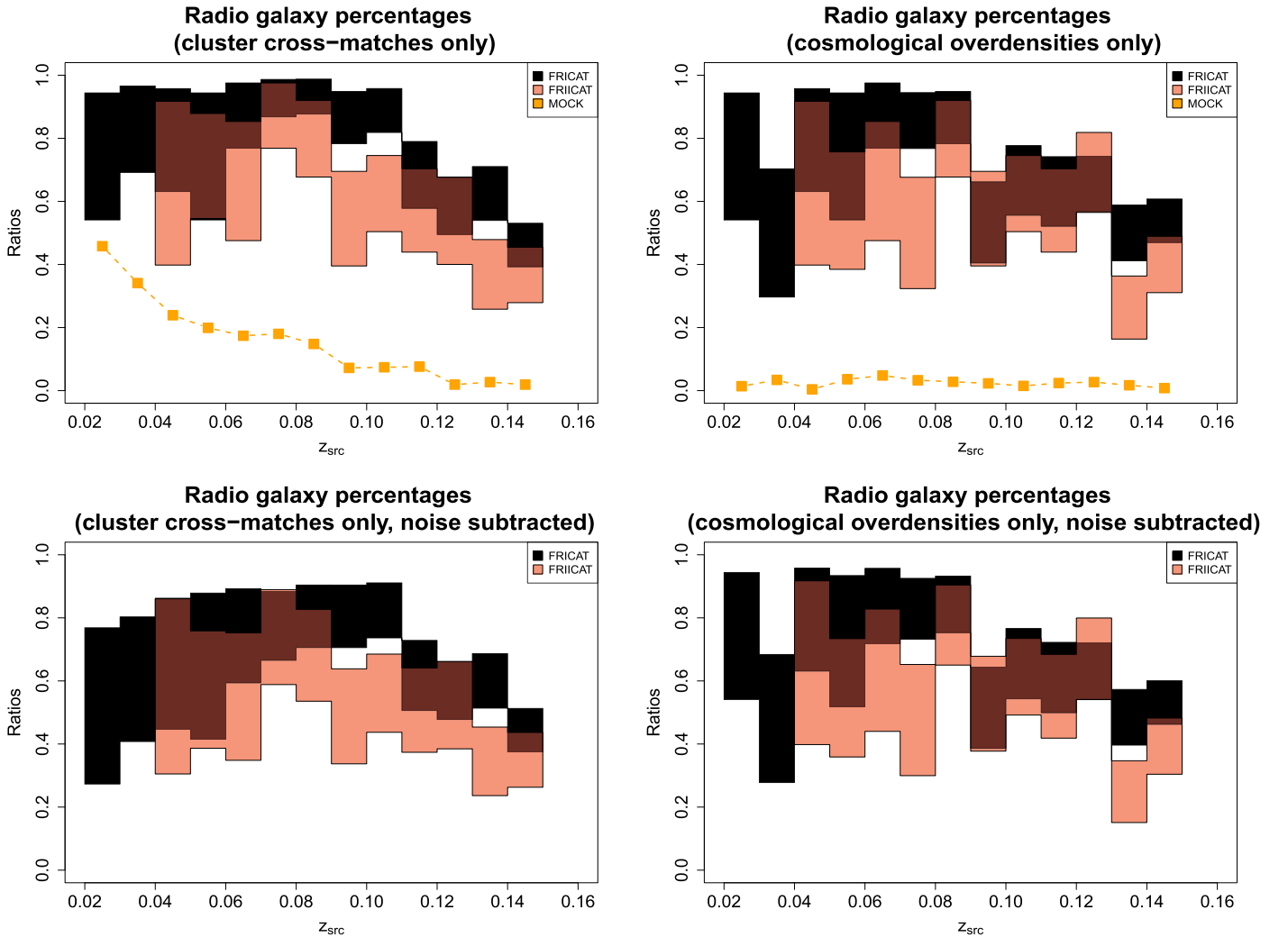
It is crucial to highlight that results on the MOCK catalog provide an estimate of the false positives we could get adopting our algorithms when claiming that a source belongs to a galaxy-rich large-scale environment. Testing our methods over the MOCK catalog helped us to estimate their “noise.”

Both FR I and FR II radio galaxy catalogs have a completeness larger than 90% in total, and almost 100% in the low-redshift bins. However, the SDSS footprint covers only  $\sim 1/3$  of the sky and we needed to estimate the uncertainties on the ratios/percentages of radio galaxies taking into account the underlying population. Thus, assuming a binomial distribution,

where, for each bin of redshift, magnitude, and/or luminosity, finding a radio galaxy in a galaxy-rich large-scale environment is a “success,” we computed binomial confidence intervals corresponding to  $1\sigma$ , adopting the procedure described by Cameron (2011).

In each plot where ratios of radio galaxies found in galaxy-rich large-scale environment are shown, we report both the comparison between confidence intervals and results obtained with the MOCK catalog as well as ratios with noise subtracted. To take into account the noise subtraction we simply define the number of “successes” as the number of radio galaxies found in a galaxy-rich large-scale environment per bin minus the average number of MOCK sources rescaled for the total number of radio galaxies in that bin.

For example, if the total number of radio galaxies in the range  $z_1 < z_{\text{src}} < z_2$  is  $n$ , the number lying in a galaxy-rich



**Figure 10.** Fraction of radio galaxies (FR Is marked in black and FR IIs in red) and MOCK sources (orange) in galaxy-rich large-scale environments as a function of redshift  $z_{\text{src}}$ . In the top left panel we show the ratios computed only adopting the cross-matching analysis with the T12 catalog of groups and clusters, and in the top right panel those calculated using only the cosmological overdensity. The efficiency of the former procedure decreases significantly with redshift while the latter is less affected by it. It is also evident how both methods show the gap between the fraction for real sources in galaxy-rich large-scale environments and the fraction for those in the MOCK catalog: the main result of our analysis. Lower panels show the same ratios but taking into account the noise subtraction as described in Section 4.

large-scale environment is  $k$ , and  $\langle k_m \rangle = 5.2$  is the average number of MOCK sources found in a galaxy-rich large-scale environment using the same procedure on  $N_s$  simulations computed with  $n$  sources, the number of successes used to compute the noise-subtracted confidence intervals is  $k' = k - \langle k_m \rangle$ , assuming that the uncertainty on the simulations is negligible due to their high number.

## 5. Results

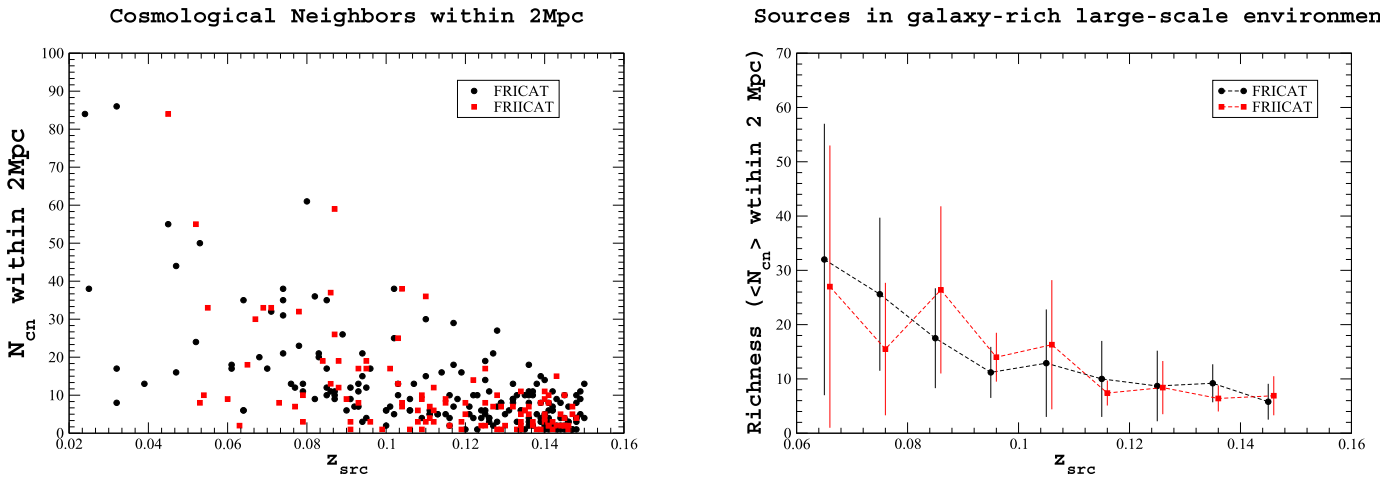
### 5.1. FR Is and FR IIs in Galaxy-rich Large-scale Environments

Results of our analysis are discussed here. The number of radio galaxies lying in galaxy-rich large-scale environments with respect to their total number per bin of redshift, and considering all the criteria previously described (i.e., cross-matches with cluster/group catalog, cosmological overdensities, cross-matches with the GMBCG), is shown in Figure 7. There are 29 FR Is out of 195 and 16 FR II out of 115 in galaxy-rich large-scale environments at  $z_{\text{src}} \leq 0.08$ , corresponding to the completeness limit of the T12 catalog. All these

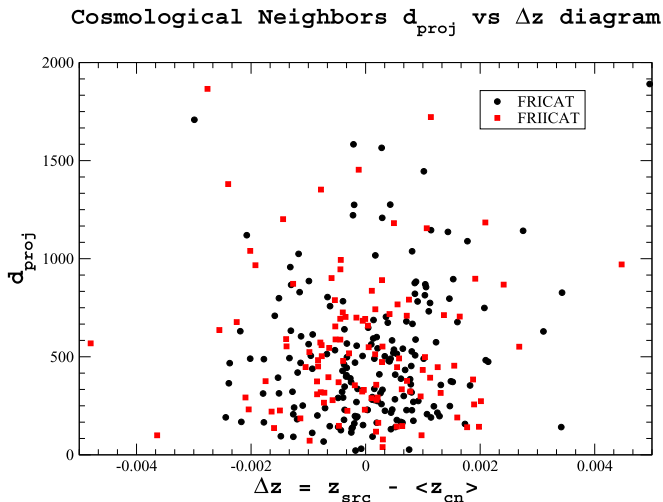
FR Is and 14 out of 16 FR IIs lie in galaxy-rich environments. The two FR IIs not belonging to galaxy-rich environments are optically classified as LERGs. The three FR II HERGs in our sample at  $z_{\text{src}} < 0.08$  all lie in galaxy groups/clusters. Ratios between the number of FR Is and FR IIs belonging to galaxy-rich environments with respect to their total number in the FRICAT and FRIICAT, respectively, are then shown in Figure 8 together with those in the MOCK sample to which the same criteria were applied.

