THE SPATIAL CLUSTERING OF LOW LUMINOSITY AGN

ANCA CONSTANTIN AND MICHAEL S. VOGLEY
Department of Physics, Drexel University, Philadelphia, PA 19104
Submitted to ApJ

ABSTRACT

We present the first multi-parameter analysis of the spatial clustering properties of narrow line emitting galaxy nuclei. We spectroscopically identify and classify active galactic nuclei (AGN) from the Sloan Digital Sky Survey (SDSS) DR2 main galaxy sample, which yields the most precise measurement to date of AGN clustering. Estimates of the two-point correlation function (CF) reveal that Seyferts are clearly less clustered than normal galaxies, while the clustering amplitude of LINERs is consistent with that of the parent galaxy population. The similarities of the distributions of host properties (color and concentration index) of Seyferts and LINERs suggest that the difference in their clustering amplitudes is not driven by the morphology-density relation. We find that the luminosity of [O I] emission shows the strongest influence on AGN clustering, with low $L_{[O I]}$ sources having the highest clustering amplitude, $r_0$. This trend is much stronger than the previously detected dependence on $L_{[O III]}$, which we confirm. There is a strong correspondence between the clustering patterns of objects of given spectral type and physical properties including $L_{[O I]}$, $L_{[O III]}$, the electron density of the emitting gas $n_e$, and the obscuration level. LINERs, which exhibit high $r_0$, show the lowest luminosities and obscuration levels, and relatively low $n_e$, suggesting that these objects harbor black holes that are relatively massive yet weakly active or inefficient in their accretion, probably due to the insufficiency of their fuel supply. Seyferts, which have low $r_0$, are very luminous in both [O I] and [O III] and show large $n_e$, suggesting that in these systems the black holes are less massive but accrete quickly and efficiently enough to clearly dominate the ionization. Star-forming galaxies – the H II’s – are weakly clustered; because these nuclei are hosted mainly by blue, late type hosts, this trend can be understood as a consequence of both the morphology-density and star formation rate-density relations. However, the spectral properties of the H II’s suggest that these systems hide in their centers, amidst large amounts of obscuring material, black holes of generally low mass whose activity remains relatively feeble. Our own Milky Way may be a typical such case.

Subject headings: galaxies:active—galaxies:emission lines—galaxies:clustering—galaxies:statistics

1. INTRODUCTION

Clustering analysis can be a useful tool to employ in investigating the formation and evolution mechanisms of various astrophysical objects. In the case of AGN, the pattern of spatial clustering and its implications for the environments of AGN provides constraints on galaxy evolution models, and in particular on how well AGN trace the normal, non-active galaxy distribution. Comparisons of the AGN environments with those of control samples of regular galaxies may indicate differences that reveal what determines the activity in galaxy nuclei, and even gauge its strength and time cycle, thus yielding a phenomenological link between black hole growth and galaxy formation and evolution processes.

Estimates of the two-point correlation function (CF) provide important diagnostics for the black hole (BH) masses and duty cycle of fueling in AGN. If structure formation is hierarchical, then the rarer and most massive virialized dark matter (DM) halos cluster the most strongly [e.g., Kaiser 1984], thus the clustering strength of a given object type constrains the mass and number density of the halos in which they reside. The DM halo mass can further be connected to the physical properties of the resident galaxies and their central black holes through empirical scaling relations: the black hole mass – stellar velocity dispersion $M_{BH} - \sigma_*$ relation that holds for normal galaxies [Gebhardt et al. 2000; Ferrarese & Merritt 2000] as well as for the active ones [Gebhardt et al. 2000; Ferrarese et al. 2001; Nelson et al. 2004; Onken et al. 2004; Greene & Ho 2005], the correlations between bulge properties and $M_{BH}$ [e.g., Magorrian et al. 1998; Graham et al. 2001], and the correlation between circular velocity and central velocity dispersion [Ferrarese 2002; Baes et al. 2003] that links dark matter halos and supermassive BHs. The number density of the DM halos can be used to infer the duty-cycle, or the accretion life-time, as this is given by the inverse ratio of this number density to that of the observed objects [Martini & Weinberg 2001]. Such predictions for the AGN’s lifetime, and its implications for the efficiency of BH formation [Haiman & Hui 2001], can test the fundamental assumptions of semi-analytical models of AGN and galaxy formation [Kauffmann & Haehnelt 2000; Di Matteo et al. 2003].

