

Large-scale environments of narrow-line Seyfert 1 galaxies

E. Järvelä^{1,2}, A. Lähteenmäki^{1,2,3}, H. Lietzen³, A. Poudel⁴, P. Heinämäki⁴, and M. Einasto³

¹ Aalto University Metsähovi Radio Observatory, Metsähovintie 114, 02540 Kylmälä, Finland
e-mail: emilia.jarvela@aalto.fi

² Aalto University Department of Electronics and Nanoengineering, PO Box 15500, 00076 Aalto, Finland

³ Tartu Observatory, Observatooriumi 1, 61602 Tõravere, Estonia

⁴ Tuorla Observatory, University of Turku, Väisäläntie 20, 21500 Piikkiö, Finland

Received 6 June 2017 / Accepted 3 July 2017

ABSTRACT

Studying large-scale environments of narrow-line Seyfert 1 (NLS1) galaxies gives a new perspective on their properties, particularly their radio loudness. The large-scale environment is believed to have an impact on the evolution and intrinsic properties of galaxies, however, NLS1 sources have not been studied in this context before. We have a large and diverse sample of 1341 NLS1 galaxies and three separate environment data sets constructed using Sloan Digital Sky Survey. We use various statistical methods to investigate how the properties of NLS1 galaxies are connected to the large-scale environment, and compare the large-scale environments of NLS1 galaxies with other active galactic nuclei (AGN) classes, for example, other jetted AGN and broad-line Seyfert 1 (BLS1) galaxies, to study how they are related. NLS1 galaxies reside in less dense environments than any of the comparison samples, thus confirming their young age. The average large-scale environment density and environmental distribution of NLS1 sources is clearly different compared to BLS1 galaxies, thus it is improbable that they could be the parent population of NLS1 galaxies and unified by orientation. Within the NLS1 class there is a trend of increasing radio loudness with increasing large-scale environment density, indicating that the large-scale environment affects their intrinsic properties. Our results suggest that the NLS1 class of sources is not homogeneous, and furthermore, that a considerable fraction of them are misclassified. We further support a published proposal to replace the traditional classification to radio-loud, and radio-quiet or radio-silent sources with a division into jetted and non-jetted sources.

Key words. galaxies: active – galaxies: Seyfert – large-scale structure of Universe

1. Introduction

Active galactic nuclei (AGN) are the most energetic non-transient phenomena in our Universe. They radiate over the whole electromagnetic spectrum, from radio to TeV energies, and are powered by a supermassive black hole accreting matter. AGN can be divided into various classes based on their observed properties. The most numerous classes are the luminous and usually more distant quasars, bright enough to easily overshadow their host galaxy, and fainter Seyfert galaxies with often detectable host galaxies. Seyfert galaxies were originally divided into two subclasses based on their optical spectra (Khachikian & Weedman 1974); Seyfert 1 galaxies (Sy1) show broad permitted emission lines arising from the broad-line region (BLR) close to the nucleus and narrow forbidden emission lines originating from the narrow-line region (NLR) further away. In Seyfert 2 galaxies (Sy2) the broad emission lines are not directly detectable and they show only the narrow lines in their spectra. It was early on suggested that Sy1 and Sy2 galaxies are similar sources seen at different angles and through varying amounts of obscuring matter (Osterbrock 1978). Soon after, several intermediate types, for example, Sy 1.2 and 1.8 – corresponding to different viewing angles – were introduced supporting this scenario (Osterbrock 1981).

A subclass of Sy1 galaxies with narrow permitted lines arising from the BLR was described by Osterbrock & Pogge (1985). These narrow-line Seyfert 1 (NLS1) galaxies are defined by their optical spectra; both, permitted and forbidden emission lines are narrow ($FWHM(H\beta) < 2000 \text{ km s}^{-1}$, Goodrich 1989) and [O III] is relatively weak ($F([O \text{ III}])/F(H\beta) < 3$, Osterbrock & Pogge 1985). In addition relative Fe II emission

is strong in the majority of NLS1 sources (Osterbrock & Pogge 1985).

The black hole masses in NLS1 galaxies are low or intermediate ($M_{\text{BH}} < 10^8 M_{\odot}$, Peterson et al. 2000). They accrete at extraordinarily high rates, from 0.1 Eddington ratio to super-Eddington accretion (Boroson & Green 1992), and lie below the normal relations of the black hole mass and the properties of the galactic bulge – σ_* (stellar velocity dispersion) and L_{bulge} (Mathur et al. 2001, but some studies also disagree, for example, Woo et al. 2015). It has been suggested that NLS1 galaxies are young objects in the early stages of their evolution (Mathur et al. 2001).

The majority of NLS1 sources are radio-quiet; only ~7% are radio-loud and 2–3% percent very radio-loud (Komossa et al. 2006). NLS1 galaxies typically show a very compact unresolved radio core, however, lately, evidence of pc- and kpc-scale structures has been found (Gliozzi et al. 2010; Doi et al. 2012; Richards & Lister 2015; Gu et al. 2015; Doi et al. 2013, 2015; Richards et al. 2015; Lister et al. 2016). Subluminal and superluminal speeds measured in some sources suggest Lorentz factors and viewing angles similar to BL Lac objects (BLO) and flat-spectrum radio quasars (FSRQ; Lister et al. 2016). The Fermi Gamma-ray Space Telescope has detected gamma-ray emission in some NLS1 galaxies (e.g. Abdo et al. 2009), thus confirming the presence of powerful relativistic jets. NLS1 galaxies detected in gamma-rays or showing extended radio emission seem to have, on average, more massive black holes ($M_{\text{BH}} > 10^7 M_{\odot}$, e.g. Doi et al. 2012; Foschini et al. 2015) than the NLS1 population in general (e.g. this paper and Järvelä et al. 2015). However, these samples are not complete; more studies will be needed to clarify this possible connection.

So far, host galaxy studies of NLS1 galaxies have mostly concentrated on the radio-quiet sources. They have been found to preferably, but not exclusively, be hosted by disk-like galaxies. Only a few studies of the host galaxies of radio-loud NLS1 sources exist; gamma-ray emitting 1H0323+342 resides in a one-armed spiral galaxy (Zhou et al. 2007) or possibly in a system disturbed by merging (Antón et al. 2008; León Tavares et al. 2014), and the host galaxy of another gamma-ray emitting NLS1, FBQS J1644+2619, seems to be a barred lenticular galaxy (Olguín-Iglesias et al. 2017). Mathur (2000) suggested that all NLS1 galaxies could be sources rejuvenated by a recent merger. Later studies (e.g. Ohta et al. 2007; Ryan et al. 2007) found no evidence for abundant interaction or merging in NLS1 galaxies, thus it seems improbable that all NLS1 galaxies would be a result of a merger or interaction. However, they might play a role; mergers and interaction are known to be able to trigger the nuclear activity (e.g. Barth et al. 2008; Ellison et al. 2011).

Studying NLS1 galaxies is challenging because of the seemingly heterogeneous nature of the population. One additional way of addressing the issue is to study the environments of NLS1 galaxies compared to other AGN classes and within the NLS1 population. The environment – at all scales – affects the intrinsic properties and evolution of galaxies, and consequently their nuclear activity. We can divide the environment into several scales: 1) the host galaxy, which is the closest environment of an AGN; 2) the local environment, including the neighbouring galaxies and the group and cluster the galaxy belongs to; and 3) the large-scale environment, tracing the largest, supercluster-scale, structures in the Universe.

The connection between the AGN and its host galaxy has been extensively studied (e.g. Storchi-Bergmann 2008; van de Ven & Fathi 2010; Pović et al. 2012; Fabian 2012; King & Pounds 2015). The host galaxy regulates the gas supply of the black hole and thus has a direct impact on its activity level. The AGN in turn affects the host galaxy via various feedback mechanisms, for example, the radiation pressure, the jet, and winds and outflows. The AGN can induce both negative and positive feedback, and thus steer the evolution of the galaxy by, for example, quenching or enhancing star formation.

The local environment can have an impact on galaxy dynamics and evolution. In a galaxy – cluster interaction the galaxy might lose a fraction or almost all of its gas due to ram-pressure stripping, thus quenching star formation and advancing the evolution (Ebeling et al. 2014; Steinhauser et al. 2016). In a dense enough environment, a galaxy might undergo a number of minor and major mergers that distort its morphology. Mergers may cause gas infall, triggering circumnuclear star formation and feeding the black hole, and replenish or strip gas reservoirs. The connection between the mergers and triggering an AGN is not clear; some studies suggest that both minor (Taniguchi 1999; Barth et al. 2008) and major (Urrutia et al. 2008; Ellison et al. 2011) mergers are able to trigger the nuclear activity, whereas others do not find them connected (Corbin 2000; Cisternas et al. 2011; Kocevski et al. 2012). It has been argued that mergers, especially major ones, predominantly trigger the most luminous AGN (Treister et al. 2012; Villforth et al. 2014).

The connection between cluster-scale environment and galaxy morphology was discovered early on (Hubble & Humason 1931; Dressler 1980) and has been established by subsequent studies (e.g. Park & Choi 2009; Chen et al. 2017). This relation is thought to be a consequence of the more frequent galaxy – galaxy interactions in denser regions. Mergers and interaction transform the galaxy morphology towards the early

type, and the speed of this evolution depends on the environment density.

The position of the galaxy in the cosmic web of superclusters, filaments and voids affects its properties as well. The effect is similar to the one found for cluster-scale environments; the galaxies residing in denser large-scale environments – close to or in filaments and superclusters – are preferably ellipticals, whereas the fraction of spirals increases the further one departs from these higher-density regions (Lietzen et al. 2012; Einasto et al. 2014; Chen et al. 2017; Kuutma et al. 2017; Pandey & Sarkar 2017). Moreover, the properties of the galaxy groups are affected by the large-scale environment (Poudel et al. 2017).

In this paper we study the large-scale environments of NLS1 galaxies using diverse and large, statistically significant samples. We use three environment data sets at three different redshift ranges between 0 and 0.62 to compare the large-scale environments of NLS1 sources with the large-scale environments of other types of AGN, as well as study the differences in the environments within the NLS1 population. In addition, we are interested in seeing whether the large-scale environment is connected to any of the intrinsic properties of NLS1 galaxies.

This paper is organised as follows. In Sect. 2 we introduce the sample and the data used for the analyses, and calculate some additional parameters. We present the results in Sect. 3, discuss the main findings in Sect. 4, and finally conclude in Sect. 4.4. Throughout the paper we assume a cosmology with $H_0 = 73 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_{\text{matter}} = 0.27$ and $\Omega_{\text{vacuum}} = 0.73$ (Spergel et al. 2007).

2. Data

2.1. Sample

Our seed sample included all 2011 NLS1 galaxies from Zhou et al. (2006) and 38 additional sources chosen from Komossa et al. (2006), Whalen et al. (2006), Yuan et al. (2008) and Foschini (2011) based on radio loudness. These additional sources were included because they are part of the Metsähovi NLS1 monitoring programme (Lähteenmäki et al. 2017). NLS1 galaxies in Zhou et al. (2006) were selected from Sloan Digital Sky Survey¹ (SDSS, York et al. 2000) Data Release 3 and had $z \lesssim 0.8$ and the FWHM of the broad $H\beta$ or $H\alpha \lesssim 2200 \text{ km s}^{-1}$ at the 10σ or higher confidence level. Most of these sources have not been detected at radio frequencies. Yuan et al. (2008) presents a sample of 23 radio-loud NLS1 galaxies selected from SDSS Data Release 5 under the same restrictions as in Zhou et al. (2006). The sample in Komossa et al. (2006) included 11 radio-loud NLS1 galaxies found by cross-correlating the Catalogue of Quasars and Active Nuclei (Veron-Cetty & Veron 2003) with several radio and optical catalogs using the cross-matcher application developed within the German Astrophysical Virtual Observatory (GAVO)² project. The sample was limited by the requirement $H\beta < 2000 \text{ km s}^{-1}$. Foschini (2011) and Whalen et al. (2006) both present a sample of both radio-loud and radio-quiet NLS1 galaxies.

2.2. Multiwavelength and spectral data

We retrieved all the multiwavelength data available from ASI Science Data Center (ASDC³). Data obtained from ASDC were

¹ www.sdss.org

² www.g-vo.org

³ www.asdc.asi.it

already corrected for galactic extinction. The archival data are not simultaneous. Studying correlations between wavebands of variable sources, such as AGN, should ideally be performed with data that are no more than a couple of weeks apart or even less for more rapidly varying jetted sources. The time delays between the wavebands should also be taken into account, but this would require frequent multiwavelength monitoring of the sources and is, in practise, often impossible. Details of the wavebands used are given below. The data in all wavebands are not complete.

Radio data (radio flux density, F_R) are from the National Radio Astronomy Observatory (NRAO) Very Large Array (VLA⁴) Faint Images of the Radio Sky at Twenty-Centimeters (FIRST⁵) survey and NRAO VLA Sky Survey (NVSS⁶), both at 1.4 GHz. In the case of multiple detections we chose the one closest to the source coordinates. The search radius used for the radio surveys was 1 arcmin, so it is possible that in cases when the NLS1 source was not detected in radio, the radio data are actually that of a nearby radio source, but the number of the false sources is very low. The detection limit of FIRST survey is 1 mJy.