In Figure 9 we also report the ratios between the number of radio galaxies (FR Is in black and FR IIs in red) and MOCK sources (orange) living in galaxy-rich large-scale environments to their total number as a function of the absolute magnitude in the  $R$  band,  $M_R$ .

It is worth highlighting that in Figure 8 the fraction of radio galaxies found in galaxy-rich large-scale environments decreases significantly with  $z_{\text{src}}$ . We therefore reanalyzed all radio galaxy samples, comparing results obtained using the T12 cluster cross-match procedure with those found when searching only for cosmological overdensities. The improvements given



**Figure 11.** Left: the total number of cosmological neighbors  $N_{\text{cn}}$  within 2 Mpc as a function of the central source redshift  $z_{\text{src}}$ . There are no differences between the two FR catalogs. The *richness* estimated according to the cosmological overdensity procedure appears the same. Right: the *richness* estimated according to the cosmological overdensity procedure as a function of redshift. The uncertainty on  $\langle N_{\text{cn}} \rangle$  is computed from the distribution of  $N_{\text{cn}}$  within the same  $z_{\text{src}}$  bin for each source class. The first bin is larger and includes all sources up to  $z_{\text{src}} = 0.065$ .



**Figure 12.** Projected distance  $d_{\text{proj}}$  as function of the redshift difference  $\Delta z$  between that of the central radio galaxy,  $z_{\text{src}}$ , and the average values of redshifts and coordinates computed with the sample of cosmological neighbors. FR I radio galaxies are marked with black circles and FR IIs are shown in red, as in all plots.

by the GMBCG cross-matches are negligible. Results from this comparison are shown in Figure 10.

It is clear how both methods show the gap between the fraction for real sources in galaxy-rich large-scale environments and the fraction for those in the MOCK catalog. However, it is also quite evident how the fraction of MOCK sources claimed to be in galaxy-rich large-scale environments rises when  $z_{\text{src}}$  decreases. This shows how the efficiency of cosmological overdensities is not dependent on  $z_{\text{src}}$ . On the other hand, this strong  $z_{\text{src}}$  dependence in the T12 cluster cross-matches turns into a higher probability of finding a source lying in a galaxy-rich large-scale environment at low  $z_{\text{src}}$ . This effect has to be taken into account when comparing our results with those available in the literature.

Despite the methods and procedures or criteria and thresholds adopted, which are the same for all samples and catalogs, our main result is that the fraction of FR IIs in rich

environments could appear systematically lower than that of FR Is but radio galaxies in both radio classes inhabit galaxy-rich large-scale environments in the local universe independently of their radio morphology.

### 5.2. Richness

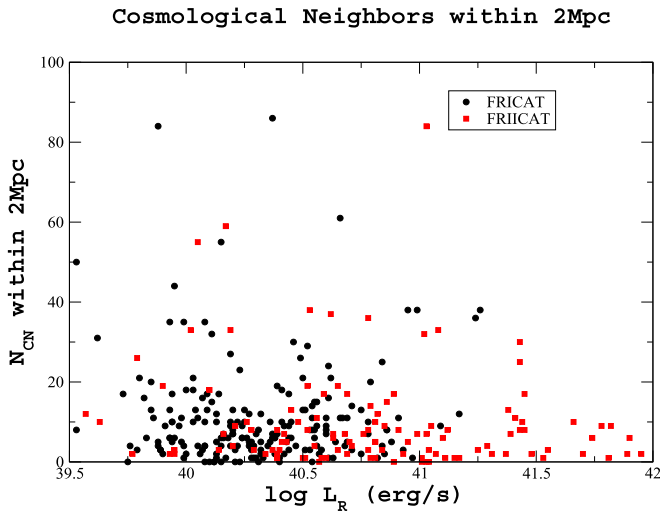
We test whether the richness of their environment is also the same. Since the galaxy density  $N_{\text{gal}}$  reported in the T12 catalog of groups and clusters could be misleading and underestimated, as shown in Figure 6, we proceeded as follows. We computed the total number of cosmological neighbors  $N_{\text{cn}}$  within 2 Mpc as a function of  $z_{\text{src}}$ , and we show in Figure 11 that there are no differences between FR I and FR II radio galaxies. In the same figure we also present the average number of cosmological neighbors  $\langle N_{\text{cn}} \rangle$  within 2 Mpc as a function of  $z_{\text{src}}$  to highlight the lack of differences between radio galaxies with different morphology.

Using the cosmological overdensities, we also computed the projected distance  $d_{\text{proj}}$  as a function of  $\Delta z$  between the redshift of the central galaxy (i.e.,  $z_{\text{src}}$ ) and the average values of redshifts  $\langle z_{\text{cn}} \rangle$  and of coordinates in the sample of cosmological neighbors. In Figure 12 FR Is do not appear different from FR IIs. Furthermore, it is clear how the threshold of  $\Delta z = 0.005$  is not extremely conservative and how both classes of radio galaxies lie closer to the centers of galaxy groups and clusters. Restricting this threshold to lower values (e.g., 0.003) does not affect our main results.

Finally, we also verified whether  $N_{\text{cn}}$  is related to the radio luminosity  $L_R$  at 1.4 GHz. As shown in Figure 13, even if the distribution of radio luminosity for the FRICAT and the FRIICAT is quite different (see, e.g., Capetti et al. 2017a, 2017b, for more details), for a given value of  $L_R$  we have similar value of  $N_{\text{cn}}$  for both classes.

### 5.3. LERGs and HERGs in Galaxy-rich Large-scale Environments

We then explored the large-scale environments and their richness adopting the optical classification. To carry out this analysis a couple of problems related to selection effects, which could introduce biases, should be properly mentioned. We



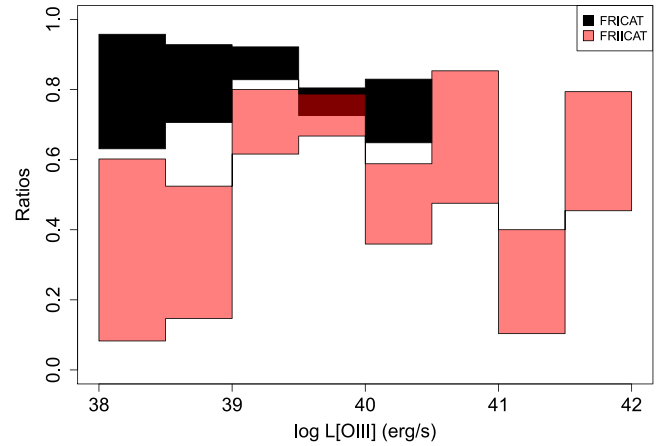
**Figure 13.** Same as Figure 11, where the total number of cosmological neighbors  $N_{\text{cn}}$  within 2 Mpc is reported as a function of the radio luminosity  $L_R$  at 1.4 GHz.

stress the fact that radio galaxies selected in flux-limited samples at high  $z_{\text{src}}$ , where it is difficult to find surrounding galaxy-rich environments, are mostly FR II HERGs, while at low  $z_{\text{src}}$  they are almost all FR Is and FR II LERGs, which are generally found in environments denser in galaxies than the former. The same situation occurs when considering radio and/or optical flux-limited surveys, where high-luminosity sources (i.e., mostly HERGs) will appear to inhabit a less dense environment than low-luminosity ones (i.e., LERGs). This will also appear as a function of the stellar mass ( $M_{\text{star}}$ ) since brighter sources also have higher values of  $M_{\text{star}}$ , which is generally estimated from the absolute magnitude (Ching et al. 2017; Miraghaei & Best 2017). Thus to test whether FR II HERGs and FR II LERGs live in environments with different galaxy density it could be useful to investigate this aspect as a function of redshift, as previously reported. This will also guarantee an analysis independent of the efficiency of clustering algorithms with  $z$  and of the cosmological evolution of the two source classes.

We then plotted the ratio of FR Is and FR IIs in galaxy-rich environments to their total number as a function of the [O III] luminosity  $L_{[\text{O III}]}$  in Figure 14. At higher values of  $L_{[\text{O III}]}$  (i.e., above  $\sim 10^{40} \text{ erg s}^{-1}$ ) most FR IIs are HERGs, but again it seems that their fraction in galaxy-rich environments does not depend strongly on  $L_{[\text{O III}]}$ . The small difference for  $\log L_{[\text{O III}]} < 39 \text{ erg s}^{-1}$  is simply due to the smaller number of FR IIs in the FRIICAT. For the sake of completeness we also show in Figure 15 the total number of cosmological neighbors  $N_{\text{cn}}$  within 2 Mpc as a function of the [O III] luminosity:  $L_{[\text{O III}]}$  and not neat differences appear in the range  $40 < \log(L_{[\text{O III]})/\text{erg s}^{-1}) < 40.5$  between the FR Is and FR IIs, where radio galaxy catalogs include both LERGs and HERGs.

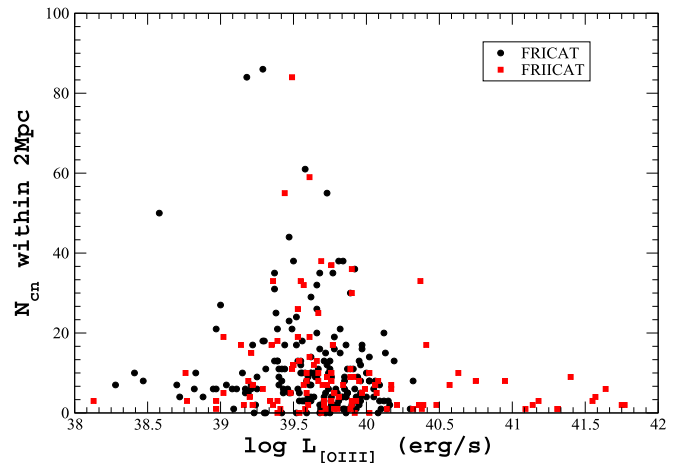
It is worth highlighting that our study is based on extremely homogeneous samples with respect to other analyses present in the literature, but being restricted to the local universe (i.e.,  $z_{\text{src}} \leq 0.15$ ) the radio galaxy catalogs include only a limited number of HERGs (only 14, all in the FRIICAT). Thus results of a comparison of LERGs and HERGs have to be treated with caution, being less statistically strong.