Studying the spatial clustering properties of different spectroscopically-classified AGN species may link their ionization mechanisms to properties of their hosts and environment. Through comparison of the clustering amplitudes of accreting-type sources such as Seyferts and the starlight-ionized H II galaxies, it is possible to contrast properties of the dark matter halos that host different kinds and levels of galactic nuclear activity. Particularly interesting is the clustering of the ubiquitous Low Ionization Emission-Line Regions (LINERs, e.g. Heckman 1980), whose physical origin, whether thermal...
or non-thermal, remains ambiguous despite exhaustive analysis of their properties (Ho 1999; Ho et al. 2003; Filho et al. 2004; Maoz et al. 2003; Constantin et al. 2006).

Galaxy mergers are predicted to play a crucial role in triggering AGN activity (e.g., Barnes & Hernquist 1992; Di Matteo, Springel & Hernquist 2005). If galaxy nuclei become more AGN-like when major mergers occur, then active galaxies of different spectral signatures, corresponding to different degrees of contributions from a non-thermal ionization process, may cluster differently, reflecting variation in their environments. Thus, a comparison of the clustering properties of the active and non-active galaxies is an important tool for testing the merger-accretion connection.

Through its wealth of both spectroscopic and photometric data, SDSS (York et al. 2000) is ideal for investigating the dependence of clustering properties on a variety of emission characteristics. The SDSS observations offer the unique opportunity to examine simultaneously the dependence of environment on the spectral classification of emission-line galaxy nuclei, and on various other features that characterize or even determine a detectable accretion process, in conjunction with properties of their host galaxies. Previous studies with similar aims (Kauffmann et al. 2003, 2004; Miller et al. 2003; Wake et al. 2004) compared passive galaxies with AGN as a general class and did not investigate the properties of sub-populations such as Seyferts and LINERs separately. In these studies, the optical classification of an object as an AGN relies, especially among the narrow-line emitters, mainly on the [N II]/Hα vs. [O III]/Hβ diagnostic diagram. There are suspicions that such rather lenient definitions may not best represent the accretion-type population (Hao et al. 2007). Using such definitions, the AGN population is found by Miller et al. (2003) and Wake et al. (2004) to occupy a uniform fraction of galaxies, to be common to a large range of galactic environments, and to follow the distribution of the whole galaxy population, and therefore to be unbiased tracers of mass in the universe. However, when divided by their luminosities, the active nuclei show markedly different clustering properties (Wake et al. 2004), different star-formation histories in their hosts (Kauffmann et al. 2003), and different environmental preference (Kauffmann et al. 2004). Such findings are intriguing and thus merit further clarification. In particular, it is crucial to establish and quantify the connections between the intrinsic physical properties of the nuclear line-emitting regions and those of their host galaxies and dark matter halos. We pursue these issues here by including important spectral diagnostics that probe the astrophysics of AGN.

In this paper, we use SDSS DR2 data in an attempt to detect and analyse differences in the clustering properties of (1) active and non-active galaxies, (2) nuclei of distinct types of activity, i.e., accretion versus star-formation, and (3) different spectroscopic types of AGN, e.g., Seyferts versus LINERs. In order to carry out such analysis with a good understanding of the possible biases introduced by detection and spectral identification, we present in Section 2 a systematic examination of the different AGN selection criteria and their sensitivity to the dominance of an accreting-type nuclear power source. In Section 3 we estimate and compare the redshift-space CF for subsamples of objects of different spectral classes (Section 3.1.1), and investigate the degree to which these measurements are influenced by the level of nuclear activity, as traced by the [O III] and [O I] luminosities (Section 3.2), properties of the line-emitting gas, e.g., intrinsic obscuration, as traced by the neutral hydrogen column densities $N_{HI}$ and the electron density $n_e$ (Section 3.1.3), and host properties (Section 3.2). We discuss in Section 4 possible implications of our findings on the features that distinguish Seyferts from LINERs, on the peculiar nature of the Transition objects, that border the definition of AGN (Seyferts or LINERs) and H II, and on the potentially great similarity between the Galactic Center and a typical H II system. Section 5 summarizes the findings and conclusions of this work.

2. THE DATA

This study is based on data from the Sloan Digital Sky Survey (SDSS) Data Release 2 (DR2; Abazajian et al. 2003). The parent sample from which we draw various categories of objects is a subset of the main galaxy sample used for large-scale structure studies (Strauss et al. 2002), and includes objects with Petrosian r magnitudes $14.5 < m_r < 17.77$ after correction for Galactic extinction based on maps of Schlegel, Finkbeiner & Davis (1998), and with a redshift distribution that extends from $0.005$ to $0.30$, with a median of $z \approx 0.1$. Thus, the sample does not include low redshift quasars or broad emission-line AGN (Richards et al. 2002; Schneider et al. 2002, 2003).