Infrared data (infrared flux density, F_{IR}) are from Wide-field Infrared Survey Explorer (WISE⁷) AllWISE Data Release⁸. WISE bands are W1–W4 with wavelengths 3.4, 4.6, 12, and 22 μm , respectively.

Optical data (optical flux density, F_O) are from SDSS Data Release 10⁹ (DR10), or, if DR10 data were not available, from DR7¹⁰. SDSS has five bands, $u, g, r, i,$ and z with wavelengths 355.1, 468.6, 616.6, 748.0, and 893.2 nm, respectively. As a representative of the optical emission we used the g -band.

X-ray data (X-ray flux density, $F_{X\text{-ray}}$) are from ROSAT¹¹ all-sky survey bright source catalogue (RASS¹²) and the WGA Catalog of ROSAT Point Sources (WGACAT¹³) at soft X-rays (0.1–2.4 keV). For sources that had both RASS and WGACAT observations we chose the closest one to the source coordinates.

Spectral data were obtained from Zhou et al. (2006). They are from SDSS DR3 and include: 1) the full width at half maximum of the broad $H\beta$ emission line, $FWHM(H\beta)$, 2) the flux of the [OIII] $\lambda 5007$ emission line, $F([OIII])$, 3) optical Fe II strength relative to the broad component of $H\beta$, R4570, and 4) the monochromatic luminosity at 5100 Å, λL_{5100} .

The broadness of the $H\beta$ emission line can be used as a proxy for the black hole mass because the BLR clouds it arises from are gravitationally bound to the black hole. In the case of a disk-like BLR, the $FWHM(H\beta)$ can additionally give us information about the orientation of the source. The intensity of the forbidden [O III] line arising from the NLR correlates with the underlying photoionizing continuum and can be used to estimate the intrinsic AGN power (e.g. Simpson 1998). The relative strength of the Fe II to the broad $H\beta$ is known to anticorrelate with the [O III] strength. This is hypothesized to be driven by the changing Eddington ratio. Increasing Eddington ratio could lead to changes in the structure of the accretion flow and thus affect the ionizing continuum seen by the NLR (e.g. Shen & Ho 2014). Furthermore, R4570 is possibly related to the radio emission;

⁴ <http://www.vla.nrao.edu/>

⁵ www.sundog.stsci.edu

⁶ <http://www.cv.nrao.edu/nvss/>

⁷ www.nasa.gov/wise

⁸ <http://wise2.ipac.caltech.edu/docs/release/allwise/>

⁹ <https://www.sdss3.org/dr10/>

¹⁰ <http://www.sdss.org/dr7/>

¹¹ <http://heasarc.gsfc.nasa.gov/docs/rosat/roskof.html>

¹² <http://heasarc.gsfc.nasa.gov/docs/rosat/rass.html>

¹³ <http://heasarc.gsfc.nasa.gov/wgacat/>

Yuan et al. (2008) found that the Fe II emission is on average stronger in radio-loud NLS1 galaxies than in the NLS1 population in general. λL_{5100} is used as an estimate of the continuum level. If a relativistic jet is present, it contributes to λL_{5100} . However, it is not possible to estimate the level of the jet contamination with the available data.

2.2.1. Radio-loudness parameter

We computed the radio-loudness (R) parameter for all of our sources with radio and optical data, which in total is 237 sources. We used the commonly defined R value; the ratio of radio flux density (F_R) and optical flux density (F_O). For F_R we used the radio flux density from the FIRST/NVSS survey at 1.4 GHz and for F_O we used the SDSS g -band (468.6 nm) optical flux density because it is closest to the commonly used wavelength of 440 nm. The flux density values are not simultaneous. We did not apply K-correction since it is inaccurate in this case. First, it depends on the spectral shape of the source. The spectral shape of NLS1 galaxies, especially in the radio band, varies from steep to inverted/convex in different sources, and for individual sources at different epochs (e.g. Lähteenmäki et al. 2017), thus the correct K-correction is source- and time-dependent. Second, NLS1 galaxies are variable in both radio and optical bands, and this variability affects the R value much more than the K-correction does; the effect of the K-correction on the radio-loudness at our $z_{\text{max}} = 0.62$ is <30% (K-correction computed as in Foschini 2011), while the variance in R induced by the variable flux density can be as high as 750% (e.g. Gabányi et al. 2017). The boundary between the radio-loud and radio-quiet sources is arbitrary, and due to variability, the sources may move about the boundary, from radio-loud to radio-quiet, and vice versa.

Even though the parameter is not necessarily the best or even an entirely correct way of estimating the radio characteristics of AGN, we use it for dividing our samples into subsamples and for comparing our results with earlier studies. According to the traditional definition 113 of our sources are radio-loud ($10 < R < 100$; RL) and 27 of them are very radio-loud ($R > 100$; VRL), leaving 97 radio-quiet ($R < 10$; RQ) sources. Altogether, 237 of 1341 sources are radio-detected (RD). In this sample we have 1104 sources that have never been detected at radio frequencies and are therefore seemingly radio-silent (RS). However, recent radio observations performed at Metsähovi Radio Observatory at 22 and 37 GHz indicate that at least some RS NLS1 galaxies are detectable (Lähteenmäki et al., in prep.). Some RS NLS1 sources are clearly misclassified, but without extensive radio surveys it is impossible to say which fraction. It should be noted that the obscure boundary between RL and RQ sources, their variability, and the misclassification of some RS sources, affect the data analysis performed using these subsamples.

2.2.2. Black hole masses

We estimated black hole masses (M_{BH}) using the $FWHM(H\beta)$ – luminosity mass scaling relation (Greene & Ho 2005)

$$M_{\text{BH}} = (4.4 \pm 0.2) \times 10^6 \left(\frac{\lambda L_{5100}}{10^{44} \text{ erg s}^{-1}} \right)^{0.64 \pm 0.02} \left(\frac{FWHM(H\beta)}{10^3 \text{ km s}^{-1}} \right)^2 M_{\odot}. \quad (1)$$

The λL_{5100} and $FWHM(H\beta)$ values were taken from Zhou et al. (2006).

We used this method instead of methods based on $M_{\text{BH}} - \sigma_*$ (stellar velocity dispersion of the bulge) or the $M_{\text{BH}} - M_{\text{bulge}}$

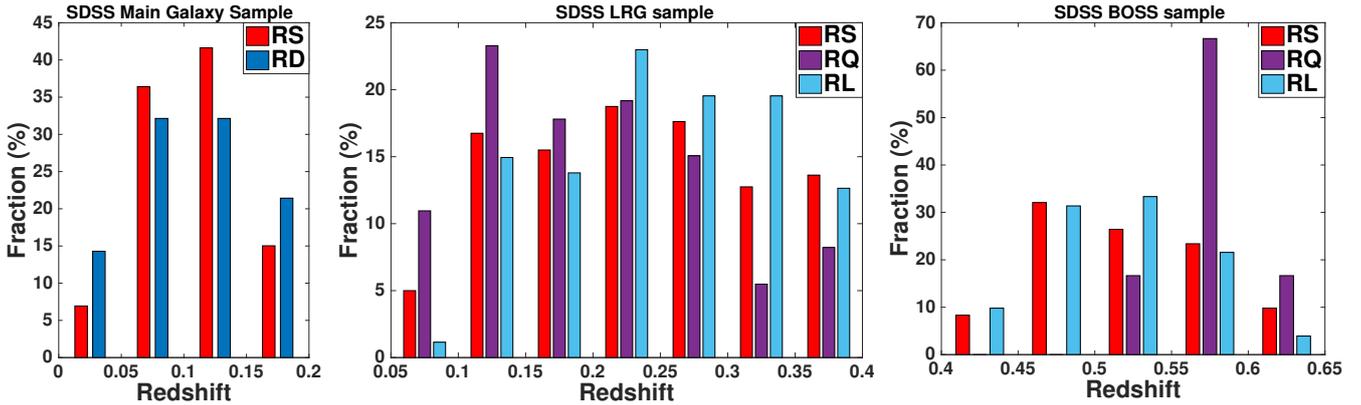


Fig. 1. Redshift distributions of the three samples. *Left:* SDSS Main Galaxy sample, 229 sources. *Middle:* SDSS LRG sample, 960 sources, of which 171 overlap with the SDSS Main Galaxy sample. *Right:* SDSS BOSS sample, 323 sources.

(the mass of the bulge) relation (Bentz et al. 2009) because σ_* or L_{bulge} data are not available for most NLS1 sources. They also tend to lie below the normal $M_{\text{BH}} - \sigma_*$ and $M_{\text{BH}} - L_{\text{bulge}}$ relations (Mathur et al. 2001; Laor 2001; Zhou et al. 2006). However, this method does not take into account possible inclination effects caused by the geometry of the BLR and the viewing angle (Decarli et al. 2011), and for sources with relativistic jets this method overestimates the black hole mass (Wu et al. 2004) because the jet contamination to λL_{5100} cannot be estimated due to lack of data. The method is not precise but can be used as an order of magnitude estimate in statistical studies. We were able to estimate M_{BH} for 1312 sources.

2.3. Environment data samples

We use three sets of environment data in our analyses.

SDSS Main Galaxy sample and luminosity-density field (LDF). In this study we use the catalogue of NLS1 galaxies by Zhou et al. (2006), constructed from the spectroscopic sample of the SDSS Data Release 3 and consisting of 2011 NLS1 sources with redshifts $z \leq 0.8$. We cross-matched these galaxies with the galaxy and group catalogues by Tempel et al. (2014b) and filament catalogues by (Tempel et al. 2014a) and found 229 matches with redshifts $z \leq 0.2$. There remain ~ 300 NLS1 sources with $z < 0.2$ that could not be matched. The galaxy catalogue by Tempel et al. (2014b) is based on the SDSS Data Release 10 (Ahn et al. 2014) and uses only the main contiguous area of the SDSS Legacy Survey. It consists of 588193 galaxies and 82458 groups spanning an area of 7221 square degrees in the sky with redshifts $z \leq 0.2$. It also provides luminosity-density estimates with 1, 2, 4 and 8 h^{-1} Mpc smoothing for each galaxy and can be used to study the effect of environment on NLS1 galaxies. The redshift range of this sample is $z = 0.0133\text{--}0.1987$ and the mean $z = 0.1078$.

SDSS LRG LDF is a three-dimensional low-resolution LDF constructed using a sample of luminous red galaxies (LRG) in SDSS Data Release 7 (Abazajian et al. 2009). The mean density around each source is calculated in a volume of 3 h^{-1} Mpc, giving an estimate of its surroundings on a supercluster-scale. 16 h^{-1} Mpc smoothing was used. This field is limited to 225–1000 h^{-1} Mpc. For a detailed description, see Appendix A, Lietzen et al. (2011), and Liivamägi et al. (2012). We have SDSS LRG LDF data for 960 sources. The mean redshift for our sample is $z = 0.2340$, minimum $z = 0.0726$, and maximum

$z = 0.3996$. This sample and the SDSS Main Galaxy sample have 171 overlapping NLS1 sources.

SDSS BOSS LDF is constructed using the SDSS Baryon Oscillation Spectroscopic Survey (BOSS) Constant MASS (CMASS) sample from Data Release 12 (Alam et al. 2015), limited to 1200–1600 h^{-1} Mpc ($z = 0.43\text{--}0.62$). The field was calculated in a 3 h^{-1} Mpc grid, with a 8 h^{-1} Mpc smoothing. For a detailed description see Appendix A and Lietzen et al. (2016). We have these data for 323 sources. The mean, minimum, and maximum redshifts for our sample are $z = 0.5224$, $z = 0.4412$ and $z = 0.6193$, respectively.

The group and filament data are available only for the Main Galaxy sample. The LDF, or for short, density, data is available for all samples but are not straightforwardly comparable with each other because of different smoothing scales and different selection criteria for the galaxies. The LDF is based on mean densities, not the actual physical large-scale structures. Because of this, the regions with intermediate densities can be either real physical filaments or transition zones between superclusters and voids. From now on the whole sample refers to all sources for which we have any environmental data, that is a total of 1341 sources.

The redshift distributions of the samples are shown in Fig. 1 and the average redshifts are given in Table 1. RL sources lie on average at higher redshifts than RQ sources. This might be due to either an observational bias or evolution of AGN. It is unlikely that this could be caused by an optical observational bias since DR3, based on which most of our sources were selected, is 95% complete down to $m_g = 22.2^{14}$, and most NLS1 galaxies are much brighter than this. Most of our radio data are from the FIRST survey which is complete down to 1 mJy, so we might miss a population of faint non-detectable radio emitters incorrectly classified as RS, but it is improbable that this could explain the whole difference. With future observatories we will be able to detect an abundant population of very radio-faint NLS1 galaxies. However, in general, these sources are expected to be optically faint and difficult to classify as NLS1 galaxies (Berton et al. 2015a), and they probably do not contaminate our sample, either. Thus the evolution of AGN is a more plausible explanation; AGN activity was more prominent in the earlier Universe (e.g. Wolf 2005; Beckmann & Shrader 2012). Therefore it is consistent that RL sources have higher redshifts.