### Radio galaxy percentages



**Figure 14.** Same as Figures 8 and 9 where ratios are expressed as a function of the [O III] luminosity  $L_{[\text{O III}]}$ . This allows us to highlight the HERGs in the FRIICAT that are at high values of  $L_{[\text{O III}]}$ .

### Cosmological Neighbors within 2Mpc



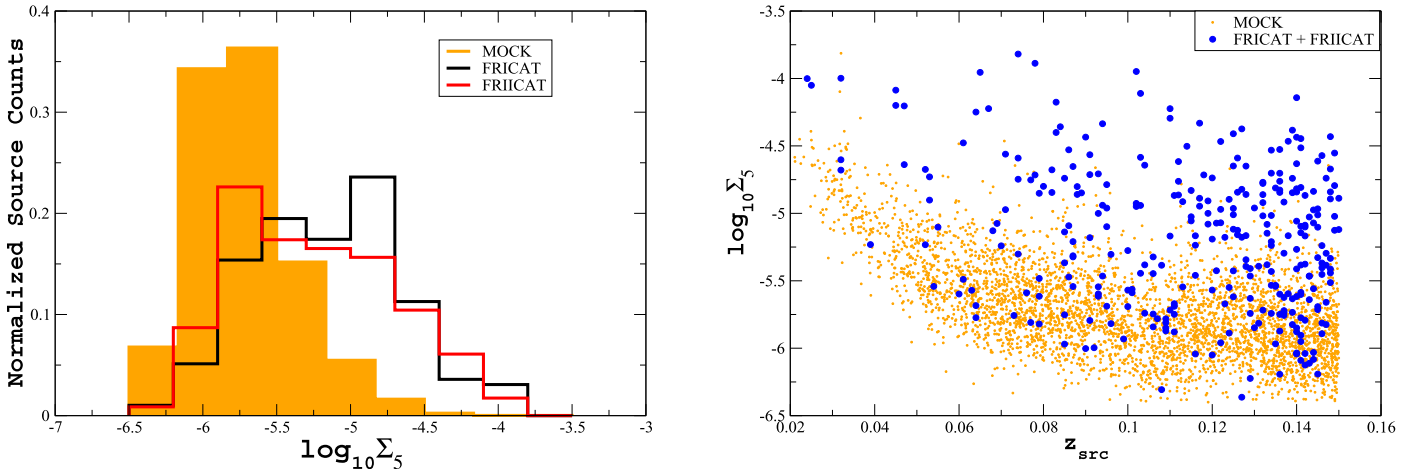
**Figure 15.** Same as Figure 13 where the total number of cosmological neighbors  $N_{\text{cn}}$  within 2 Mpc is reported as a function of the [O III] emission-line luminosity  $L_{[\text{O III}]}$ .

The total number of FR II LERGs lying in galaxy-rich large-scale environments at  $z_{\text{src}} < 0.11$  is 31 out of a total of 40 (i.e., 77%), similar to FR II HERGs in the same redshift range. Below  $z_{\text{src}} = 0.11$  the total number of HERGs in galaxy-rich environments is five out of a total of seven, while this fraction decreases to two out of the seven at higher redshifts.

Using the GMBCG we found that 33 FR Is (17%) appear associated with a cluster hosting a BCG candidate, and lie at  $\Delta z \leq 0.005$  as estimated by using only spectroscopic redshifts. They also belong to galaxy clusters with more than eight members. In particular, 10% of their total number lie at projected distances of less than 1 kpc, and are themselves the BCG candidates. A similar situation occurs again for FR IIs. Eighteen (16% of the FRIICAT) belong to a BCG cluster, as previously stated, and 11 (10%) are BCG candidates. There are no FR II HERGs that are positionally associable with a BCG listed in the GMBCG, even if HERGs are always the most luminous sources of all their cosmological neighbors.



### The $\Sigma_5$ distribution



**Figure 16.** Left: normalized distribution of the  $\Sigma_5$  parameter, i.e., the fifth nearest neighbor density, tracing the dark matter halo density in large-scale environments (Sabater et al. 2013; Worpel et al. 2013). It is evident how radio galaxies tend to have, on average, larger values of  $\Sigma_5$  than those computed for fake sources in the MOCK catalog. Right: log of the  $\Sigma_5$  parameter as a function of the central source redshift  $z_{\text{src}}$ . In this plot radio galaxies (blue circles) are reported together with MOCK sources (orange circles). It is clear how the  $\Sigma_5$  estimator is affected by a redshift dependence due to the Malmquist bias.

#### 5.4. The $\Sigma_k$ Comparison

We also investigated the distribution of the  $\Sigma_k$  parameter, i.e., the  $k$ th nearest neighbor density (see, e.g., Best 2004), for both radio galaxies and MOCK sources.

The  $\Sigma_k$  parameter is defined as the ratio between the number of sources  $k$  and the projected area  $\pi r_k^2$ , where  $r_k$  is the projected distance between the central galaxy and the  $k$ th nearest neighbor. We computed it for  $k = 5$  (i.e.,  $\Sigma_5$ ), adopting the distance (in kiloparsecs) between the central galaxy and the fifth closest *candidate elliptical galaxy* (Ching et al. 2017).

As discussed extensively in the literature, this parameter can be used as a beacon to trace the dark matter halo density (Sabater et al. 2013; Worpel et al. 2013) and it also appears to correlate with halo mass (Haas et al. 2012). According to all previous analyses the distribution of  $\Sigma_5$  as reported in Figure 16 (left panel) also shows that both FR Is and FR IIs live in galaxy-rich large-scale environments, with larger values of dark matter halo density than random MOCK sources (see Figure 16). However, the right panel of the same figure marks a  $z_{\text{src}}$  dependence of  $\Sigma_5$ , possibly due to the Malmquist bias, which could affect the use of this estimator.

## 6. Summary and Conclusions

We have presented a detailed statistical analysis of the large-scale environments of radio galaxies. The main advantages of our study, with respect to those previously carried out in the literature, are (i) the sample selection, which is extremely homogeneous over a wide range of frequencies, and (ii) the large variety of clustering algorithms adopted for our analysis, which always provide consistent results.

In particular, for the radio galaxies, we used the FRICAT and FRIICAT catalogs, complete at a level of confidence higher than 90%. Thanks to their selection criteria these catalogs are not contaminated by compact radio objects, such as compact steep-spectrum sources and FR 0 (Baldi et al. 2015, 2018), which

show a different cosmological evolution to FR Is and FR IIs and potentially lie in different environments.

We proved that identifying a galaxy-rich environment only using a cross-matching analysis with a catalog of groups and clusters can introduce biases due to the redshift dependence of the algorithm (e.g., FoF) used to build it. Thus our analysis was carried out by performing a direct search around radio galaxies and using them as beacons.

We investigated the large-scale environment of radio galaxies, adopting both their radio morphological classification (FR I versus FR II; Fanaroff & Riley 1974) and the optical spectroscopic classification (HERG versus LERG; Hine & Longair 1979). This allowed us to search for a link between their environment, their radio extended structure, and their accretion modes. However, it is worth mentioning that due to the selection criteria of our radio galaxy catalogs, the limited number of HERGs listed in the FRIICAT does not allow us to firmly compare our results with those present in the literature.

Our main result is that, in the local universe, FR Is and FR IIs as well as HERGs (all FR IIs in our catalog) and LERGs live in galaxy-rich large-scale environments that have the same richness, independently of the redshift range considered or their radio luminosity or absolute magnitude. This is also independent of the thresholds and algorithms used to identify galaxy-rich large-scale environments.

More than 70% of the FR Is and more than 55% of all FR IIs in our catalogs lie in galaxy-rich large-scale environments. The probability of finding an FR I lying in a dense neighborhood appears larger than that for FR IIs. This claim could also be biased by the possible presence of fossil groups where FR IIs could reside, and that could be only revealed with X-ray observations. As previously stated, the numbers of HERGs in our samples prevent us from drawing a strong statistical conclusion, when considering the optical classification, even if five out of seven HERGs up to  $z_{\text{src}} = 0.11$  lie in groups/clusters of galaxies and this fraction decreases to two out of seven at higher redshifts. On the other hand, finding HERGs in

denser environments is consistent with optical observations carried out on high-redshift radio galaxies, generally used as beacons to search for protoclusters (see, e.g., Miley & De Breuck 2008 for a recent review), which appear to be all HERGs (see, e.g., Röttgering et al. 1997; De Breuck et al. 2001, 2006; Jarvis et al. 2001; Bornancini et al. 2007).

We also found that  $\sim 17\%$  of the FR Is are associated with a cluster hosting a BCG candidate and lie within  $\Delta z \leq 0.005$ , estimated by using only spectroscopic redshifts reported in the GMBCG catalog. All these FR Is belong to galaxy clusters with more than eight members, and 20 of them (10% of the total number) are themselves the BCGs. A similar situation occurs again for FR II radio galaxies, where  $\sim 16\%$  belong to a BCG cluster and 10% are the BCGs. There are no FR II HERGs in our sample associated with a BCG when considering the GMBCG. However, all HERGs in our sample that lie in galaxy-rich large-scale environments are indeed BCGs according to the luminosity distribution of their cosmological neighbors.

Finally, it is worth highlighting that to carry out our analysis we developed a method based on the number of counts of cosmological neighbors in the large-scale environments of selected sources, measuring cosmological overdensities. This method has been also successfully tested and compared with several clustering algorithms generally used to perform blind searches of galaxy groups and clusters in large optical and infrared surveys.