The core of the analysis we present here exploits the source spectroscopic properties and, consequently, the source spectral type definition. The spectroscopic measurements we employ are drawn from a catalog of absorption and emission line fluxes and equivalent widths (EW) of a superset of the DR2 galaxy targets which fulfill the galaxy sample selection criteria. These data are kindly made publicly available by the MPA/JHU collaboration at http://www.mpa-garching.mpg.de/SDSS/Brinchmann et al. (2004), and are analyzed by, e.g., Kauffmann et al. (2003, 2004). This dataset has the advantage that the line emission component is separated from the host galaxy light; based on stellar population synthesis templates that span a relatively large range in age and metallicity, the absorption-line contribution is identified and subtracted from the total galaxy spectrum (Tremonti et al. 2004).

Other observed properties of the objects used in this study were measured using the standard SDSS photometric and spectroscopic pipelines and were obtained directly from the SDSS archive. We employ Petrosian magnitudes, and in calculating the absolute magnitudes from the apparent magnitudes $m_r$ and the redshift $z$ we apply the formula

$$M_r = m_r - 5 \log_{10}(r(1+z)/1\text{Mpc}) - 25 - K(z) + 5 \log_{10} h + Q(z).$$

Here $K(z)$ is the r-band magnitude galaxy K-correction relative to the $z = 0.1$ value, as calculated by Blanton et al. (2003a), and made publicly available as part of the NYU Value-Added Catalog (Blanton et al. 2005, http://wassup.physics.nyu.edu/vagc/), $r(1+$
z) is the luminosity distance, with r the comoving distance, and $Q(z)$ accounts for the average evolution in galaxy luminosities in the past, with $Q(z) = 1.6(z - 0.1)$ (Blanton et al. 2003; Tegmark et al. 2004). Throughout this paper we assume a cosmology with $\Omega = 0.3$, $\Lambda = 0.7$, and $H_0 = 100 h$ km s$^{-1}$ Mpc$^{-1}$. The distances are quoted in $h^{-1}$ Mpc.

In this study we conduct a comparative analysis of the clustering properties of the emission-line galaxies by employing subsamples of galaxies, defined by absolute magnitude limits. Figure 1 illustrates the distribution of the whole galaxy sample (247,080 objects) in the redshift–absolute-magnitude space. The continuous curves delineate the sample boundaries corresponding to our choice of apparent magnitude limits $14.5 < m_r < 17.7$. The bright cutoff excludes bright galaxies that may have been shedded into many pieces, particularly in early SDSS photometry, and/or galaxies that have been excluded from spectroscopic target selection due to fiber saturation effects, while the faint limit corresponds to that of the main galaxy sample (Strauss et al. 2002).

Following Park et al. (2003), we define a “Best” volume-limited sample of SDSS galaxies that contains the maximum number of galaxies, with absolute-magnitude limits $-21.6 < M_r < -20.2$ that correspond to a redshift range of $0.05 < z < 0.12$, or a comoving distance range of $148.5 < r < 350.1 h^{-1}$ Mpc, when the apparent magnitude cuts are applied. The definition of our “Best” volume is shown in Figure 1 as a dashed line. The sky area covered by the parent spectroscopic sample is illustrated in Figure 1 of Abazajian et al. (2004). It is important to note that the subsets of spectroscopically-defined AGN drawn from this volume-limited galaxy sample will not themselves be truly volume-limited samples because of selection effects induced by their spectroscopic definition; the host galaxies are restricted to the described range of absolute magnitude and redshift, but the redshift distributions of AGN may not be uniform because of the requirement that the emission lines be detected. These effects are discussed in detail in Section 2.2 and are taken into account in our clustering analysis.

The spectroscopic subsamples of interest for this study suffer from the fiber collision problem. When two targets lie closer than 55", only one member of such galaxy pair will have a spectrum available, unless later spectroscopic tiles overlap that same area of sky. Because classification as an AGN requires a spectrum for each object, we cannot make any “correction” for such collisions, as is commonly done for large-scale structure studies of SDSS galaxy samples (Zehavi et al. 2002). The impact of this problem on our estimates of the correlation functions is discussed later in Section 3.