The average M_{BH} values for the samples are given in Table 1 and the black hole mass distribution of the whole sample is

¹⁴ <http://classic.sdss.org/dr3/>

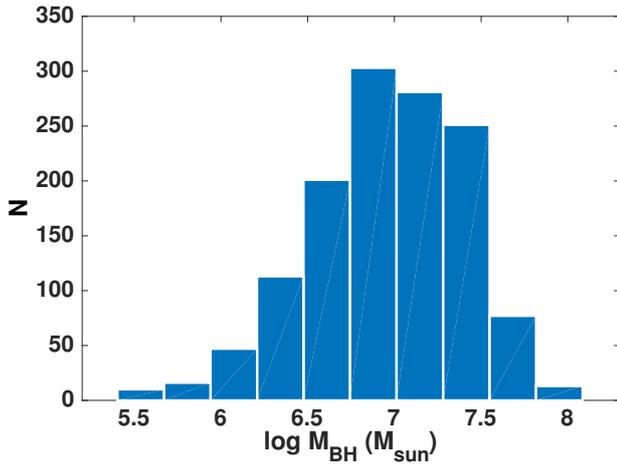
Table 1. Average redshifts and black hole masses of different subsamples.

Sample	SDSS MGS			SDSS LRG			SDSS BOSS		
	N	z	$\log M_{\text{BH}} (M_{\odot})$	N	z	$\log M_{\text{BH}} (M_{\odot})$	N	z	$\log M_{\text{BH}} (M_{\odot})$
All	229	0.108 ± 0.003	6.57 ± 0.03	960	0.234 ± 0.003	6.86 ± 0.01	323	0.522 ± 0.003	7.37 ± 0.02
RS	173	0.109 ± 0.003	6.55 ± 0.03	799	0.235 ± 0.003	6.85 ± 0.01	266	0.523 ± 0.003	7.35 ± 0.02
RD	56	0.103 ± 0.006	6.63 ± 0.05	161	0.227 ± 0.007	6.91 ± 0.03	57	0.522 ± 0.007	7.47 ± 0.04
RQ				73	0.201 ± 0.010	6.91 ± 0.05	6	0.577 ± 0.010	7.72 ± 0.09
RL				74	0.249 ± 0.009	6.91 ± 0.04	37	0.514 ± 0.009	7.46 ± 0.05
VRL				13	0.247 ± 0.026	6.96 ± 0.10	14	0.521 ± 0.013	7.36 ± 0.07
RL + VRL				87	0.249 ± 0.009	6.91 ± 0.04	51	0.515 ± 0.007	7.44 ± 0.04
XD	149	0.105 ± 0.004	6.58 ± 0.03	418	0.219 ± 0.004	6.87 ± 0.02	97	0.522 ± 0.006	7.39 ± 0.03

Table 2. SDSS Main Galaxy sample NLS1 statistics.

	N	Average density	N	N_{gal}^a	$R_{\text{vir}}^b (h^{-1} \text{ Mpc})$	$R_{\text{max}}^c (h^{-1} \text{ Mpc})$	N	$D_{\text{F}} (\text{Mpc})^d$
All	229	2.60 ± 0.13	51	10.96 ± 1.92	0.28 ± 0.02	0.50 ± 0.05	157	2.74 ± 0.20
RS	173	2.61 ± 0.15	41	10.44 ± 1.99	0.29 ± 0.02	0.53 ± 0.06	119	2.75 ± 0.23
RD	56	2.56 ± 0.27	10	12.91 ± 5.45	0.23 ± 0.05	0.41 ± 0.13	38	2.71 ± 0.40

Notes. ^(a) Average richness of a group. ^(b) Virial radius of the group, projected harmonic mean. ^(c) Maximum radius of the group, projected harmonic mean. ^(d) Distance to the closest filament axis.


Fig. 2. Black hole mass distribution of the whole sample of 1341 sources.

shown in Fig. 2. At the closest redshift range in the SDSS Main Galaxy Sample the RD sources harbour slightly more massive black holes than RS sources. In the SDSS LRG sample there is a small difference between the RS and RD sources, but no difference between RQ and RL sources. At the farthest redshift range, RS sources have the lowest black hole masses, while RL sources have masses only a little higher. RQ sources have the highest average black hole mass, but the sample size is small. The trend of increasing black hole mass with higher redshift is clear and is caused by the correlation of λL_{5100} and redshift, that is, NLS1 galaxies are brighter farther away.

3. Results

In this section we present the results of the various statistical studies we performed to investigate the impact of the large-scale environment. The large-scale environment analysis of NLS1 galaxies was done separately for the three large-scale environment density data samples and redshift ranges, and is presented in Sects. 3.1.1–3.1.3. We studied the changes in the large-scale

environments of subsamples of NLS1 galaxies, and compare them with other AGN classes. We performed principal component analysis (PCA) for the whole sample and selected subsamples. PCA was run separately for the SDSS LRG (Sect. 3.2.1) and BOSS (Sect. 3.2.2) samples. In Sect. 3.2.3 we study how the principal components are correlated with the intrinsic properties of NLS1 sources.

3.1. Large-scale environment

3.1.1. SDSS Main Galaxy sample

This sample with 229 sources is rather small so we divided it into only two subsamples, RS and RD, to obtain more reliable statistics. The averages of the luminosity-density with the $8 h^{-1}$ Mpc smoothing scale, the group richness, the group radii and the distance to the closest filament axis for the whole sample and the subsamples are shown in Table 2. The error is the standard error of the mean. The data are not complete; the number of sources used to calculate each average value is given in Table 2 in the “ N ” column. The average luminosity-densities for the subsamples, using any of the available smoothing scales, are similar, as is the average richness of the groups NLS1 galaxies reside in. Although only 52 out of 229 sources belong to a group ($N_{\text{gal}} \geq 3$), 22 are in pairs and 155 are field galaxies. The average black hole mass in isolated NLS1 galaxies is slightly higher (6.63 ± 0.03) than in NLS1 galaxies in pairs or in groups (6.45 ± 0.05). It should be noted that most of the NLS1 sources reside in poor groups, and a few individuals in unusually rich clusters dominate the average. Two of our sources reside in the same cluster, so average N_{gal} , R_{vir} and R_{max} are calculated using 51 groups. The average radii, R_{vir} and R_{max} , of the groups are typical for groups of this size (Bahcall 1996). RS and RD sources are similar in their properties except for the size of the groups they are in; RS NLS1 galaxies reside in groups with larger virial/maximum radii than RD sources. Since the average richness of the groups is similar, this means that the groups that RD sources reside in have a higher number density. In the SDSS filament catalogue we use (Tempel et al. 2014a), the filament radius is fixed

Table 3. Average density of the whole SDSS LRG sample and the NLS1 subsamples, and percentage in voids, intermediate density regions, and superclusters.

	N	Average density	$LD^a < 1$ (%)	$1 < LD < 3$ (%)	$LD > 3$ (%)
All	960	1.50 ± 0.04	44	44	12
RS	799	1.48 ± 0.05	45	43	12
RD	161	1.61 ± 0.10	40	47	13
RQ	73	1.48 ± 0.08	40	52	8
RL	74	1.66 ± 0.16	41	43	16
VRL	13	2.01 ± 0.36	38	38	23
RL and VRL	87	1.71 ± 0.14	40	43	17
Sy1 galaxies ^b	1095	1.73 ± 0.04	34	51	15
Sy2 galaxies ^b	2494	1.65 ± 0.03	35	52	13

Notes. ^(a) LD = mean luminosity-density of the LDF. ^(b) Lietzen et al. (2011).

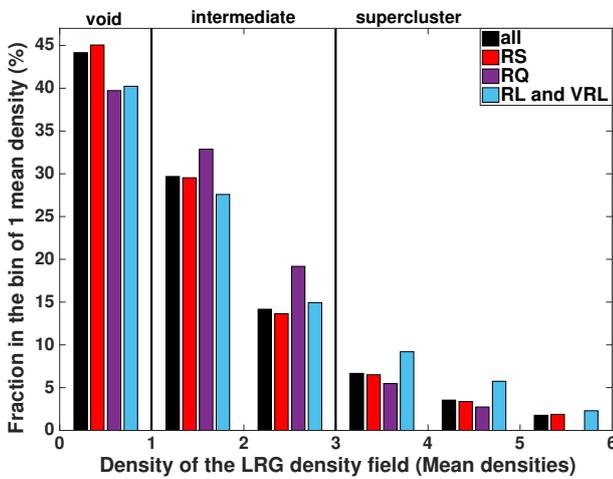


Fig. 3. Distributions of the whole sample and RQ, RQ, and RL plus VRL subsamples of NLS1 galaxies in the different environments defined with the SDSS LRG LDF. The last bin (5–6) includes all sources with density > 5 .

to $R_F = 0.5 h^{-1}$ Mpc. Out of 157 NLS1 galaxies with filament data, 40 have $D_F < 0.5 h^{-1}$ Mpc, thus most of the sources in our sample reside in a void.

3.1.2. SDSS LRG LDF

We have SDSS LRG LDF data for 960 sources, of which 171 are included in the SDSS Main Galaxy sample as well. We are interested in how the subsamples and NLS1 galaxies in general are distributed in the large-scale environment, and if the distribution is similar to the distributions of Seyfert galaxies and other types of AGN in Lietzen et al. (2011). Density distributions for the whole sample, RS, RQ, and RL plus VRL subsamples are shown in Fig. 3. Table 3 shows average densities for the subsamples and the fraction of sources in voids, intermediate density regions, and superclusters. The error is the standard error of the mean. Sy1 and Sy2 galaxies from Lietzen et al. (2011) are listed for comparison. We use the same definition for voids, intermediate density regions, and superclusters as in Lietzen et al. (2011); in a void the luminosity-density is less than 1.0, in an intermediate-density region the luminosity-density is between 1.0 and 3.0, and in a supercluster the luminosity-density is larger than 3.0. These limits were defined by spatially overlapping LRG LDF with the SDSS Main Sample LDF (Lietzen et al. 2011). As mentioned

before, the intermediate luminosity-density regions defined this way might correspond to either physical filaments or other areas of intermediate mean density; for example, the boundaries between superclusters and voids.

The average density for the whole sample is smaller than for any sample in Lietzen et al. (2011), and the distribution to voids, intermediate-density regions, and superclusters is clearly distinct compared to the samples in Lietzen et al. (2011). These results imply that the large-scale environment of NLS1 galaxies is different when compared to other AGN. The difference from Sy1 and Sy2 galaxies in the average densities is not so pronounced, but is significant nonetheless. This divergence is supported by the differences in their spatial distributions; there are more NLS1 galaxies than Sy1 and Sy2 galaxies in voids, and less in intermediate-density regions. Interestingly the fraction is almost the same in superclusters.

Amongst the NLS1 subsamples the average density increases with increasing radio loudness, and the distribution of sources to voids, intermediate-density regions, and superclusters changes with varying radio loudness. For RL and RQ sources the fraction in voids is the same (40%), but a larger percentage of RQ than RL sources reside in intermediate-density regions (52% vs. 43%), and vice versa for superclusters (8% vs. 16%). The fraction of the RS sources residing in voids (45%) and intermediate-density regions (43%) is very high. But surprisingly, a bigger fraction of RS than RQ sources reside in superclusters (12% vs. 8%). This might indicate, and be explained, by the misclassification of a fraction of RS sources, meaning that the RS subsample is mixed. The two-sample Kolmogorov-Smirnov (KS) test also suggests this; according to the KS-test the density distributions of our subsamples – RS vs. RD, RQ vs. RL and RS vs. RL – are drawn from the same distribution.

Lietzen et al. (2011) studied the large-scale environments of various types of AGN using the same SDSS LRG LDF that is used in this paper. Their sample included subsamples of radio-quiet (radio luminosity, $L_R < 10^{25}$ W Hz⁻¹) and radio-loud ($L_R > 10^{25}$ W Hz⁻¹) quasars, BL Lac objects, Sy1 and Sy2 galaxies, and radio galaxies divided into flat-spectrum, FR I, and FR II sources. Their Seyfert samples do not include NLS1 galaxies since they only chose sources with $[OIII]/H\beta > 3$, and for NLS1 sources $[OIII]/H\beta < 3$ by definition. Their main result was that radio galaxies tend to reside in denser large-scale environments than radio-quiet quasars, which can be explained by the AGN evolution scenario.

Overlap of Main Galaxy sample and LRG LDF. There are 171 overlapping sources in the SDSS Main Galaxy sample and LRG LDF. With this subset of sources we can study how the

Table 4. Average LRG luminosity-densities of the overlapping sources of the SDSS Main Galaxy sample and LRG, and subsamples defined by the group richness.

	N	Average density
All	171	1.57 ± 0.08
$N_{\text{gal}}^a = 1$	122	1.44 ± 0.09
$N_{\text{gal}} > 1$	49	1.88 ± 0.15
$N_{\text{gal}} > 2$	34	1.99 ± 0.19
$N_{\text{gal}} > 10$	10	2.72 ± 0.31

Notes. ^(a) Group richness.