## 7. Literature Comparison

The comparison between the results achieved in our analysis, even if limited to the local universe (i.e.,  $z_{\text{src}} \leq 0.15$ ), and the claims present in the literature (all reported below in italics) can be summarized as follows.

1. *HERGs are found almost exclusively in low-density environments, while LERGs occupy a wider range of densities, independent of FR morphology (Gendre et al. 2013).*

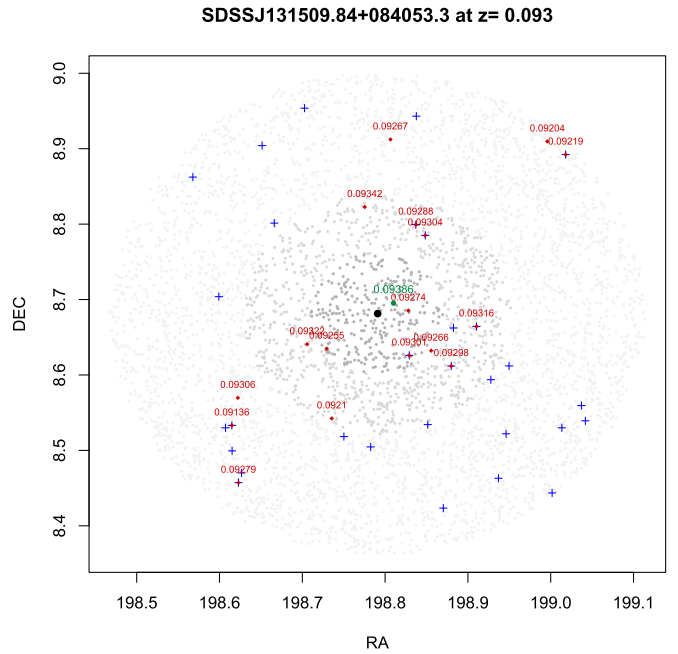
Indeed we proved that LERGs and HERGs in the local universe (at least up to  $z_{\text{src}} \leq 0.1$ ) live in galaxy-rich large-scale environments with same richness. This conclusion is based on the number of cosmological neighbors within 2 Mpc (see Figure 11). At higher redshifts it was not possible for us to establish a firm conclusion given the low number of HERGs in our catalogs. Our result is consistent with X-ray observations of FR IIs (see, e.g., Hardcastle & Worrall 2000).

2. *There is a significant overlap in the environment between LERGs and HERGs, and no clear driving factor between the FRI and FR II sources is found even when combining radio luminosity and accretion mode (Gendre et al. 2013).*

We confirmed this statement.

3. *FRI radio galaxies lie in higher-density environments, on average, than FRIIs (Miraghaei & Best 2017).*

On the contrary, we showed that fraction of FR Is living in galaxy-rich large-scale environments could be slightly larger than that of FR IIs, but the richness of their environments is certainly consistent at all redshifts sampled by our analysis (see right panel of Figure 11). The difference with respect to literature works could be due to radio sources that are contaminants of selected samples, a bias that does not affect our radio galaxy catalogs. The



**Figure 17.** Large-scale environment of SDSS J131509.84+084053.3, a radio galaxy classified as FR II HERG. The total number of cosmological neighbors within 1 Mpc (i.e., 11) is a clear example of a HERG associated with a galaxy group/cluster at  $z_{\text{src}} = 0.093$ . Gray circles mark all SDSS sources within a 2 Mpc distance from the central radio galaxy, while blue crosses mark the candidate elliptical galaxies in the same field, and red circles with the redshift reported above are cosmological neighbors.

major difference with respect to literature analyses, which strengthens our result, is that the sample selection carried out for both radio galaxy catalogs is extremely homogeneous when considering their radio morphology.

4. *The environments of LERGs display higher density than those of HERGs (Miraghaei & Best 2017).*

No differences were found in the environments of LERGs and HERGs in our analysis. The case of the FR II HERG SDSS J131509.84+084053.3 with 11 cosmological neighbors within 1 Mpc is shown in Figure 17 as an example of a galaxy-rich environment. However, we again highlight that our claim is limited to  $z_{\text{src}} \leq 0.15$ , where the number of HERGs is only a tiny fraction of the whole FRIICAT (i.e.,  $\sim 10\%$  of the FRIICAT), but it does not suffer from possible selection effects due to different cosmological evolution of HERGs and LERGs.

5. *High-luminosity radio galaxies with weak or no emission lines (LERGs) lie in more massive haloes than non-radio galaxies of similar stellar mass and color (Ching et al. 2017). The HERGs are typically in lower-mass haloes than LERGs.*

The distribution of the  $\Sigma_5$  parameter for our radio galaxy catalogs is not in agreement with this statement, at least in the  $z_{\text{src}}$  range considered. As in the previous case, we remark that our result, even if based on a statistically homogeneous sample selection, is limited by the number of HERGs in the radio galaxy catalogs (i.e., about 10% of all FR II sources).

6. *At low redshifts, there is a correlation between radio luminosity and the cluster environment for LERGs but not for HERGs (Ineson et al. 2015).*

No difference in the richness of their environments was indeed found as a function of their  $L_R$ , not even

between FR Is and FR IIs. However, we also remark that the sample selected by Ineson et al. (2015) spans a wider range of radio luminosities than our radio galaxy catalogs, which could better reveal trends between radio power and richness. It is worth highlighting that in our analysis the richness is estimated by the number of cosmological neighbors instead of using X-ray observations. The advantage of our approach is that it is difficult to get X-ray observing time for large, homogeneously selected samples of sources. This is challenging since good spatial resolution is needed to separate diffuse X-ray emission surrounding radio galaxies due to the intergalactic medium (IGM) from the extended emission, mainly detected along the radio axis, due to inverse Compton scattering of seed photons arising from the cosmic microwave background in radio lobes (Scharf et al. 2003; Celotti & Fabian 2004; Erlund et al. 2007; Smail et al. 2012; Massaro et al. 2013b, 2018; Stuardi et al. 2018).

Finally, it is worth mentioning that, in the future, we could also extend the radio galaxy catalogs at redshifts larger than 0.15, losing a small fraction of their completeness, but making our analyses more suitable for comparison with others present in the literature (e.g., Ineson et al. 2013, 2015).

## 8. Future Perspectives

Our analysis is only based on radio and optical information, while X-ray observations will be crucial to obtain a more complete view of the large-scale environments of radio galaxies. By carrying out these observations we could perform a unique, unbiased, survey of radio galaxies that is extremely homogeneous in terms of source selection.

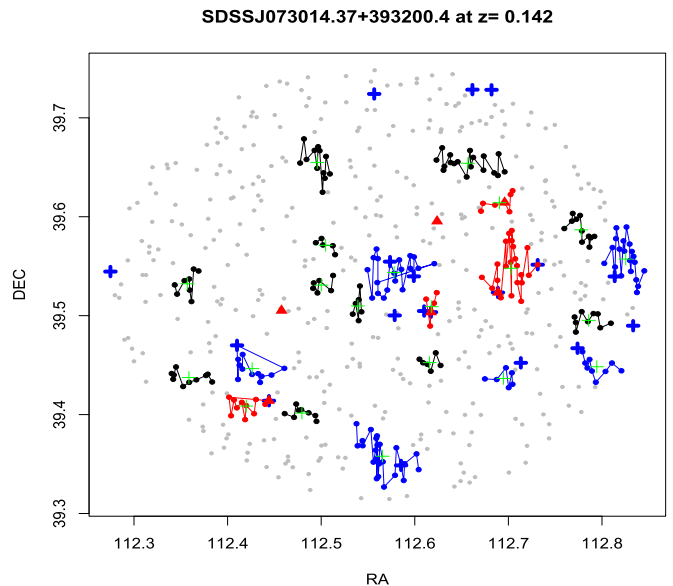
An X-ray survey will allow us to: (1) know the real fraction of FR Is lying in groups or existing in galaxy clusters, taking into account fossil groups; (2) measure mass, temperature, and luminosity of the IGM; (3) determine the location of the radio galaxy with respect to the center of the group/cluster—information immediately obtainable from the IGM distribution traced in the X-rays; (4) discover the fraction of radio galaxies lying in cooling core groups/clusters and test whether there is a gradient of temperature close to the radio galaxy, revealing active feedback processes, to name a few examples.

We thank the anonymous referee for useful comments that led to improvements in the paper. F.M. also wishes to thank Prof. M. Paolillo for his suggestions on the calculation of the binomial confidence intervals and Dr. C. C. Cheung for valuable discussions on this project initially planned during the IAU 313 on the Galapagos islands. This work is supported by the “Departments of Excellence 2018–2022” Grant awarded by the Italian Ministry of Education, University and Research (MIUR) (L. 232/2016). This research has made use of resources provided by the Compagnia di San Paolo for the grant awarded on the BLENV project (S1618\_L1\_MASF\_01) and by the Ministry of Education, Universities and Research for the grant MASF\_FFABR\_17\_01. This investigation is supported by the National Aeronautics and Space Administration (NASA) grants GO4-15096X, AR6-17012X, and GO6-17081X. F.M. acknowledges financial contribution from the agreement ASI-INAF n.2017-14-H.0. Funding for SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation,