2.1. Subsample Definition

2.1.1. Active vs. Non-Active: the Spectral Classification

The fraction of galaxies with an AGN is difficult to assess observationally, because it strongly depends on the quality and quantity of the available spectral data. Sub-classifying emission-line nuclei typically requires a comparison of pairs of line-flux ratios of spectral features that are relatively insensitive to reddening and metallicity, are able to distinguish between different ionization levels, and provide a useful segregation of different object types. There is a set of line ratios that satisfy these criteria, and that have been widely used in quantifying emission properties and distinguishing between accretion-powered and starlight-powered nuclei: [O III]/Hβ, [N II]/λ6583/Hα, [S II]/λ6716,6731/Hα, and [O I]/λ6300/Hα (e.g., Baldwin, Phillips, & Terlevich 1981; Veilleux & Osterbrock 1987).

Figure 2 illustrates how galaxies that exhibit all 6 of these emission lines at moderately high signal-to-noise, e.g., with fractional errors $\leqslant 50\%$, segregate as a function of these combinations of line ratios. Following the Ho et al. (1997a) criteria, as listed in Table I we classify the objects as H II galaxies, Seyferts, and LINERs, with their Pure LINER and Transition subsets. H II’s are the galaxy nuclei that populate mainly the left branches of these diagrams, as they usually show weak low-ionization transitions in [N II], [S II], and especially [O I]. By contrast, Seyferts and LINERs occupy the right branch as they exhibit relatively strong forbidden features. The distinction between Seyferts and LINERs is usually given by their ionization level, which is easily gauged by the [O III]/Hβ ratio, with Seyferts showing larger values. The Transition systems are differentiated from the Pure LINERs based on their low [O I]/Hα ratio. Such divisions offer a basis for categorizing the line-emitters. However, as is apparent from these three heavily populated diagrams, the nebular properties span a continuous range; no breaks appear in the diagrams to indicate a clear-cut transition between one type to another.

There are also relatively successful theoretical attempts to separate the accretion-type galaxy nuclei from star-forming sources, using criteria that can be described using these same spectra diagnostics diagrams. Based on models that account for large realistic metallicity, ionization parameter, and dust depletion ranges, Kewley et al. (2001) find that simple hyperbolic curves can be used to split the line-emitting galaxies into AGN and star-forming galaxies. As can be seen from Figure 2, the curves claimed as upper limits on the location of H II-like objects exclude only a small fraction of LINERs from the AGN class, and are otherwise reasonably consistent with the Ho et al. (1997a) classification criteria.

Previous studies of low luminosity narrow-lined AGN in SDSS samples used simpler criteria than we use here. Based on the shape of the distribution of objects in the [N II]/Hα vs. [O III]/Hβ diagram, Kauffmann et al. (2003) propose an empirical cut between AGN and H II-type galaxies, while some other studies (Miller et al. 2003; Wake et al. 2004) classify an object as an AGN based only on values of the [N II]/Hα flux ratio (i.e., based on the condition that [N II]/Hα > 0.6). These classifications overlook the emission-line behavior as a function of [S II]/Hα, and [O I]/Hα, which may discriminate better than [N II]/Hα between the young stars and AGN as dominant ionization sources. Of the three low-ionization lines, [O I] is the most sensitive to the shape of the ionizing spectrum, requiring a significantly hard radiation field, i.e., that of an AGN, to sustain a sufficiently extensive partially ionized zone in clouds optically thick to Lyman continuum, and thus to produce a strong such feature. On the other hand, high [N II]/Hα line ratios in AGN may be caused by potential selective enhancements of nitrogen (Storchi-Bergmann & Pastoriza 1989).
and thus are not always caused by an accretion-type radiation field.

Figure 3 illustrates the behavior as a function of [S II]/Hα and [O I]/Hα of the objects classified as AGN based on [N II]/Hα ratio like in Miller et al. (right panels), and based on the Kauffmann et al. line that separates the two branches of the objects distribution in [N II]/Hα vs. [O III]/Hβ diagram (left panels). For the Miller et al. classification, the density contours suggest that these AGN are governed by a LINER-like behavior.

Using the spectral types defined in Table 1, the fraction of LINERs is at least 30%, while Seyferts make up for less than 9% of these AGN, in both their ‘2-line’ AGN (only detection of 2 line-emissions, [N II] and Hα is required) and “4-line” AGN (additional detection of [O III] and Hβ emission lines is required) samples. Also, the fraction of Miller et al. AGN with low [S II]/Hα and [O I]/Hα ratios, which are more likely to be dominated by star-forming activity, is as high as 20%, which is more than double the number of Seyferts. AGN defined using the Kauffmann et al. criteria severely leak into the H II locus in both [S II]/Hα and [O I]/Hα diagrams. 40% of these sources, or 60% of the objects situated between the Kewley et al. and Kauffmann et al. lines (the blue dots in the [N II]/Hα vs. [O III]/Hβ diagram), would be classified as H II using the criteria in Table 1. Seyferts remain poorly represented in the Kauffmann et al. sample as well (less than 5%), while LINERs, although constituting the majority of the potentially accretion-dominated sources (25%), are greatly outnumbered by H II’s.