Table 5. Spearman rank correlations and probability values (in parentheses) for the overlapping sources of the SDSS Main Galaxy sample and SDSS LRG sample.

	LRG LD
LD 1 ^a	0.049 (0.529)
LD 8 ^b	0.765 (0.000)
N_{gal}^c	0.226 (0.003)
D_F^d	-0.003 (0.974)
R_{vir}^e	0.150 (0.309)
M_{NFW}^f	0.288 (0.048)

Notes. Correlations in boldface have $p < 0.05$. ^(a) Luminosity-density, $1 h^{-1}$ Mpc smoothing scale. ^(b) Luminosity-density, $8 h^{-1}$ Mpc smoothing scale. ^(c) Group richness. ^(d) Distance to the closest filament axis, Mpc. ^(e) Virial radius of the group, projected harmonic mean, h^{-1} Mpc. ^(f) Total mass of the group, NFW profile, h^{-1} Mpc.

physical properties derived from the SDSS Main Galaxy sample correlate with the SDSS LRG LDF. The average densities of the LRG LDF for the whole overlapping subset and subsamples defined by the group richness are shown in Table 4, and the correlation results between the Main Galaxy sample properties and LRG luminosity-density are presented in Table 5.

Table 4 and the significant correlation in Table 5 clearly show that groups with more members reside in denser large-scale regions. A similar trend was found in Poudel et al. (2017). The LRG luminosity-density ($16 h^{-1}$ Mpc smoothing scale) correlates with the Main Galaxy sample luminosity-density with the smoothing scale of $8 h^{-1}$ Mpc, but not with the $1 h^{-1}$ Mpc smoothing scale, which is of course due to the differences in the smoothing scales. In addition, the LRG luminosity-density correlates with the total group mass, which is due to the increasing group richness towards the denser large-scale environment.

3.1.3. SDSS BOSS CMASS LDF

We have SDSS BOSS CMASS LDF data for 323 sources. There are no overlapping sources with the SDSS Main Galaxy sample or the SDSS LRG LDF since the redshift range is different (0.43–0.62). The results for the BOSS LDF sample are displayed in Table 6. The average densities of the RQ and RL subsamples are significantly different, even when taking into account the comparatively large errors¹⁵. None of the RQ sources reside in a supercluster, although the RQ subsample contains only six sources and is therefore probably not a good representative of the overall RQ population. 16% of RL sources are located in superclusters. Interestingly the average density of RS sources lies

between those of RQ and RL sources, and 7% of RS sources reside in supercluster environments. This apparent contradiction might result from the small sample size of RQ sources and the mixed nature of the RS subsample. However, in both LRG and BOSS LDFs, the fraction of RS sources in superclusters is larger than the fraction of RQ sources. The main result of the BOSS LDF analysis agrees with the LRG LDF analysis; the radio-loudness increases with the environmental density.

Due to different smoothing scales, the mean density values of the LRG and BOSS LDFs are not straightforwardly comparable, and there are no previous studies of the average densities of other types of AGN in the BOSS LDF. In the BOSS LDF only the boundary between a supercluster and the outside of a supercluster has been defined. This was done by matching the supercluster volumes with the SDSS Main Sample superclusters (Lietzen et al. 2016). It is noteworthy that at this redshift range there is a clear deficit of RQ sources and an excess of RL sources when compared to the LRG sample. This issue was discussed in Sect. 2.3.

We also compared the average densities of NLS1 galaxies with those of their candidate parent population samples from Berton et al. (2015b). Unfortunately we were only able to obtain the large-scale environment data for six sources, rendering the samples statistically insignificant.

3.2. PCA

PCA is a statistical method used to simplify large amounts of data, and to study the underlying correlations that do not show in basic correlation analyses, and might include multiple parameters. Using an orthogonal transformation it converts a set of possibly correlated variables into a set of linearly uncorrelated variables called principal components, or eigenvectors (EV). The first principal component accounts for as much of the variability in the data as possible. The second principal component has as large a variance as possible while still being orthogonal to the first principal component, and so on. This method makes it possible to find underlying connections and the most dominant variables in a data set, and possibly helps to identify the physical properties connected with each EV. While PCA is a powerful tool to study and systematise extensive data sets it should be noted that the results of PCA always depend on the set of parameters and the sample used. A good explanation and overview of the PCA can be found in Abdi & Williams (2010).

Boroson & Green (1992) used PCA to study the optical properties of 87 quasi-stellar objects (QSO); they found the EV1 to be dominated by the anticorrelation between the strength of Fe II, and the strength of [O III] $\lambda 5007$ and $FWHM(H\beta)$, while EV2 distinguished between the strength of He II $\lambda 4686$ and optical luminosity. In Boroson (2002), 75 sources were added to the original sample, yielding consistent results. They suggested that the EV1 correlations are driven by the Eddington ratio, L/L_{Edd} , and EV2 by the accretion rate. Xu et al. (2012) studied a sample of NLS1 and broad-line Seyfert 1 galaxies (BLS1) using PCA, and found their EV1 and EV2 to be in good agreement with Boroson & Green (1992) and Boroson (2002). Grupe (2004) performed PCA for a sample of 110 soft-X-ray-selected AGN – of which about half were NLS1 galaxies – and found the EV1 to be similar to the EV1 in Boroson (2002), and the EV2 to strongly correlate with the black hole mass. In Järvelä et al. (2015) we used PCA to study a pure sample of 292 NLS1 sources; in our study the EV1 was dominated by the anticorrelation between F_{O} and F_{IR} , and M_{BH} and $FWHM(H\beta)$, and it did not correlate with the Eddington ratio. In contrast, our EV2

¹⁵ Standard error of the mean.

Table 6. Average density of the whole SDSS BOSS sample and the NLS1 subsamples, and the fraction outside of and in superclusters.

	N	Average density	$LD^a < 6$ (%)	$LD > 6$ (%)
All	323	2.29 ± 0.15	92	8
RS	266	2.13 ± 0.15	93	7
RD	57	3.05 ± 0.54	86	14
RQ	6	1.12 ± 0.51	100	0
VRL	14	2.35 ± 0.64	93	7
RL and VRL	51	3.27 ± 0.60	84	16

Notes. ^(a) LD = mean luminosity-density of the LDF.

was the “traditional” EV1, distinguishing between R4570 and $FWHM(H\beta)$, and correlating strongly with the Eddington ratio. The correlation space defined by $FWHM(H\beta)$, $F([OIII]) \lambda 5007$, R4570, and C IV $\lambda 1549$ has been established as the 4DE1 parameter space explaining a host of differences observed in AGN, and possibly tracing the general AGN evolution (Marziani et al. 2006; Sulentic et al. 2007).

We performed weighted PCA separately for the LRG and the BOSS samples using the `pca`¹⁶ function in MATLAB Statistics and Machine Learning Toolbox. This is the first time a parameter describing the large-scale environment has been used in the PCA of AGN, allowing us to study its connection to the AGN correlation space. The variables were selected so that there exists no strong correlations between them, because including already correlated variables to PCA skews the results. The variables used in all PC analyses were F_O , $FWHM(H\beta)$, R4570, $F([O III])$, and the large-scale environment density, either from LRG or BOSS. Values for $FWHM(H\beta)$, R4570, and $F([O III])$ were taken from Zhou et al. (2006). In addition to these variables we used F_R in the PCA of RD sources and F_{X-ray} in the PCA of an X-ray-detected (XD) sample. With these additional samples we wanted to study whether the correlation space is consistent among samples with disparate selection criteria; the diversity of the RD sample should be kept in mind. We excluded the infrared wavebands since they are tightly correlated to the optical emission and M_{BH} which is correlated with $FWHM(H\beta)$.

The results are shown in Tables 7–12, and in Figs. 4 and 5. In the tables, the coefficients have been grouped together based on their sign, that is, whether they correlate or anticorrelate with the EV and each other. In this way, it is easier to see which properties dominate each EV and what their mutual relations are. We used the same notation in Järvelä et al. (2015). This information is also presented in biplots; the direction and the length of each vector corresponds to the level of contribution of a variable to an EV, and shows the relations between the variables. The direction of the coefficients is arbitrary, that is, the sign of the coefficient is insignificant as such and only their respective direction matters.

3.2.1. LRG PCA

We performed the LRG PCA with those sources that have all the necessary data (935 sources), and with two samples with F_R (139 sources) or F_{X-ray} (408 sources) as an additional parameter. The results for the first three principal components are displayed and the PCA coefficients listed in Tables 7–9, and the corresponding biplots are shown in Fig. 4.

EV1 accounts for 26–32% of the variance. In all samples EV1 is dominated by F_O and $F([OIII])$, which in the whole and RD samples are anticorrelated with R4570. F_R does not play

Table 7. Results of the LRG PCA: EV1.

–	Sample	+
R4570 -0.39	All (30%)	$F([OIII])$ 0.65 F_O 0.57 $FWHM(H\beta)$ 0.31 density 0.07
R4570 -0.44	RD (26%)	$F([OIII])$ 0.61 F_O 0.47 $FWHM(H\beta)$ 0.36 density 0.27 F_R 0.17
R4570 -0.06	XD (32%)	$F([OIII])$ 0.59 F_O 0.59 F_{X-ray} 0.54 density 0.10 $FWHM(H\beta)$ 0.09

Notes. The coefficients have been grouped together based on their sign.

Table 8. Results of the LRG PCA: EV2.

–	Sample	+
R4570 -0.60 F_O -0.42 $F([OIII])$ -0.27 density -0.12	All (25%)	$FWHM(H\beta)$ 0.62
F_O -0.53 R4570 -0.43 $F([OIII])$ -0.35	RD (23%)	$FWHM(H\beta)$ 0.51 F_R 0.37 density 0.14
R4570 -0.70 F_{X-ray} -0.21 density -0.10 F_O -0.05	XD (23%)	$FWHM(H\beta)$ 0.67 $F([OIII])$ 0.08

Notes. The coefficients have been grouped together based on their sign.

a significant role in EV1 of the RD sample. In the XD sample, R4570 is insignificant, and F_{X-ray} correlates with F_O and $F([OIII])$. The EV1 of this pure NLS1 sample is not exactly similar to the traditional EV1 distinguishing between $FWHM(H\beta)$ and $F([OIII])$, and R4570, although $FWHM(H\beta)$ slightly contributes to EV1 in the whole and RD samples. Interestingly, in the XD sample, $FWHM(H\beta)$ is insignificant.

EV2 23–25% of the variance is explained by EV2, which is clearly dominated by the anticorrelation of $FWHM(H\beta)$ and

¹⁶ <https://se.mathworks.com/help/stats/pca.html>

Table 9. Results of the LRG PCA: EV3.

–	Sample	+
F_O -0.09 $F([\text{OIII}])$ -0.07	All (20%)	density 0.98 $FWHM(H\beta)$ 0.14 R4570 0.05
$FWHM(H\beta)$ -0.29 $F([\text{OIII}])$ -0.04	RD (17%)	density 0.68 F_R 0.57 R4570 0.36 F_O 0.03
$F_{X\text{-ray}}$ -0.11 F_O -0.11 R4570 -0.04	XD (17%)	density 0.99 $FWHM(H\beta)$ 0.06 $F([\text{OIII}])$ 0.03

Notes. The coefficients have been grouped together based on their sign.

Table 10. Results of the BOSS PCA: EV1.

–	Sample	+
R4570 -0.29 density -0.02	All (34%)	$F([\text{OIII}])$ 0.64 F_O 0.56 $FWHM(H\beta)$ 0.44
R4570 -0.32 density -0.31 F_R -0.08	RD (28%)	$F([\text{OIII}])$ 0.59 F_O 0.57 $FWHM(H\beta)$ 0.35
density -0.27	XD (38%)	F_O 0.60 $F_{X\text{-ray}}$ 0.53 $F([\text{OIII}])$ 0.48 $FWHM(H\beta)$ 0.18 R4570 0.16

Notes. The coefficients have been grouped together based on their sign.

R4570. In the whole and RD samples F_O correlates with R4570, but in the XD sample it is negligible. Since EV2 is led by the R4570 – $FWHM(H\beta)$ anticorrelation, it is probably similar to EV1 found in Boroson & Green (1992), Boroson (2002) and Xu et al. (2012).

EV3 explains 17–20% of the variance. Density has the strongest impact on EV3 in all samples. All the other parameters are negligible, except F_R , which correlates with the density in the RD sample. This supports the results obtained in Sect. 3.1 inferring that the large-scale environment and radio properties are connected. Otherwise the fact that EV3 is completely dominated by the large-scale environment parameter, with the other variables affecting very little, further indicates that the large-scale environment density does not appear to be connected to the intrinsic properties of NLS1 galaxies, at least with this sample and set of parameters.

3.2.2. BOSS PCA

For the sources with the BOSS LDF data, we performed the PCA as described in Sect. 3.2.1, substituting the LRG LDF data with the BOSS LDF data. The whole sample consists of 312 sources, the RD sample of 51 sources, and the XD sample of 91 sources. The PCA coefficients for the first three principal components are shown in Tables 10–12, and the biplots are shown in Fig. 5.