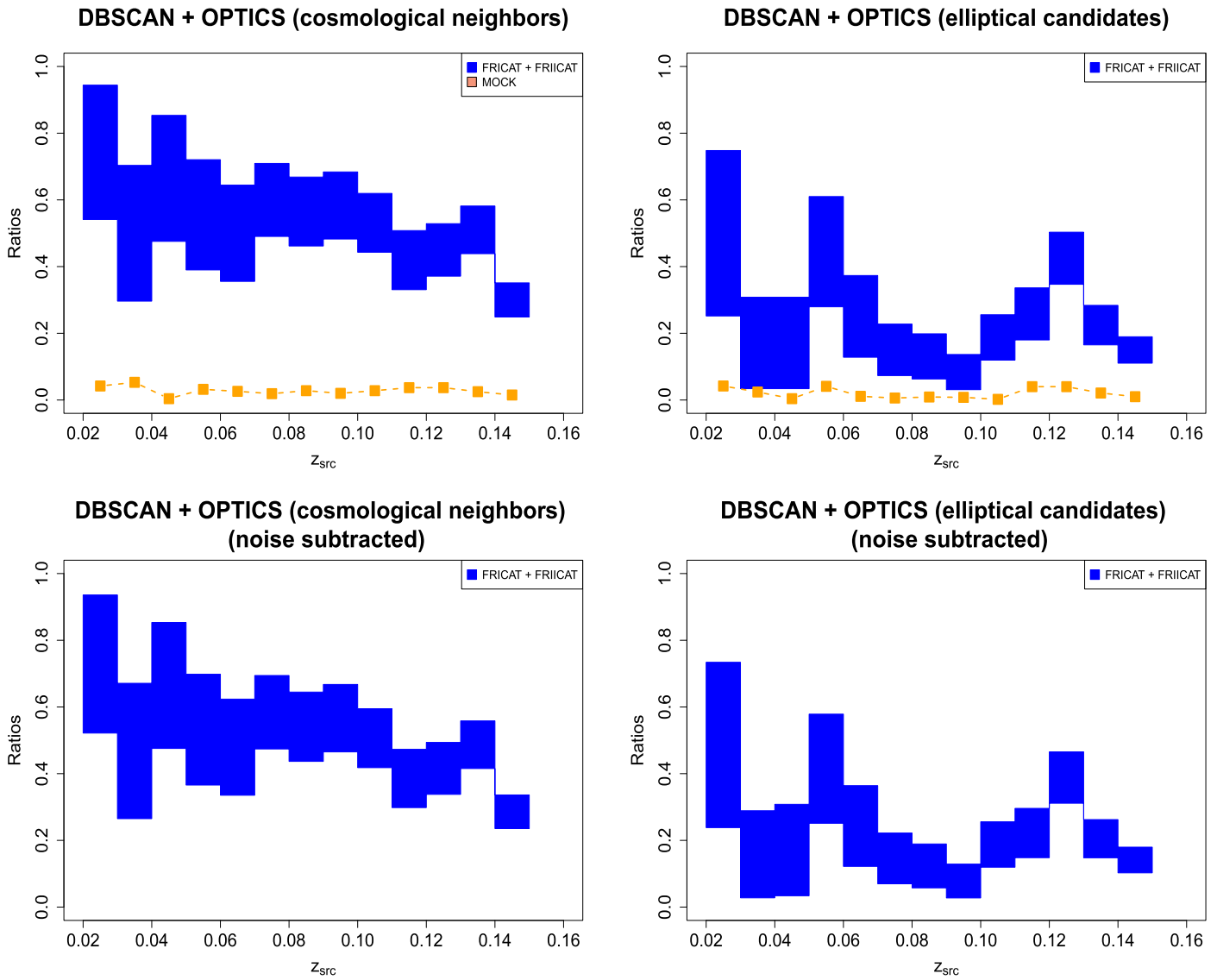
the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium for the 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. TOPCAT and STILTS astronomical software (Taylor 2005) were used for the preparation and manipulation of the tabular data and the images.

## Appendix Comparison with Other Clustering Algorithms

We carried out our analysis also considering several clustering algorithms, generally adopted to search for groups and clusters of galaxies in large optical surveys. This allowed us (i) to claim that our analysis is independent of the method adopted and (ii) to test efficiency and/or biases and/or limits of different methods.



**Figure 18.** Galaxy-rich large-scale environment detected by the DBSCAN algorithm optimized with the OPTICS procedure for a radio galaxy in our sample. We marked in red those clusters for which one of its members is a cosmological neighbor as previously defined, and in blue those hosting a candidate elliptical galaxy, while black clusters are those composed of simple SDSS galaxies in the field. Cluster centers are indicated with green crosses. The source analyzed lies in the center of the field, and in this case it belongs to a cluster with at least a candidate elliptical galaxy. Gray circles in the background mark the position of all SDSS optical sources classified as galaxies in the field of a 2 Mpc distance from the central radio galaxy.



**Figure 19.** Fractions of radio galaxies found in galaxy-rich environments computed with the DBSCAN+OPTICS algorithm only. The gap between real and fake sources rises clearly when the definition of galaxy-rich large-scale environments depends on the number of cosmological neighbors (top left panel), while it is only evident at  $z_{\text{src}}$  larger than 0.1 when using candidate elliptical galaxies (top right panel). Lower panels show the same ratios but taking into account the noise subtraction as described in Section 4.

The clustering algorithms adopted for our comparison are (i) density-based spatial clustering of applications with noise (DBSCAN), (ii) Voronoi tessellation, and (iii) minimum spanning tree (MST). These have been used in addition to the analyses carried out with the T12 cross-matches and/or the cosmological overdensities. As in the two previous cases, the level of significance of all clustering algorithms was determined on the basis of results achieved on the MOCK sample.

In this comparison analysis we searched for source overdensities hosting cosmological neighbors and/or a *candidate elliptical galaxy*, defining galaxy-rich large-scale environments as those above a certain threshold optimized as described below. As a last, additional, test we also run the KDE estimator on the candidate elliptical galaxies within 2 Mpc of each radio galaxy.

The gap between real and MOCK sources rises clearly when the definition of galaxy-rich large-scale environments depends on the number of cosmological neighbors, while it is only

evident at redshifts larger than  $\sim 0.1$  when using candidate elliptical galaxies. This occurs for almost all clustering algorithms adopted due to a large scatter in the optical colors of elliptical galaxies at low redshifts. However, it was crucial to test the methods also using candidate elliptical galaxies for future applications of proposed procedures at higher redshifts than  $z_{\text{src}} = 0.15$ , where the SDSS spectroscopic coverage decreases significantly.

Finally, we note that our additional tests with the clustering algorithms were run keeping FRICAT and FRIICAT separated but results are shown here for the whole sample of radio galaxies given their similar galaxy-rich large-scale environment.

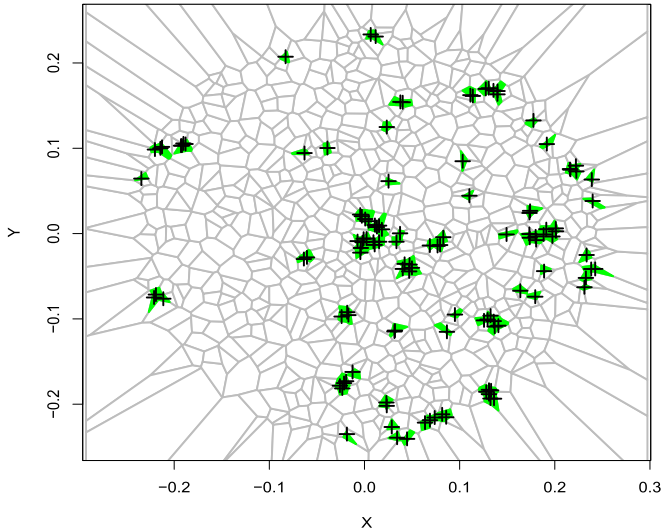
#### *Density-based Spatial Clustering of Applications with Noise (DBSCAN)*

Density-based spatial clustering of applications with noise (DBSCAN, Ester et al. 1996) is a supervised clustering algorithm able to locate regions of high source density. The



SDSS J075506.67+262115.9 at  $z = 0.123$ 

## Voronoi Tessellation



**Figure 20.** Voronoi cells built for SDSS J075506.67+262115.0 at  $z_{\text{src}} = 0.123$  are shown in gray, while regions with area smaller than the average value of the top 5% computed on 100 replicas of the field built assuming the same number of galaxies, with a uniform distribution, are marked in black. Green crosses indicate the location of sources in the high-density cells.

main advantage of this algorithm is the possibility of discovering source clusters of an arbitrary shape by handling the noise. DBSCAN depends on only two parameters:  $\epsilon$ , the maximum radius of a neighborhood, and  $k$ , the minimum number of points within a  $\epsilon$ -neighborhood. A source cluster is defined as a maximal set of density-connected points. On the other hand, its major drawback is that results are highly dependent on the choice of  $\epsilon$  and  $k$ . To avoid this problem we adopted the OPTICS procedure (Ankerst et al. 1999), proposed as an implementation of DBSCAN, to overcome the difficulty with this initial parameter selection.

The OPTICS procedure works by implementing the DBSCAN algorithm for an infinite number of distance parameters  $\epsilon_i$ , which are smaller than a generating distance. This implementation does not produce source clusters explicitly, but generates a so-called *reachability plot*, i.e., an ordering of the data objects representing the density-based clustering structure. In the *reachability plot* points belonging to a source cluster show up as *valleys*; the deeper the valley the denser the source cluster is. Thus we adopted as the input  $\epsilon$  for DBSCAN the mean of 75% all local maxima in the *reachability plot* and then applied the DBSCAN algorithm to locate source clusters.

In Figure 18 we show as connected points all source clusters found combining the DBSCAN algorithm with the OPTICS implementation; if a *cosmological neighborhood* is included within a source cluster found, the cluster itself is marked in red, while blue source clusters are those to which candidate elliptical galaxies belong; all others are simply shown in black.

In order to consider a source as surrounded by a galaxy-rich large-scale environment we set thresholds on the number of cosmological neighbors or candidate elliptical galaxies lying within a cluster detected by applying the DBSCAN+OPTICS algorithm to the MOCK sample within the top 5% of cases, as shown in Figure 19.

Voronoi tessellation (Lee & Schacter 1980) is a clustering algorithm that, in partitioning a considered region, creates so-called Voronoi cells on the basis of the distances between points (i.e., sources) present in the region itself. The area of each Voronoi cell is inversely proportional to the source density in the neighborhood (i.e., smaller areas correspond to regions of higher source density).

In Figure 20 we show an example of the Voronoi cells computed for the radio galaxy SDSS J075506.67+262115.0 at  $z = 0.123$ . To mark areas with high source density for each object analyzed, we first counted the number of galaxies within a radius of 2 Mpc and then we simulated 100 replicas of that region, assuming a uniform galaxy distribution. We set a threshold for the area of the Voronoi cells equal to the top 5% of the simulated fields as a galaxy-rich region.

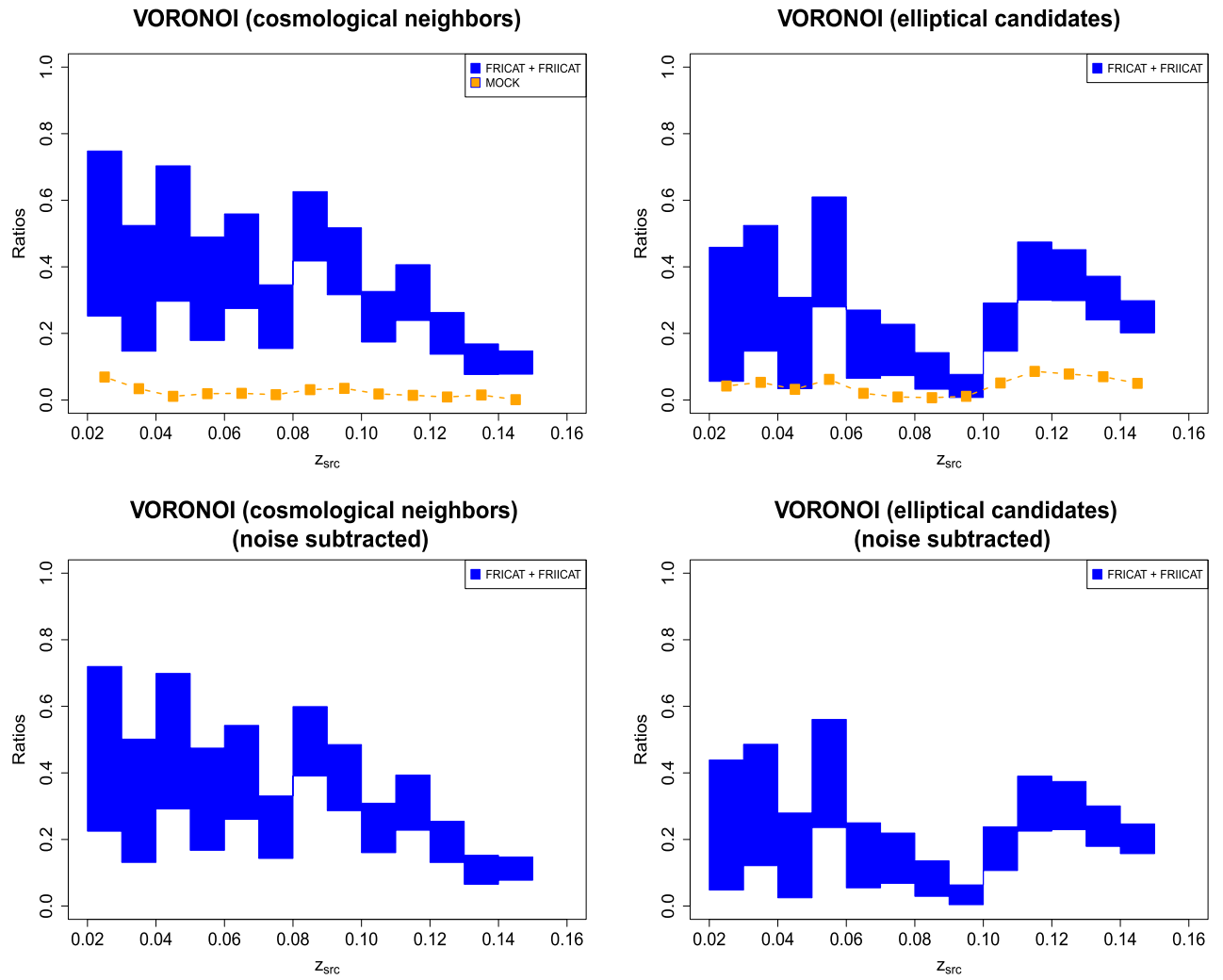
Then, as adopted for the previous methods, we considered as galaxy-rich large-scale environments those Voronoi cells having a number of cosmological neighbors or candidate elliptical galaxies larger than the top 5% of those detected in the MOCK sample. Results of this algorithm are shown in Figure 21, where the gap between real and fake sources is quite evident using cosmological neighbors, while it becomes less significant, in particular at  $z_{\text{src}} > 0.1$ , when using candidate elliptical galaxies.

This clustering algorithm is among those most used in searching for galaxy clusters/groups in photometric and spectroscopic surveys (see, e.g., Ramella et al. 2001). It generally uses both galaxy positions and magnitudes to find clusters as significant density fluctuations above the background. As applied here it is a non-parametric procedure and does not apply any smoothing of the data set.

## The Minimum Spanning Tree

The MST is a clustering algorithm used to search for candidate sources in gamma-ray images (Campana et al. 2008, 2013), as recently occurred for DBSCAN (Tramacere & Vecchio 2013), but it is also used to search for groups/clusters of galaxies in photometric surveys (Barrow et al. 1985). Photon arrival directions or galaxy positions are treated as the nodes of a two-dimensional graph, over which the tree with the minimal length is constructed. Edges with a length (in this case: angular distance) larger than the MST average value are removed, leaving several disconnected source clusters, which are further selected by their characteristics. Parameters used for the selection are the number of nodes  $N$  of the cluster, its clustering degree  $g$  (i.e., the ratio between the mean edge length in the MST and the average edge length in the local cluster) and the magnitude  $M$  (defined as  $M = gN$ ).

As in the previous cases we chose the case of the MST procedure applied to one radio galaxy (see Figure 22). Then, to consider a source as surrounded by a galaxy-rich large-scale environment, thresholds were chosen by applying the MST algorithm to the MOCK sample and selecting the number of cosmological neighbors or candidate elliptical galaxies of the top 5% of cases. Results of the MST procedure are shown in Figure 23.

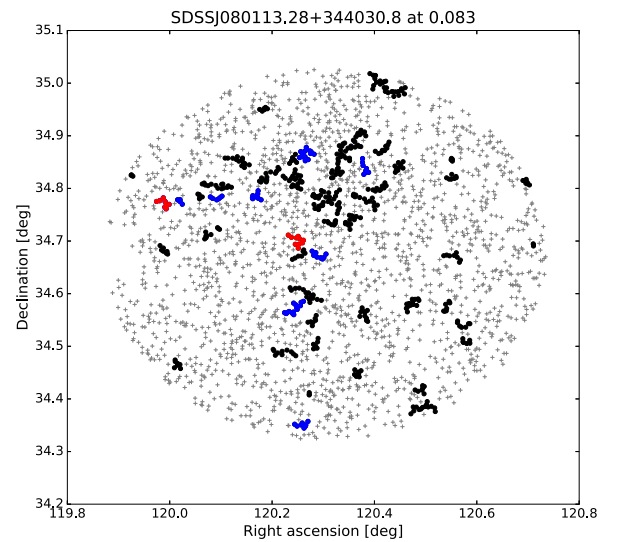


**Figure 21.** Top left and top right panels are the same as Figure 19, where source fractions are computed with the Voronoi tessellation algorithm only. The gap between real and fake sources is again clear when the definition of galaxy-rich large-scale environments depends on the cosmological neighbors present in each Voronoi cell, while it is only evident at  $z_{\text{src}} > 0.1$  when counting candidate elliptical galaxies (left panel). Lower panels show the same ratios but taking into account the noise subtraction as described in Section 4.

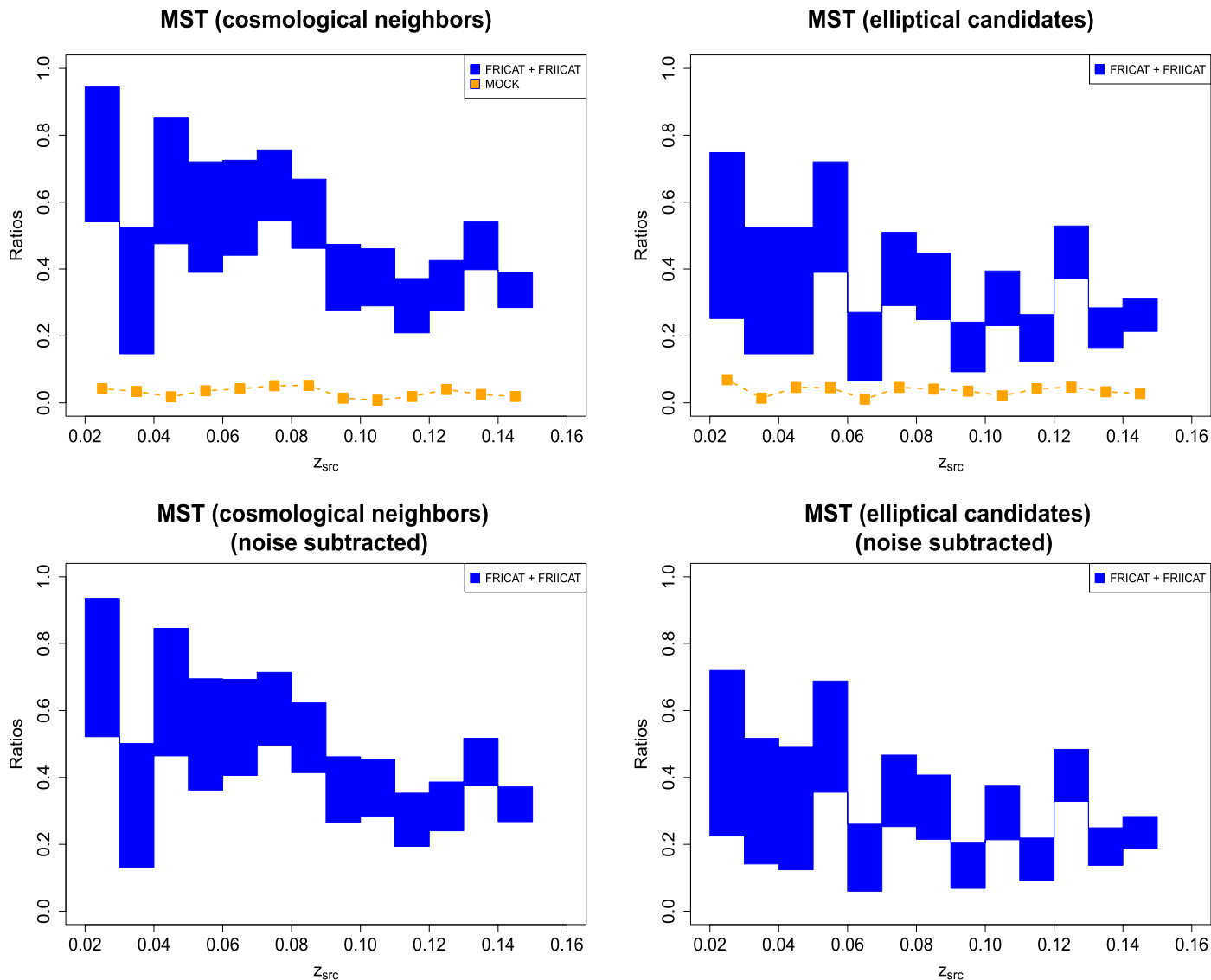
#### Galaxy Density Estimated with Kernel Density Estimation

Kernel density estimation (KDE) is a method that provides an effective procedure to estimate the probability function of a multivariate variable without any assumption on the shape of the *parent* distribution (Richards et al. 2004). KDE divides the data set into a square grid and convolves the discrete data with a kernel function, estimating the density for each one of them. Isodensity contours drawn from its application and associated with 20% of the highest density level are plotted in Figure 24 for the radio galaxy SDSS J073505.25+415827.5 at  $z_{\text{src}} = 0.087$ . All sources lying inside the top 20% contour level are marked in black while contours are labeled in green. Background and foreground galaxies in the field are also shown in gray. This is also an example where the central source belongs to one of the regions of high galaxy density.