Table 11. Results of the BOSS PCA: EV2.

–	Sample	+
R4570 -0.69 F_O -0.50 $F([\text{OIII}])$ -0.14	All (23%)	$FWHM(H\beta)$ 0.40 density 0.32
R4570 -0.64 F_O -0.36 $F([\text{OIII}])$ -0.17 density -0.03	RD (21%)	F_R 0.53 $FWHM(H\beta)$ 0.38
R4570 -0.63 $F_{X\text{-ray}}$ -0.09	XD (25%)	$FWHM(H\beta)$ 0.58 density 0.43 $F([\text{OIII}])$ 0.25 F_O 0.07

Notes. The coefficients have been grouped together based on their sign.

Table 12. Results of the BOSS PCA: EV3.

–	Sample	+
$F([\text{OIII}])$ -0.06	All (21%)	density 0.87 R4570 0.41 $FWHM(H\beta)$ 0.20 F_O 0.16
	RD (17%)	density 0.76 F_R 0.43 F_O 0.27 R4570 0.25 $FWHM(H\beta)$ 0.24 $F([\text{OIII}])$ 0.19
$F([\text{OIII}])$ -0.45	XD (14%)	density 0.51 R4570 0.50 $FWHM(H\beta)$ 0.41 $F_{X\text{-ray}}$ 0.34 F_O 0.04

Notes. The coefficients have been grouped together based on their sign.

EV1 28–38% of the variance is explained by EV1. It is similar to the EV1 of the LRG PCA; in every sample $F([\text{O III}])$ and F_O are the strongest contributors. They are anticorrelated with R4570 in the whole and RD samples, and correlated with $F_{X\text{-ray}}$ in the XD sample. Overall the R4570 contribution is weaker than in LRG PCA EV1. In the RD sample, the contribution of the density is about the same as of R4570, but neither is particularly strong.

EV2 accounts for 21–25% of the variance. The overall trend is that EV2 is led by the anticorrelation between R4570 and $FWHM(H\beta)$, suggesting that it is similar to the traditional EV1 as in the LRG PCA. In addition to this anticorrelation, F_O correlates with R4570 in the whole and RD samples. F_R has a strong contribution in the RD sample, and the density parameter comes forth in the XD sample.

EV3 14–21% of the variance is explained by EV3. In the whole and RD samples the large-scale environment density is the sole strong contributor. EV3 of the XD sample is slightly different and distinguishes between $F([\text{O III}])$, density, and R4570.

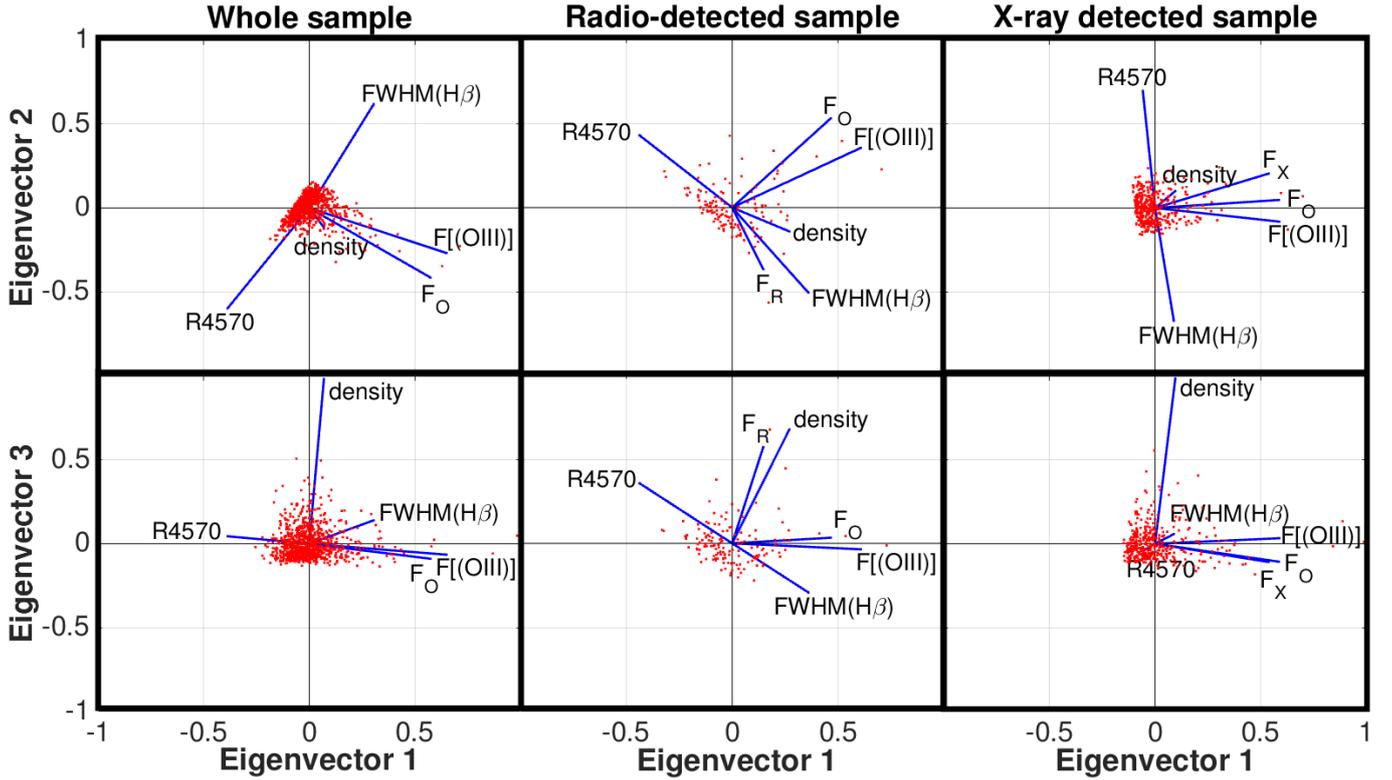


Fig. 4. Biplots of EVs 1, 2, and 3, and variables of the LRG PCA.

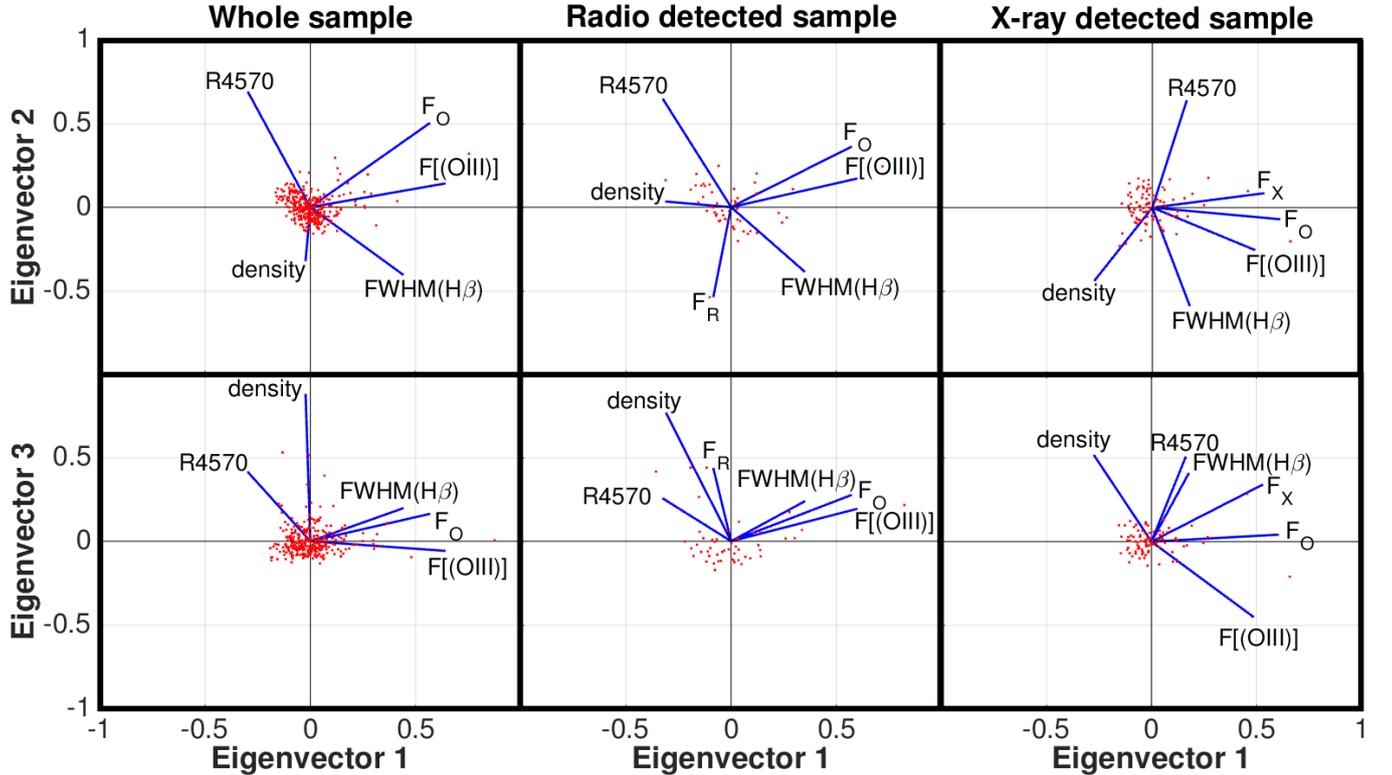


Fig. 5. Biplots of EVs 1, 2 and 3, and variables of the BOSS PCA.

3.2.3. EV correlations

In order to gain more insight into the connection between the EVs and the physical properties of our sources, we calculated the Spearman rank correlation coefficients between EVs of both

LRG and BOSS PCA, and selected intrinsic AGN properties – Eddington ratio ($L_{\text{bol}}/L_{\text{Edd}}$), M_{BH} , and λL_{5100} – for the whole, RD, and XD samples. To compute the Eddington ratio we used the estimations $L_{\text{bol}} = 9\lambda L_{5100}$ (Kaspi et al. 2000) and $L_{\text{Edd}} = 1.3 \times 10^{38} M_{\text{BH}}/M_{\odot}$ (Xu et al. 2012). Logarithmic mean,

Table 13. Spearman rank correlation between EVs and the intrinsic properties (probability values in parentheses): whole, RD and XD samples with SDSS LRG LDF data.

		$\log L/L_{\text{Edd}}$	$\log M_{\text{BH}}$	$\log \lambda L_{5100}$
EV1	all	-0.506 (0.000)	0.291 (0.000)	-0.083 (0.011)
	RD	-0.452 (0.000)	0.360 (0.000)	0.030 (0.724)
	XD	-0.080 (0.109)	0.209 (0.000)	0.118 (0.017)
EV2	all	-0.764 (0.000)	0.394 (0.000)	-0.112 (0.001)
	RD	0.717 (0.000)	-0.314 (0.000)	0.138 (0.107)
	XD	-0.827 (0.000)	0.453 (0.000)	-0.108 (0.029)
EV3	all	-0.157 (0.000)	0.100 (0.002)	-0.003 (0.924)
	RD	-0.419 (0.000)	0.294 (0.001)	0.015 (0.859)
	XD	-0.140 (0.005)	-0.042 (0.394)	-0.141 (0.005)

Notes. Correlations in boldface have $p < 0.05$.

Table 14. Spearman rank correlation between EVs and the intrinsic properties (probability values in parentheses): whole, RD and XD samples with SDSS BOSS LDF data.

		$\log L/L_{\text{Edd}}$	$\log M_{\text{BH}}$	$\log \lambda L_{5100}$
EV1	all	-0.407 (0.000)	0.749 (0.000)	0.487 (0.000)
	RD	-0.246 (0.082)	0.639 (0.000)	0.432 (0.001)
	XD	-0.100 (0.362)	0.666 (0.000)	0.769 (0.000)
EV2	all	-0.709 (0.000)	0.061 (0.286)	-0.412 (0.000)
	RD	-0.732 (0.000)	0.059 (0.678)	-0.372 (0.007)
	XD	-0.701 (0.000)	0.436 (0.000)	-0.017 (0.874)
EV3	all	-0.060 (0.294)	0.329 (0.000)	0.307 (0.000)
	RD	0.051 (0.721)	0.517 (0.000)	0.559 (0.000)
	XD	-0.354 (0.001)	0.359 (0.001)	0.189 (0.074)

Notes. Correlations in boldface have $p < 0.05$.

minimum, and maximum values for the Eddington ratio in the whole sample are -0.15 , -0.88 , and 1.20 , respectively. The correlation results are shown in Tables 13 and 14.

LRG EV1 correlates with the Eddington ratio to some extent in the whole and RD samples, but not in the XD sample. There is a weak correlation with the black hole mass in all samples. EV2 of our sample is closest to the “traditional EV1”, and correlates strongly with the Eddington ratio in all samples, and significantly, but not so strongly, with the black hole mass. EV3 does not correlate strongly with any of the properties, which is understandable since the large-scale environment does not directly affect any singular properties, but rather the overall evolution of the galaxies. The only exception to this is the RD sample in which EV3 shows some correlation with the Eddington ratio.