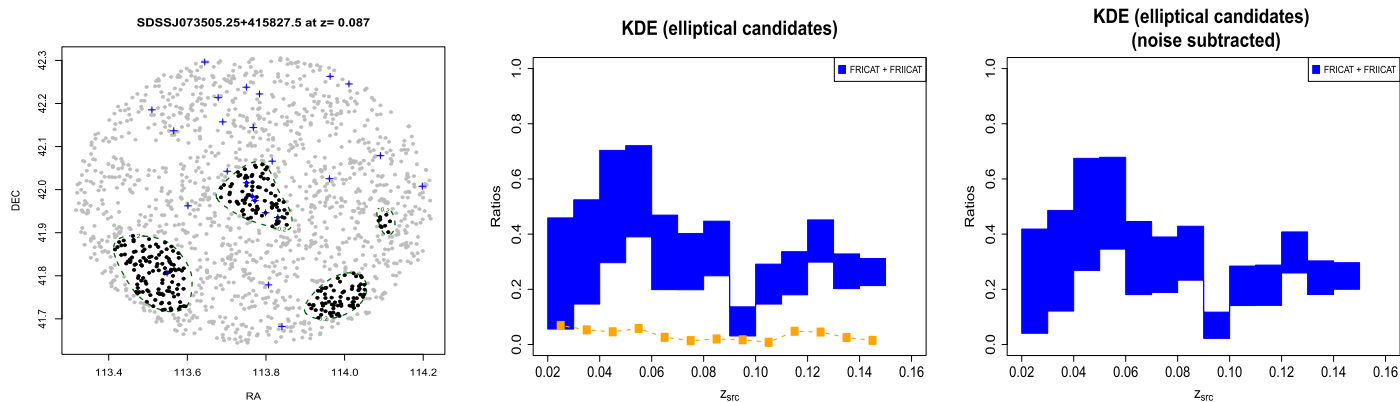
The threshold chosen to consider a source surrounded by a galaxy-rich large-scale environment was arbitrarily set to the top 20% of regions of high source density found applying the KDE algorithm and including at least the number of candidate elliptical galaxies present in the top 5% of the distribution of the MOCK sources, for the same  $z$  bin. The fractions of sources lying in galaxy-rich large-scale environments are plotted in



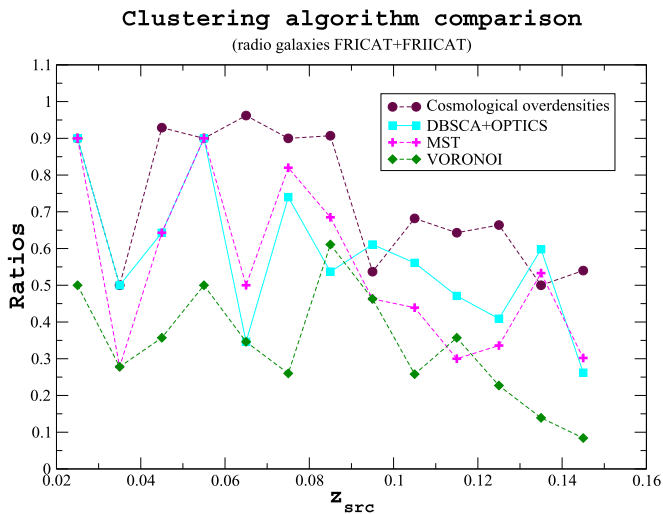
**Figure 22.** Clusters detected using the MST algorithm for the 2 Mpc region of SDSS J080113.28+344030.8 at  $z = 0.083$ . Those including at least one cosmological neighbor are shown in red, while those with at least one candidate elliptical galaxy appear in blue; the remaining ones are marked in black and overlaid on the background and foreground galaxies in the field shown in gray.



**Figure 23.** Top left and top right panels are the same as Figures 19 and 21, where source fractions are computed with the MST algorithm only. The gap between real and fake sources appears well marked when the definition of galaxy-rich large-scale environments depends on the number of cosmological neighbors present in each MST cluster, while it only rises at  $z_{\text{src}} > 0.1$  when adopting the threshold for the number of candidate elliptical galaxies present in 5% of the MOCK sample (left panel). Lower panels show the same ratios but taking into account the noise subtraction as described in Section 4.



**Figure 24.** All the SDSS galaxies in the 2 Mpc circular region centered on the position of SDSS J073505.25+415827.5 at  $z = 0.087$  are shown in gray, while regions with galaxy density larger than 20% are marked with green contours, computed by applying the KDE. The sources included within the green contours are marked as black filled circles. Blue crosses mark the location of candidate elliptical galaxies. Following Figures 19, 21, and 23, source ratios are calculated with the KDE algorithm only. The gap between real and fake sources is again significant at almost all redshifts. Fluctuations in the samples of radio galaxies are mainly due to their limited number of sources. The right panel, as in previous figures, shows the same ratios as the central panel but with noise subtracted.



**Figure 25.** Comparison between the results obtained on the whole radio galaxy sample (i.e., FRICAT + FRIICAT) with different clustering algorithms (see Appendix for a full description of these methods). Results achieved with the cosmological overdensities appear to be the best procedure, while Voronoi tessellation seems to be the least efficient. DBSCAN with the OPTICS implementation as well as MST seem to produce the same results. Independently of the algorithm adopted, radio galaxies live in galaxy-rich large-scale environments. Uncertainties are not reported in this plot since the main goal is to compare the efficiency of the procedures.

Figure 24 as a function of their redshift, where the gap between real and fake objects is again clear, particularly above  $z_{\text{src}} = 0.1$ , as occurs for the previous clustering algorithms.

#### Comparison between Cosmological Overdensities and Clustering Algorithms

The comparison between all three algorithms and their results obtained with the cosmological overdensities are shown in Figure 25. Here we report the total fraction of radio galaxies (FRICAT and FRIICAT together) living in galaxy-rich large-scale environments as a function of their redshift. All these methods are unaffected by a strong  $z_{\text{src}}$  dependence that could bias the results at low redshifts, as occurs for the cluster cross-matches (see Figure 10). Considering that their threshold is set at 5%, they are all in agreement with the statement that radio galaxies tend to live in a denser environment than occurs randomly. MST and DBSCAN appear to have the same efficiency, with the latter working better at higher redshifts. Voronoi tessellation seems to be systematically weaker than all the others that find source clusters, under the assumption that all sources lie in galaxy groups/clusters.

Finally, we note that one of the major drawback of all clustering algorithms adopted, including the FoF method, which was used to create the T12 cluster catalog, is that they can find regions with higher galaxy density than that of background/foreground sources in the field but in most of these cases such structures include objects at similar redshifts, which could be physically connected. The use of the cosmological overdensity procedure, as well as the use of the clustering algorithms combined with the number of cosmological neighbors lying in detected clusters, mitigates this bias.

#### ORCID iDs

F. Massaro <https://orcid.org/0000-0002-1704-9850>  
A. Capetti <https://orcid.org/0000-0003-3684-4275>

R. D. Baldi <https://orcid.org/0000-0002-1824-0411>  
I. Pillitteri <https://orcid.org/0000-0003-4948-6550>  
R. Campana <https://orcid.org/0000-0002-4794-5453>  
A. Paggi <https://orcid.org/0000-0002-5646-2410>

#### References

- Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2012, *ApJS*, 203, 21  
Allen, S. W., Dunn, R. J. H., Fabian, A. C., Taylor, G. B., & Reynolds, C. S. 2006, *MNRAS*, 372, 21  
Ankerst, M., Breunig, M. M., Kriegel, H., & Sander, J. 1999, in Proc. Int. Conf. on Management of Data, SIGMOD' 99 (New York: ACM), 49  
Baldi, R. D., & Capetti, A. 2008, *A&A*, 489, 989  
Baldi, R. D., & Capetti, A. 2010, *A&A*, 519A, 48  
Baldi, R. D., Capetti, A., & Giovannini, G. 2015, *A&A*, 576A, 38  
Baldi, R. D., Capetti, A., & Massaro, F. 2018, *A&A*, 609A, 1  
Balmaverde, B., Baldi, R. D., & Capetti, A. 2008, *A&A*, 486, 119  
Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1985, *MNRAS*, 216, 17  
Belsole, E., Worrall, D. M., Hardcastle, M. J., & Croston, J. H. 2007, *MNRAS*, 381, 1109  
Bender, R., Doebereiner, S., & Moellenhoff, C. 1988, *A&A*, 74, 385  
Bender, R., Surma, P., Doebereiner, S., Moellenhoff, C., & Madejsky, R. 1989, *A&A*, 217, 35  
Bennett, A. S. 1962, *MmRAS*, 68, 163  
Bennett, C. L., Larson, D., Weiland, J. L., & Hinshaw, G. 2014, *ApJ*, 794, 135  
Berlind, A. A., et al. 2006, *ApJS*, 167, 1  
Best, P. N. 2004, *MNRAS*, 351, 70  
Best, P. N. 2007, *NewAR*, 51, 168  
Best, P. N., & Heckman, T. M. 2012, *MNRAS*, 421, 1569  
Best, P. N., Kauffmann, G., Heckman, T. M., et al. 2005, *MNRAS*, 362, 25  
Best, P. N., von der Linden, A., Kauffmann, G., Heckman, T. M., & Kaiser, C. R. 2007, *MNRAS*, 379, 894  
Bicknell, G. V. 1994, *ApJ*, 422, 542  
Biviano, A. 2000, in Proc. IAP Meeting, Constructing the Universe with Clusters of Galaxies, ed. F. Durret & D. Gerbal (Paris: IAP)  
Bornancini, C. G., De Breuck, C., de Vries, W., et al. 2007, *MNRAS*, 378, 551  
Buttiglione, S., Capetti, A., Celotti, A., et al. 2010, *A&A*, 509, 6  
Cameron, E. 2011, *PASP*, 28, 128  
Campana, R., Bernieri, E., Massaro, E., Tinebra, F., & Tosti, G. 2013, *ApSS*, 347, 169  
Campana, R., Massaro, E., Gasparrini, D., Cutini, S., & Tramacere, A. 2008, *MNRAS*, 383, 1166  
Capetti, A., Massaro, F., & Baldi, R. D. 2017a, *A&A*, 598A, 49  
Capetti, A., Massaro, F., & Baldi, R. D. 2017b, *A&A*, 601A, 81  
Carilli, C. L., & Barthel, P. D. 1996, *A&ARv*, 7, 1  
Celotti, A., & Fabian, A. C. 2004, *MNRAS*, 353, 523  
Chiaberge, M., Capetti, A., & Celotti, A. 2002, *A&A*, 394, 791  
Ching, J. H. Y., Croom, S. M., Sadler, E. M., et al. 2017, *MNRAS*, 469, 4584  
Condon, J. J., Cotton, W. D., Greisen, E. W., et al. 1998, *AJ*, 115, 1693  
D'Abrusco, R., Longo, G., & Walton, N. A. 2009, *MNRAS*, 396, 223  
D'Abrusco, R., Massaro, F., Paggi, A., et al. 2014, *ApJS*, 215, 14  
De Breuck, C., Klamer, I., Johnston, H., et al. 2006, *MNRAS*, 366, 58  
De Breuck, C., van Breugel, W., Röttgering, H., et al. 2001, *AJ*, 121, 1241  
Edge, D. O., Shakeshaft, J. R., McAdam, W. B., Baldwin, J. E., & Archer, S. 1959, *MmRAS*, 68, 37  
Eisenstein, D. J., Weinberg, D. H., Agol, E., et al. 2011, *ApJ*, 142, 72  
Eke, V. R. 2004, *MNRAS*, 348, 866  
Erlund, M. C., Fabian, A. C., Blundell, K. M., Moss, C., & Ballantyne, D. R. 2007, *MNRAS*, 379, 498  
Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. 1996, in Proc. of II Int. Conf. on Knowledge Discovery and Data Mining, ed. E. Simoudis, J. Han, & U. Fayyad (Menlo Park, CA: AAAI Press), 226  
Faber, S. M., Tremaine, S., Ajhar, E. A., et al. 1997, *AJ*, 114, 1771  
Fabian, A. C. 2012, *ARA&A*, 50, 455  
Fanaroff, B. L., & Riley, J. M. 1974, *MNRAS*, 167, 31  
Gendre, M. A., Best, P. N., Wall, J. V., & Ker, L. M. 2013, *MNRAS*, 430, 3086  
Gladders, M. D., Lopez-Cruz, O., Yee, H. K. C., & Kodama, T. 1998, *ApJ*, 501, 571  
Gladders, M. D., & Yee, H. K. C. 2000, *ApJ*, 120, 2148  
Haas, M. R., Schaye, J., & Jeesson-Daniel, A. 2012, *MNRAS*, 419, 2133  
Hao, C. N., Mao, S., Deng, Z. G., Xia, X. Y., & Wu, H. 2006, *MNRAS*, 370, 1339  
Hao, J., McKay, T. A., Koester, B. P., et al. 2010, *ApJS*, 191, 254  
Hardcastle, M. J., Ching, J. H. Y., Virdee, J. S., et al. 2013, *MNRAS*, 429, 2407



- Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2006, *MNRAS*, **370**, 1893
- Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2007, *MNRAS*, **376**, 1849
- Hardcastle, M. J., Evans, D. A., & Croston, J. H. 2009, *MNRAS*, **396**, 1929
- Hardcastle, M. J., & Worrall, D. M. 2000, *MNRAS*, **319**, 562
- Heckman, T. M., & Best, P. N. 2014, *ARA&A*, **52**, 589
- Hill, G. J., & Lilly, S. J. 1991, *ApJ*, **367**, 1
- Hine, R. G., & Longair, M. S. 1979, *MNRAS*, **188**, 111
- Huchra, J. P., & Geller, M. J. 1982, *ApJ*, **257**, 423
- Ineson, J., Croston, J. H., Hardcastle, M. J., et al. 2013, *ApJ*, **770**, 136
- Ineson, J., Croston, J. H., Hardcastle, M. J., et al. 2015, *MNRAS*, **453**, 2682
- Jackson, N., & Rawlings, S. 1997, *MNRAS*, **286**, 241
- Jarvis, M. J., Rawlings, S., Eales, S., et al. 2001, *MNRAS*, **326**, 1585
- Laing, R. A., & Bridle, A. H. 2008, *ASPC*, **386**, 70
- Laing, R. A., Jenkins, C. R., Wall, J. V., & Unger, S. W. 1994, in ASP Conf. Ser. 54, *The First Stromlo Symp.: The Physics of Active Galaxies*, ed. G. V. Bicknell, M. A. Dopita, & P. J. Quinn (San Francisco, CA: ASP), 201
- Lee, D. T., & Schacter, B. J. 1980, *Int. J. Comp. Info. Sci.*, **9**, 219
- Lintott, C. J., Schawinski, K., Slosar, A., et al. 2008, *MNRAS*, **389**, 1179
- Longair, M. S. 1971, *RPPh*, **34**, 1125
- Massaro, F., D'Abrusco, R., Ajello, M., et al. 2011, *ApJ*, **740L**, 48
- Massaro, F., D'Abrusco, R., Paggi, A., et al. 2013a, *ApJS*, **209**, 10
- Massaro, F., Giroletti, M., D'Abrusco, R., et al. 2014, *ApJS*, **213**, 3
- Massaro, F., Harris, D. E., Tremblay, G. R., et al. 2013b, *ApJS*, **206**, 7
- Massaro, F., Missaglia, V., Stuardi, C., et al. 2018, *ApJS*, **234**, 7
- McNamara, B. R., & Nulsen, P. E. J. 2007, *ARA&A*, **45**, 117
- Miley, G. K., & De Breuck, C. 2008, *A&ARv*, **15**, 67
- Mingo, B., Hardcastle, M. J., Croston, J. H., et al. 2014, *MNRAS*, **440**, 269
- Miraghaei, H., & Best, P. N. 2017, *MNRAS*, **466**, 4346
- Mocz, P., Fabian, A. C., & Blundell, K. M. 2011, *MNRAS*, **413**, 1107
- Moore, B., Frenk, C. S., & White, S. D. M. 1993, *MNRAS*, **261**, 827
- Prestage, R. M., & Peacock, J. A. 1988, *MNRAS*, **230**, 131
- Quillen, A. C., Bower, G. A., & Stritzinger, M. 2000, *ApJS*, **128**, 85
- Ramella, M., Boschini, W., Fadda, D., & Nonino, M. 2001, *A&A*, **368**, 776
- Richards, G. T., Nichol, R. C., Gray, A. G., et al. 2004, *ApJS*, **155**, 257
- Rines, K., Geller, M. J., Diaferio, A., & Kurtz, M. J. 2013, *ApJ*, **767**, 15
- Röttgering, H. J. A., van Ojik, R., Miley, G. K., et al. 1997, *A&A*, **326**, 505
- Sabater, J., Best, P. N., & Argudo-Fernández, M. 2013, *MNRAS*, **430**, 638
- Scharf, C., Smail, I., Ivison, R., et al. 2003, *ApJ*, **596**, 105
- Smail, I., Blundell, K. M., Lehmer, B. D., & Alexander, D. M. 2012, *ApJ*, **760**, 132
- Smolčić, V., Zamorani, G., Schinnerer, E., et al. 2009, *ApJ*, **696**, 24
- Spinrad, H., Marr, J., Aguilar, L., & Djorgovski, S. 1985, *PASP*, **97**, 932
- Stuardi, C., Missaglia, V., Massaro, F., et al. 2018, *ApJS*, **235**, 32
- Tago, E., Saar, E., Tempel, E., et al. 2010, *A&A*, **514A**, 102
- Tasse, C., Best, P. N., Röttgering, H., & Le Borgne, D. 2008, *A&A*, **490**, 893
- Taylor, M. B. 2005, in ASP Conf. Ser. 347, *adass XIV*, ed. P. Shopbell, M. Britton, & R. Ebert (San Francisco, CA: ASP), 29
- Tempel, E., Tago, E., & Liivamägi, L. J. 2012, *A&A*, **540A**, 106
- Tramacere, A., & Vecchio, C. 2013, *A&A*, **549A**, 138
- Tremblay, G. R., Chiaberge, M., Donzelli, C. J., et al. 2007, *ApJ*, **666**, 109
- Visvanathan, N., & Sandage, A. 1977, *ApJ*, **216**, 214
- White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, *ApJ*, **475**, 479
- Wing, J. D., & Blanton, E. L. 2011, *AJ*, **141**, 88
- Worpel, H., Brown, M. J. I., Jones, D. H., Floyd, D. J. E., & Beutler, F. 2013, *ApJ*, **772**, 64
- Zirbel, E. L. 1997, *ApJ*, **476**, 489