BOSS EV1 correlates strongly with M_{BH} and is clearly comparable to EV2 of previous studies. There is a slight correlation with the continuum luminosity, λL_{5100} , especially in the XD sample. In addition, the whole sample correlates with the Eddington ratio to some extent. The correlation of EV2 with the Eddington ratio is strong in all samples, indicating that it is the main driver behind this EV. EV2 is clearly the traditional AGN EV1. EV3 correlations are rather ambiguous, which can be explained by the strong contribution of the density – which does not correlate with any other properties – to EV3. The weak correlations seen between EV3 and the intrinsic properties are probably induced by other parameters and not the density.

4. Discussion

4.1. NLS1 galaxies in the AGN family

NLS1 galaxies have been found to be morphologically young, late-type sources in previous studies concentrating on smaller samples. Our findings support this scenario. In the SDSS Main Galaxy sample, 23% of NLS1 sources reside in groups. This number is low compared to galaxies in general; in the flux-limited SDSS Main Galaxy sample 48% of galaxies are in groups (Tempel et al. 2014b), and according to Bahcall (1996) 60% of galaxies are in groups or clusters, leaving 40% of galaxies to be in pairs or field galaxies. The excess of field NLS1 galaxies is consistent with their young age. Also in the framework of the density – morphology relation (Hubble & Humason 1931; Dressler 1980) – more evolved galaxies are found in denser regions – our study confirms the young nature of the NLS1 population. The average total mass of the groups with 4–15 members in Tempel et al. (2014b) is $\sim 10^{13} h^{-1} M_{\odot}$, which is of the same order as for more luminous quasars (Mandelbaum et al. 2009; Koutoulidis et al. 2013; Fanidakis et al. 2013; Shen et al. 2013) and approximately the same as for the groups our NLS1 galaxies reside in.

Our results clearly distinguish between NLS1 and BLS1 sources in terms of average densities. This discrepancy contradicts the simple orientation-based unification model of NLS1 and BLS1 galaxies in which, when assuming a disk-like BLR, the observed differences can be explained as orientation effects, and NLS1 sources can be unified with BLS1 sources (Decarli et al. 2008; Rakshit et al. 2017). If BLS1 galaxies were actually the parent population of NLS1 galaxies, we would expect their spatial distributions to be similar. Our findings also contradict the results in Ermash (2014) who used SDSS DR7 to study the spatial densities of NLS1 and BLS1 galaxies as functions of the large-scale galaxy density. They found that the ratio $N_{\text{NLS1}}/N_{\text{BLS1}}$ is constant and does not depend on large-scale galaxy density. We do not find the fraction to be constant.

However, BLS1 galaxies are diverse in their properties; for example, their bulges can be classical, composite, or pseudobulges, whereas the bulges in the majority of NLS1 galaxies are pseudobulges. This diversity has not been taken into account in most papers studying, and finding, differences between NLS1 and BLS1 sources (e.g. Crenshaw et al. 2003; Deo et al. 2006; Sani et al. 2010), nor was it taken into account in Lietzen et al. (2011). It is, therefore, possible that a subset of BLS1 galaxies is part of the parent population of NLS1 galaxies.

A similar challenge due to heterogeneous samples has been encountered before; Tran (2003) suggested that studies of the differences between Sy1 and Sy2 galaxies should be reanalysed taking into account the fact that Sy2 galaxies are a heterogeneous class consisting of Sy2 sources with powerful hidden Sy1 nuclei and of “pure” Sy2 sources with weak nuclei and very weak or nonexistent BLR (Heisler et al. 1997; Tran 2001, 2003; Marinucci et al. 2012). They argue that, for example, Schmitt et al. (2001) did not find differences because due to selection effects they compare Sy1 and hidden broad-line region (HBLR) Sy2 galaxies, which by definition should be similar. They reanalysed some of the studies and found that when Sy2 sources are divided into two populations, the results change, revealing Sy1 and HBLR Sy2 sources to be similar, and non-HBLR Sy2 sources to be different. They suggest an evolutionary link between pure Sy2 galaxies, and HBLR Sy2 and Sy1 galaxies; either non-HBLR Sy2 sources evolve to more powerful HBLR Sy2 and Sy1 sources once the AGN activity is triggered (e.g. Nicastro 2000; Wu et al. 2011), or non-HBLR Sy2

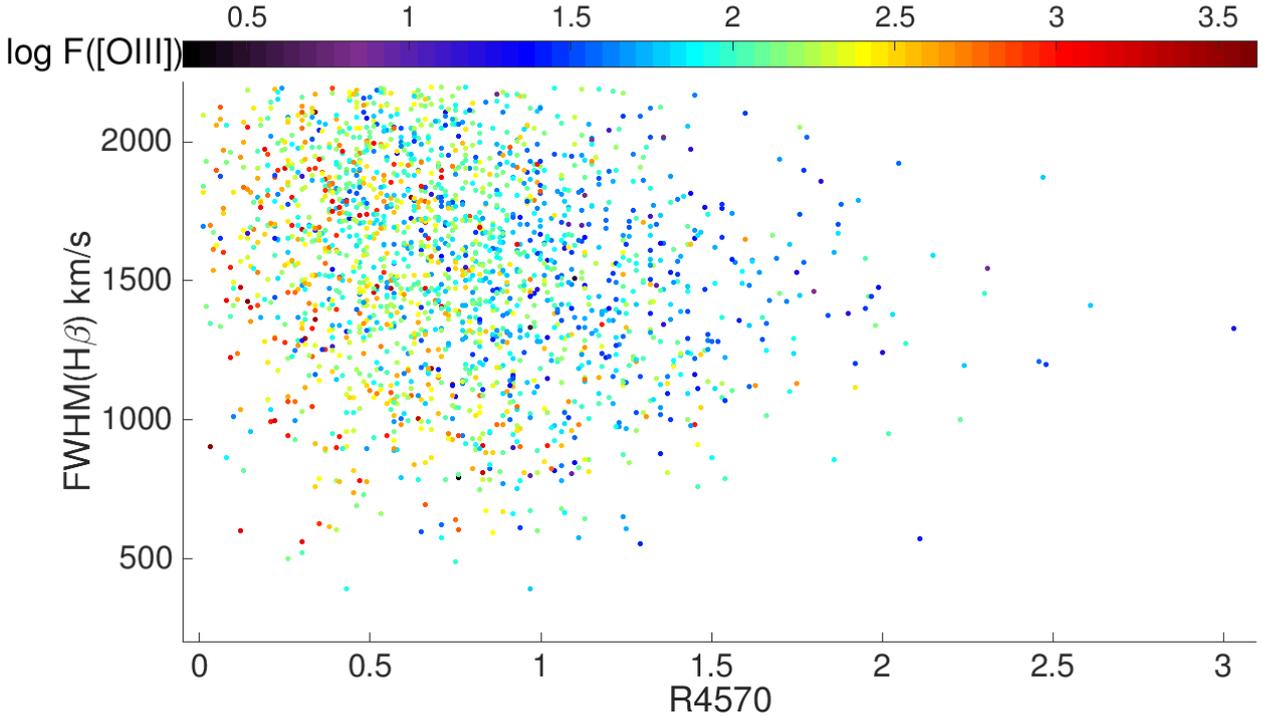


Fig. 6. Narrow $FWHM(H\beta)$ extreme of the 4DE1 continuum, populated by NLS1 galaxies which follow the same trends as AGN in general. Abscissa is the strength of Fe II compared to the broad $H\beta$, and ordinate is the $FWHM(H\beta)$. The flux density of [O III] is shown in colour; bluer colour indicates lower flux density and redder colour higher flux density.

galaxies have exhausted their fuel and returned to a dormant state (Yu & Hwang 2011; Yu et al. 2013). A similar evolutionary scenario has been proposed for NLS1 and BLS1 galaxies; assuming that the narrow broad lines are due to the undermassive BH, the fast growing NLS1 galaxies would eventually evolve to Sy1 galaxies with higher-mass black holes and thus broader broad lines, that is, BLS1 galaxies. Our results favour this idea over the unification by orientation scenario.

Moreover, a few higher-inclination steep-spectrum NLS1 sources with extended radio emission have recently been found (Doi et al. 2015; Congiu et al. 2017), proving that in these sources the narrowness of the broad lines is due to the low black hole mass and not orientation effects, supporting the scenario that no additional “misaligned” broad-line Type 1 parent population is necessarily needed. However, the parent population studies are still in the early phase and, for example, edge-on compact steep-spectrum (CSS) sources and disk-hosted radio galaxies have been proposed “to be possible” Type 1 parent populations of jetted NLS1 sources (Berton et al. 2015b, 2016, 2017). The nature of the obscured, or Type 2, parent population remains unclear; for the jetted sources it might include narrow-line CSS sources and narrow-line disk-hosted radio galaxies. Zhang & Wang (2006) suggested that non-HBLR Sy2 galaxies could be the Type 2 counterparts of NLS1 galaxies and unified based on orientation. This result, however, remains debatable since the sources have many unexplained differences, for example, non-HBLR Sy2 galaxies do not show Fe II emission lines (Yu et al. 2013).

Our analysis based on PCA supports the idea that NLS1 sources are a part of the continuous AGN spectrum; EVs 1 and 2 found in this study are a manifestation of the 4DE1 AGN correlation space – in our case without CIV – dominated by the interplay between $FWHM(H\beta)$, R4570, and [O III]. These three parameters were studied in more detail, for

example in Shen & Ho (2014). They suggest that the variations in $FWHM(H\beta)$, R4570, and [O III], and the correlations between them, can be explained based only on the Eddington ratio and the orientation. The increasing Eddington ratio leads to increasing R4570 and decreasing [O III] strength, and the scatter in $FWHM(H\beta)$ at a fixed R4570 is mainly due to the orientation of a flattened BLR, and not the black hole mass and virialised BLR. Their sample consists mainly of broad-line AGN. NLS1 galaxies represent the narrow $FWHM(H\beta)$, strong Fe II, and weak [O III] extreme of this continuum. In Fig. 6 we show a plot similar to Fig. 1 in Shen & Ho (2014). NLS1 sources follow the same trend as broad-line AGN and complete one extreme of the continuum; the flux density of [O III] decreases as the Fe II strength increases, and there does not seem to exist any vertical trends correlations with the $FWHM(H\beta)$. The same domination of $FWHM(H\beta)$, R4570 and [O III] in the PCA can be seen in Figs. 4 and 5. We cannot, however, confirm the role of the orientation as the driver of the scatter in $FWHM(H\beta)$ values since our black hole mass estimate is orientation-dependent. If the $FWHM(H\beta)$ depended only on the orientation then BLS1 sources should be the parent population of NLS1 galaxies, which, based on our results, is improbable.

4.2. NLS1 galaxies as a class

In the SDSS Main Galaxy sample only 25% of RD and 34% of RS NLS1 galaxies reside in groups. The average richness of the groups is similar, but RD sources reside in physically slightly smaller, and thus denser, groups. This indicates that not the group membership itself but the density of the group is correlated with the radio properties of NLS1 galaxies.

The density of the large-scale environment has the capacity of transfiguring the radio characteristics of NLS1 sources, probably via interactions or merging; the incidence of RL sources

increases with the increasing large-scale environment density. However, the average black hole masses of the subsamples – RS, RQ, and RL – within a certain redshift range are almost similar. This suggests that the environment is indeed causing the differences in the radio loudness. Radio loudness does not correlate with the black hole mass, or, at a given redshift; NLS1 sources with more massive black holes are not more probable to be radio loud. However, on average the black hole mass in RL sources is higher because they are farther away; there is a clear excess of RL and deficit of RQ sources at high redshift. This fits well with the hierarchical evolution model of AGN, discussed in Sect. 2.3.

The average black hole masses within the different density regions are similar. For example, in SDSS LRG LDF, RL sources have similar average black hole masses in voids, intermediate density regions, and superclusters. The same holds for RQ and RS sources. This might suggest that, at least up to this point, the accretion history of these sources has been similar in all density regions.

The distribution of sources to voids, intermediate density regions, and superclusters changes with increasing radio loudness; RL sources have a higher probability of being found in superclusters. An exception to this are RS sources. The biggest fraction of RS sources resides in voids; these are probably the “true” RS NLS1 galaxies. However, in superclusters, the fraction is clearly higher than for RQ sources. This might indicate, and be explained by the misclassification of a fraction of RS sources, meaning that the subsample is mixed. The two-sample KS test indeed suggests that the density distributions of our subsamples – RS versus RD, RQ versus RL, and RS versus RL – are drawn from the same continuous distribution, which further indicates that the subsamples are not well defined.

Misclassifications and thus heterogeneous samples are one of the biggest caveats in extensive AGN population studies, and one of the reasons why large statistical studies of NLS1 galaxies and also other AGN classes give inconsistent results. In our study there are a few possible reasons for misclassifications. Data are incomplete, even though optical data cannot be incomplete since NLS1 sources are selected based on that, but we might lose some weak radio emitters from the RQ sample to the RS sample. Due to enhanced star formation, the RL subsample probably includes a few RQ and RS sources, and the RQ subsample might include RS sources. Also the radio loudness value is time-dependent, because NLS1 sources are variable in optical and radio bands (e.g. Lähteenmäki et al. 2017; Gabányi et al. 2017). Moreover, there is clear evidence that some RS sources have jets and are thus severely misclassified (Lähteenmäki et al., in prep.).

PCA results of the whole, RD, and XD samples are mostly consistent. In the majority of cases the first two EVs are dominated by the 4DE1 components; R4570, $FWHM(H\beta)$, and $F([O\ III])$, and the third EV by the density parameter. The appearance of the “traditional” EV1, R4570 – $FWHM(H\beta)$, only as the second most important principal component is probably due to our pure NLS1 sample; the variance of $FWHM(H\beta)$ is small compared to mixed samples, but the variance in R4570 and $F([O\ III])$ is of comparable magnitude. It has been suggested that the anticorrelation of R4570 and $F([O\ III])$ is driven by the Eddington ratio, and the variance in $FWHM(H\beta)$ is only due to the orientation (Shen & Ho 2014). Interestingly, in our sample, the Eddington ratio correlates poorly with EV1 that is dominated by the R4570 – $F([O\ III])$ anticorrelation. Instead, it correlates very well with EV2, dominated by R4570 and $FWHM(H\beta)$. To some extent this might be because $FWHM(H\beta)$ is used to estimate the Eddington ratio.

Adding F_R or $F_{X\text{-ray}}$ to PCA changes the EVs slightly. Radio flux density emerges in EVs 2 and 3, correlating with $FWMH(\beta)$ and the density parameter, respectively. Correlation of the radio flux density and $FWMH(\beta)$ supports the hypothesis that radio brighter-, and also radio louder- (since radio loudness correlates with the radio flux density) sources have on average broader $FWMH(H\beta)$. $FWMH(H\beta)$ does not evolve with the redshift so this correlation is not induced by an increasing fraction of RL sources being farther away. Correlation of F_R and the density support the results from the average density calculations of the different subsamples. In the XD sample $F_{X\text{-ray}}$ and F_O emerge in the first EV. This might be induced by the initial correlation of $F_{X\text{-ray}}$ and F_O , which then enhances variance in their direction.

Overall the density parameter appears to have more impact on BOSS PCA than on LRG PCA. BOSS LDF has a smaller smoothing scale ($8\ h^{-1}\ \text{Mpc}$) than LRG LDF ($16\ h^{-1}\ \text{Mpc}$) which means that it traces the structures at a smaller scale, which might induce the difference in PCA. Another possibility is that the LRG and BOSS samples are intrinsically slightly different; the excess of RL sources at higher redshifts might cause the density to appear more important in the PCA.

4.3. From RL to jetted NLS1 galaxies

Using the more or less arbitrarily defined value of radio loudness as a dividing parameter between NLS1 subclasses results in poorly defined, mixed samples. This affects the results. The subsamples should be divided based on their real physical properties, for example, jetted and non-jetted sources, as proposed in Padovani (2016). This would require extensive studies of individual sources, which in the case of extensive samples is not viable, and the radio loudness parameter must thus be used as a proxy of the nuclear activity. However, we can begin to address this issue by studying the environments of those sources that are known to have jets.

For this we selected a sample of NLS1 galaxies whose radio emission should be dominated by non-thermal emission from the AGN. We included all gamma-ray detected NLS1 galaxies and NLS1 sources which have kpc- or pc-scale jets. In addition we included sources detected at the Metsähovi Radio Observatory at 37 GHz (Lähteenmäki et al. 2017; and Lähteenmäki et al., in prep.), and sources in which, based on the $q22$ parameter (Caccianiga et al. 2015), the radio emission is dominated by the jet emission.

The $q22$ parameter can be used to distinguish sources in which the radio loudness is due to the jet from sources in which it might be induced by enhanced star formation (Caccianiga et al. 2015). It is defined as $q22 = \log(F_{22\ \mu\text{m}}/F_{1.4\ \text{GHz}})$, and based on the fact that jet-dominated and star formation-dominated sources have different radio-to-mid-infrared flux ratios. In sources with $q22 < -0.8$ the jet is probably the main source of the radio emission; in sources with $q22 > 1$ the star formation may contribute significantly to the observed radio emission, especially in sources which have excess mid-infrared emission, that is, $W3\text{--}W4 > 2.5$. Sources in between, with $-0.8 < q22 < 1$, have probably both the star formation and the jet contributing to the radio emission.

Altogether, we have 29 sources that supposedly have a jet and for which we have the large-scale environment data; 17 of those lie in the LRG LDF and 12 in the BOSS LDF. The Metsähovi sample includes sources that were initially selected for the observing programme due to their dense environments; two of these have been detected and might considerably increase

Table 15. Average large-scale environment densities of jetted, mixed emission, and non-jetted subsamples of NLS1 galaxies.

	<i>N</i>	Average density
Jetted, LRG	17	2.48 ± 0.57
Jetted, LRG, w/o density-selected	15	1.75 ± 0.31
Mixed, LRG	82	1.68 ± 0.15
Star formation, LRG	57	1.44 ± 0.14
Jetted, BOSS	12	2.05 ± 0.81

the average density shown below. The average densities for the jetted NLS1 galaxies in the LRG sample, with and without the two density-selected sources, and in the BOSS sample are presented in Table 15. For comparison, the average densities for the mixed and star formation-dominated subsamples are presented.

The subsample sizes are relatively small and thus the results are not unambiguous, however, the average density of the jetted NLS1 sources in the LRG LDF is almost similar to the traditionally defined RL NLS1 sources (1.71 ± 0.14). Compared to mixed, and especially the star formation dominated sources, the jetted sources are more likely to reside in denser environments. This result is consistent with the results of the various RL subsamples, since it is probable that the increasing radio loudness increases the probability that the radio emission originates in the jet. It is noteworthy that the average density of the star formation-dominated subsample is similar to that of the RS subsample. It is understandable since in these sources the contribution of the AGN to the radio emission might be negligible or non-existent, and the AGN in them is probably radio quiet or silent. This result does not exclude the possibility that jetted sources could have enhanced star formation, and thus does not tell us anything about how the sources with different amounts of star formation are distributed to voids, intermediate density regions, and superclusters.

It is not surprising that NLS1 galaxies reside in less dense supercluster-scale environments when compared to the AGN samples in Lietzen et al. (2011); most of their samples consist of source types that are considered to be evolved, whereas the majority of NLS1 galaxies are thought to be young, unevolved spiral galaxies. The result supports the hierarchical galaxy evolution model in which the more evolved galaxies, and also AGN, are found in denser environments. The pronounced difference between the environments of the jetted NLS1 galaxies and other jetted AGN classes studied in Lietzen et al. (2011) is particularly interesting. Their samples included BL Lac objects (2.50 ± 0.20), flat-spectrum radio galaxies (2.60 ± 0.07), FR I radio galaxies (3.01 ± 0.07), and FR II radio galaxies (3.20 ± 0.04), all of which have significantly higher average large-scale environment densities than the jetted NLS1 galaxies (1.75 ± 0.31). This confirms that triggering a jet in diverse environments, spanning from voids to superclusters, is possible. Whereas the results indicate that the denser large-scale environment affects the incidence of jets, the triggering mechanism remains unclear. A small percentage (8–16%, Ohta et al. 2007) of NLS1 galaxies shows signs of interaction or merging (20–30% for Seyfert galaxies in general, Schmitt 2001), suggesting that secular processes dominate the growth and evolution of NLS1 sources. Extensive studies of the host galaxies of the different, especially the jetted, NLS1 sources are needed to examine whether a link between jets, and interaction and merging exists. The error in the average density of jetted BOSS sources is relatively large, and unfortunately we do not have any comparison samples, but the results are generally consistent with the LRG LDF results.

4.4. Conclusions

Our study is the first attempt to study the relationship of the large-scale environment and the intrinsic properties of NLS1 galaxies using a large, statistically significant sample. The large-scale environment is an interesting additional parameter helping us to understand how the changes at the largest cosmic scales emerge in the smaller-scale properties of AGN; for example, their host galaxies and nuclear activity. Our main conclusions are the following.

- This study supports the young, unevolved nature of NLS galaxies, and shows that the large-scale environment density affects their radio properties; RL, and also jetted NLS1 sources preferably reside in denser regions compared to RQ or non-jetted NLS1 sources.
- NLS1 galaxies are a diverse, heterogeneous class of sources which suffers from a considerable amount of misclassification. Our results support the Padovani’s view that the traditional division to RL and RQ/RS sources is a severe problem and should be replaced with the division to jetted and non-jetted NLS1 galaxies.
- Orientation is probably not a sufficient explanation for the differences in NLS1 and BLS1 galaxies. NLS1 galaxies are a distinct class of sources, and BLS1 galaxies are not their parent population. However, they could be evolutionarily connected.
- NLS1 sources are located in one extreme of the AGN 4DE1 continuum, but they clearly are distinctive in their nature among the gamma-ray emitting AGN.

Our results indicate that some of the properties of NLS1 galaxies might be affected by the large-scale environment, but it is likely that the large-scale environment has an effect at the scale of the whole galaxy evolution and is not clearly connected to the specific intrinsic properties of NLS1 galaxies, or AGN in general; it affects the long time-scale evolution of galaxies, whereas the nuclear activity of galaxies is variable and intermittent at timescales shorter than galaxy evolution. Since all of the sources in our sample are NLS1 sources and close to one another in the evolutionary sequence, the differences between the sources are not pronounced enough for drawing definite conclusions of the effects of the large-scale environment. A more diverse sample with various AGN classes is needed to study this more effectively.

In the future, the environment studies should be extended to include smaller-scale environments; for example, the group and cluster scales, which have a bigger contemporary effect on the individual galaxies than the large-scale environment. Another crucial part of the environment research, and closely connected to the local environment, is to study the host galaxy morphologies of an extensive and diverse sample of NLS1 galaxies. Gathering data at all scales of environments will help us to create a bigger picture of the connection of AGN evolution and the environment, and give insight into the triggering mechanism of the jet. Multiwavelength observations, including interferometry, together with environmental data, will be essential in correctly classifying the various NLS1 sources. This is crucial for population-wide studies, including the issues concerning the heterogeneity of the NLS1 class and the intra-class evolution, and will aid in the search of the parent population.

Acknowledgements. H.L. is funded by PUT1627 grant from Estonian Research Council. M.E. was supported by the ETAG project IUT26-2, and by the European Structural Funds grant for the Centre of Excellence “Dark Matter

in (Astro)particle Physics and Cosmology” TK133. This research made use of the NASA/IPAC Extragalactic Database (NED) which is operated by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. This publication makes use of data products from the Wide-field Infrared Survey Explorer, which is a joint project of the University of California, Los Angeles, and the Jet Propulsion Laboratory/California Institute of Technology, funded by the National Aeronautics and Space Administration. Funding for the Sloan Digital Sky Survey (SDSS) has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the US Department of Energy, the Japanese Monbukagakusho, and the Max Planck Society. The SDSS Web site is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are The University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation Group, The Johns Hopkins University, the Korean Scientist Group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington. This research made use of the ROSAT All-Sky Survey data which have been processed at MPE.

References

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, **182**, 543
- Abdi, H., & Williams, L. J. 2010, Wiley Interdisciplinary Reviews: Computational Statistics, **2**, 433
- Abdo, A. A., Ackermann, M., Ajello, M., et al. 2009, *ApJ*, **699**, 976
- Ahn, C. P., Alexandroff, R., Allende Prieto, C., et al. 2014, *ApJS*, **211**, 17
- Alam, S., Albareti, F. D., Allende Prieto, C., et al. 2015, *ApJS*, **219**, 12
- Antón, S., Browne, I. W. A., & Marchã, M. J. 2008, *A&A*, **490**, 583
- Bahcall, N. A. 1996, ArXiv e-prints [[arXiv:astro-ph/9611148](https://arxiv.org/abs/astro-ph/9611148)]
- Barth, A. J., Bentz, M. C., Greene, J. E., & Ho, L. C. 2008, *ApJ*, **683**, L119
- Beckmann, V., & Shrader, C. R. 2012, Active Galactic Nuclei (Wiley-VCH Verlag)
- Bentz, M. C., Peterson, B. M., Pogge, R. W., & Vestergaard, M. 2009, *ApJ*, **694**, L166
- Berton, M., Foschini, L., Caccianiga, A., et al. 2015a, in The Many Facets of Extragalactic Radio Surveys: Towards New Scientific Challenges, **75**
- Berton, M., Foschini, L., Ciroi, S., et al. 2015b, *A&A*, **578**, A28
- Berton, M., Caccianiga, A., Foschini, L., et al. 2016, *A&A*, **591**, A98
- Berton, M., Foschini, L., Caccianiga, A., et al. 2017, ArXiv e-prints [[arXiv:1705.07905](https://arxiv.org/abs/1705.07905)]
- Blanton, M. R., & Roweis, S. 2007, *AJ*, **133**, 734
- Boroson, T. A. 2002, *ApJ*, **565**, 78
- Boroson, T. A., & Green, R. F. 1992, *ApJS*, **80**, 109
- Caccianiga, A., Antón, S., Ballo, L., et al. 2015, *MNRAS*, **451**, 1795
- Chen, Y.-C., Ho, S., Mandelbaum, R., et al. 2017, *MNRAS*, **466**, 1880
- Cisternas, M., Jahnke, K., Inskip, K. J., et al. 2011, *ApJ*, **726**, 57
- Congiu, E., Berton, M., Giroletti, M., et al. 2017, *A&A*, **603**, A32
- Corbin, M. R. 2000, *ApJ*, **536**, L73
- Crenshaw, D. M., Kraemer, S. B., & Gabel, J. R. 2003, *AJ*, **126**, 1690
- Decarli, R., Dotti, M., Fontana, M., & Haardt, F. 2008, *MNRAS*, **386**, L15
- Decarli, R., Dotti, M., & Treves, A. 2011, *MNRAS*, **413**, 39
- Deo, R. P., Crenshaw, D. M., & Kraemer, S. B. 2006, *AJ*, **132**, 321
- Doi, A., Nagira, H., Kawakatu, N., et al. 2012, *ApJ*, **760**, 41
- Doi, A., Asada, K., Fujisawa, K., et al. 2013, *ApJ*, **765**, 69
- Doi, A., Wajima, K., Hagiwara, Y., & Inoue, M. 2015, *ApJ*, **798**, L30
- Dressler, A. 1980, *ApJ*, **236**, 351
- Ebeling, H., Stephenson, L. N., & Edge, A. C. 2014, *ApJ*, **781**, L40
- Einasto, M., Lietzen, H., Tempel, E., et al. 2014, *A&A*, **562**, A87
- Ellison, S. L., Patton, D. R., Mendel, J. T., & Scudder, J. M. 2011, *MNRAS*, **418**, 2043
- Ermash, A. A. 2014, *Astron. Rep.*, **58**, 205
- Fabian, A. C. 2012, *ARA&A*, **50**, 455
- Fanidakis, N., Macciò, A. V., Baugh, C. M., Lacey, C. G., & Frenk, C. S. 2013, *MNRAS*, **436**, 315
- Foschini, L., Colpi, M., Gallo, L., et al. 2011, in Narrow-Line Seyfert 1 Galaxies and their Place in the Universe, PoS(NLS1)024
- Foschini, L., Berton, M., Caccianiga, A., et al. 2015, *A&A*, **575**, A13
- Gabányi, K. É., Frey, S., Paragi, Z., et al. 2017, *MNRAS*, submitted
- Gliozzi, M., Papadakis, I. E., Grupe, D., et al. 2010, *ApJ*, **717**, 1243
- Goodrich, R. W. 1989, *ApJ*, **342**, 224
- Greene, J. E., & Ho, L. C. 2005, *ApJ*, **630**, 122
- Grupe, D. 2004, *AJ*, **127**, 1799
- Gu, M., Chen, Y., Komossa, S., et al. 2015, *ApJS*, **221**, 3
- Heisler, C. A., Lumsden, S. L., & Bailey, J. A. 1997, *Nature*, **385**, 700
- Hubble, E., & Humason, M. L. 1931, *ApJ*, **74**, 43
- Järvelä, E., Lähteenmäki, A., & León-Tavares, J. 2015, *A&A*, **573**, A76
- Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, *ApJ*, **533**, 631
- Khachikian, E. Y., & Weedman, D. W. 1974, *ApJ*, **192**, 581
- King, A., & Pounds, K. 2015, *ARA&A*, **53**, 115
- Kocevski, D. D., Faber, S. M., Mozena, M., et al. 2012, *ApJ*, **744**, 148
- Komossa, S., Voges, W., Xu, D., et al. 2006, *AJ*, **132**, 531
- Koutoulidis, L., Plionis, M., Georgantopoulos, I., & Fanidakis, N. 2013, *MNRAS*, **428**, 1382
- Kuutma, T., Tamm, A., & Tempel, E. 2017, *A&A*, **600**, L6
- Lähteenmäki, A., Järvelä, E., Hovatta, T., et al. 2017, *A&A*, **603**, A100
- Laor, A. 2001, *ApJ*, **553**, 677
- León Tavares, J., Kotilainen, J., Chavushyan, V., et al. 2014, *ApJ*, **795**, 58
- Lietzen, H., Heinämäki, P., Nurmi, P., et al. 2011, *A&A*, **535**, A21
- Lietzen, H., Tempel, E., Heinämäki, P., et al. 2012, *A&A*, **545**, A104
- Lietzen, H., Tempel, E., Liivamägi, L. J., et al. 2016, *A&A*, **588**, L4
- Liivamägi, L. J., Tempel, E., & Saar, E. 2012, *A&A*, **539**, A80
- Lister, M. L., Aller, M. F., Aller, H. D., et al. 2016, *AJ*, **152**, 12
- Mandelbaum, R., Li, C., Kauffmann, G., & White, S. D. M. 2009, *MNRAS*, **393**, 377
- Marinucci, A., Bianchi, S., Nicastro, F., Matt, G., & Goulding, A. D. 2012, *ApJ*, **748**, 130
- Marziani, P., Dultzin-Hacyan, D., & Sulentic, J. W. 2006, in Accretion onto Supermassive Black Holes in Quasars: Learning from Optical/UV Observations, ed. P. V. Kreitler (Nova Science Publishers), 123
- Mathur, S. 2000, *MNRAS*, **314**, L17
- Mathur, S., Kuraszkiwicz, J., & Czerny, B. 2001, *New Astron.*, **6**, 321
- Nicastro, F. 2000, *ApJ*, **530**, L65
- Ohta, K., Aoki, K., Kawaguchi, T., & Kiuchi, G. 2007, *ApJS*, **169**, 1
- Olguín-Iglesias, A., Kotilainen, J. K., León Tavares, J., Chavushyan, V., & Añorve, C. 2017, *MNRAS*, **467**, 3712
- Osterbrock, D. E. 1978, *Proc. the National Academy of Science*, **75**, 540
- Osterbrock, D. E. 1981, *ApJ*, **249**, 462
- Osterbrock, D. E., & Pogge, R. W. 1985, *ApJ*, **297**, 166
- Padovani, P. 2016, *A&ARv*, **24**, 13
- Pandey, B., & Sarkar, S. 2017, *MNRAS*, **467**, L6
- Park, C., & Choi, Y.-Y. 2009, *ApJ*, **691**, 1828
- Peterson, B. M., McHardy, I. M., Wilkes, B. J., et al. 2000, *ApJ*, **542**, 161
- Poudel, A., Heinämäki, P., Tempel, E., et al. 2017, *A&A*, **597**, A86
- Pović, M., Sánchez-Portal, M., Pérez García, A. M., et al. 2012, *A&A*, **541**, A118
- Rakshit, S., Stalin, C. S., Chand, H., & Zhang, X.-G. 2017, *ApJS*, **229**, 39
- Richards, J. L., & Lister, M. L. 2015, *ApJ*, **800**, L8
- Richards, J. L., Lister, M. L., Savolainen, T., et al. 2015, in Extragalactic Jets from Every Angle, eds. F. Massaro, C. C. Cheung, E. Lopez, & A. Siemiginowska, *IAU Symp.*, **313**, 139
- Ryan, C. J., De Robertis, M. M., Virani, S., Laor, A., & Dawson, P. C. 2007, *ApJ*, **654**, 799
- Sani, E., Lutz, D., Risaliti, G., et al. 2010, *MNRAS*, **403**, 1246
- Schmitt, H. R. 2001, *AJ*, **122**, 2243
- Schmitt, H. R., Antonucci, R. R. J., Ulvestad, J. S., et al. 2001, *ApJ*, **555**, 663
- Shen, Y., & Ho, L. C. 2014, *Nature*, **513**, 210
- Shen, Y., McBride, C. K., White, M., et al. 2013, *ApJ*, **778**, 98
- Simpson, C. 1998, *MNRAS*, **297**, L39
- Spergel, D. N., Bean, R., Doré, O., et al. 2007, *ApJS*, **170**, 377
- Steinhauser, D., Schindler, S., & Springel, V. 2016, *A&A*, **591**, A51
- Storchi-Bergmann, T. 2008, in *Rev. Mex. Astron. Astrofis. Conf. Ser.*, **32**, 139
- Sulentic, J. W., Bachev, R., Marziani, P., Negrete, C. A., & Dultzin, D. 2007, *ApJ*, **666**, 757
- Taniguchi, Y. 1999, *ApJ*, **524**, 65
- Tempel, E., Stoica, R. S., Martínez, V. J., et al. 2014a, *MNRAS*, **438**, 3465
- Tempel, E., Tamm, A., Gramann, M., et al. 2014b, *A&A*, **566**, A1
- Tran, H. D. 2001, *ApJ*, **554**, L19
- Tran, H. D. 2003, *ApJ*, **583**, 632
- Treister, E., Schawinski, K., Urry, C. M., & Simmons, B. D. 2012, *ApJ*, **758**, L39
- Urrutia, T., Lacy, M., & Becker, R. H. 2008, *ApJ*, **674**, 80
- van de Ven, G., & Fathi, K. 2010, *ApJ*, **723**, 767
- Veron-Cetty, M. P., & Veron, P. 2003, *VizieR Online Data Catalog: VII/235*
- Villforth, C., Hamann, F., Rosario, D. J., et al. 2014, *MNRAS*, **439**, 3342
- Whalen, D. J., Laurent-Muehleisen, S. A., Moran, E. C., & Becker, R. H. 2006, *AJ*, **131**, 1948
- Wolf, C. 2005, *Mem. Soc. Astron. Ital.*, **76**, 21
- Woo, J.-H., Yoon, Y., Park, S., Park, D., & Kim, S. C. 2015, *ApJ*, **801**, 38

- Wu, X.-B., Wang, R., Kong, M. Z., Liu, F. K., & Han, J. L. 2004, *A&A*, 424, 793
- Wu, Y.-Z., Zhang, E.-P., Liang, Y.-C., Zhang, C.-M., & Zhao, Y.-H. 2011, *ApJ*, 730, 121
- Xu, D., Komossa, S., Zhou, H., et al. 2012, *AJ*, 143, 83
- York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, *AJ*, 120, 1579
- Yu, P.-C., & Hwang, C.-Y. 2011, *AJ*, 142, 14
- Yu, P.-C., Huang, K.-Y., Hwang, C.-Y., & Ohyama, Y. 2013, *ApJ*, 768, 30
- Yuan, W., Zhou, H. Y., Komossa, S., et al. 2008, *ApJ*, 685, 801
- Zhang, E.-P., & Wang, J.-M. 2006, *ApJ*, 653, 137
- Zhou, H., Wang, T., Yuan, W., et al. 2006, *ApJS*, 166, 128
- Zhou, H., Wang, T., Yuan, W., et al. 2007, *ApJ*, 658, L13

Appendix A: The LDF

We calculated the luminosities of the galaxies using k -corrections from the `kcorrect` algorithm (Blanton & Roweis 2007). To set the mean luminosity the same through the whole redshift range, we calculated a distance-dependent weighting factor $W_L(d)$ for the luminosities of galaxies. We calculated the mean luminosity density as a function of distance, smoothed it, and used this as a weight. We then calculated the weighted luminosity for each galaxy as

$$L_{\text{gal,w}} = W_L(d)L_{\text{gal}}, \quad (\text{A.1})$$

where $L_{\text{gal}} = L_{\odot}10^{0.4(h^{-1}M_{\odot}-M)}$ is the observed luminosity of a galaxy with the absolute magnitude M , and M_{\odot} is the absolute magnitude of the Sun.

The coordinates of galaxies were transformed into Cartesian coordinates defined as

$$x = -d \sin \lambda \quad (\text{A.2})$$

$$y = d \cos \lambda \cos \eta \quad (\text{A.3})$$

$$z = d \cos \lambda \sin \eta, \quad (\text{A.4})$$

where d is the distance of the galaxy and η and λ are the SDSS angular coordinates. We then defined a $3 h^{-1}$ Mpc grid in these coordinates for calculating the LDF.

The luminosity-density value l_i at grid point r_i is calculated by a kernel sum

$$l_i = \frac{1}{a^3} \sum_{\text{gal}} K^{(3)}\left(\frac{r_{\text{gal}} - r_i}{a}\right) L_{\text{gal,w}}, \quad (\text{A.5})$$

where a is the kernel scale. As the kernel we used the B_3 spline function

$$B_3(x) = \frac{|x-2|^3 - 4|x-1|^3 + 6|x|^3 - 4|x+1|^3 + |x+2|^3}{12}, \quad (\text{A.6})$$

with $8 h^{-1}$ Mpc (BOSS) or $16 h^{-1}$ Mpc (LRG) smoothing scale. As the final step in constructing the LDF, we normalised the density field by converting the densities into units of mean density as

$$D_i = \frac{l_i}{l_{\text{mean}}}, \quad (\text{A.7})$$

where l_{mean} is the average over all density grid points.