This comparison shows that the Kauffmann et al. and the Miller et al. AGN classifications, although powerful through their great statistics, may not best represent accretion dominated narrow-line galaxy nuclei. Such classifications include a significant fraction of galaxies that show star-forming activity, and certainly, a mix of different types of AGN. For the purpose of understanding differences between emission-line activity originating in star formation and black hole accretion, a more stringent characterization of the dominant power source, particularly as a function of [S II] and/or [O I], is necessary.

2.1.2. Our Samples

We select from the parent sample of galaxies a subset of strong “emission-line” sources and a set of “passive” objects that show insignificant, if any, line emission activity. The emission-line sources are galaxies that exhibit all six lines Hα, Hβ, [O III],[N II] [S II], and [O I] in emission, i.e. the line fluxes have positive values and the fluxes are measured with at least 2σ confidence. We identify 54,410 such strong line-emitters, which comprise ≈ 20% of all objects. Note that our definition of emission-line objects is very restrictive, and therefore the detection fraction of such sources among all galaxies is smaller than what Ho et al. (1997a) and Miller et al. 2003 find (≥ 40%). When we relax the “6-line” definition criteria to include any kind of emission-line activity, we recover a similarly high fraction. The “passive” galaxies are defined here as the objects for which Hα, Hβ, [O II], and [N II] are not detected in emission (i.e., their EWs are positive in the emission-line catalog, indicating an absorption feature). This sample, which includes 59,197 sources, is a subset of the red, old, non-Hα emitters.

Among the emission-line galaxies described above, we use the criteria listed in Table 1 to identify 34,781 H II-type galaxies, 2,392 Seyferts, and 7,468 LINERs, among which 2,155 are Transition objects. There are a significant number of strong emission-line objects (≈18%) that do not simultaneously comply with these sets of conditions: these are objects that show large [N II]/Hα ratios (> 0.6) but low [S II]/Hα (< 0.4) and/or [O I]/Hα (< 0.08), or small [N II]/Hα and large ratios involving [S II] and/or [O I]. For the sake of simplicity, we exclude these sources from the comparative analysis of clustering of Seyfert, LINER and HII populations.

Although [O I] is generally weak and therefore difficult to measure accurately, this feature is very useful in separating the Transition objects among LINERs. Note, however, that requiring its detection in the sample definition severely diminishes the sample sizes. Without this condition on [O I] and consequently on [O I]/Hα, we identify 98,556 line emitters, or ≈ 40% of the galaxies, among which 3,052 are Seyferts, 12,779 are LINERs, and 65,971 are H II galaxies. In our clustering analysis, we compare the clustering properties of active galaxies whose definitions both include and exclude conditions on [O I] and [O I]/Hα.

2.2. Subsample Properties & Selection Effects

We begin by forming the “Best” absolute-magnitude limited sample of galaxies using the definition illustrated in Figure 1, then apply the criteria listed in Table 1 to form spectroscopically-classified subsamples. The numbers of objects in each subsample are recorded in Table 2. Figure 4 shows the redshift z distributions for these samples and comparisons of their apparent and absolute r magnitudes are presented in Figure 5.

It is clear that, although drawn from a parent sample of galaxies with a fixed range of intrinsic luminosities, subsamples of spectroscopically-classified objects are not, in general, volume limited. The necessity of detecting the spectral features employed in classification causes a redshift dependence. For example, LINERs and Seyferts are less prominent than other types of sources at z > 0.1, and seem to populate the brighter galaxies, both in apparent and absolute magnitudes, while the H II systems are shifted toward fainter apparent and absolute magnitudes.

Whether a galaxy nucleus is included into a given spectral category is strongly subject to multiple selection effects that vary with redshift. The fraction of galaxy centers exhibiting measurable emission lines decreases with redshift as a consequence of decreasing signal-to-noise. More importantly, the increase with distance of the projected size of the SDSS spectroscopic aperture causes dilution of emission-line activity by the increasing amount of host galaxy light. Thus, weak line-emitting systems are more likely to be misclassified as non-active at larger distance. These effects explain the differences seen in the redshift and magnitude distributions of the accretion-type sources (Seyferts, and at least some LINERs) and the galaxy nuclei that are mainly ionized by hot young stars (H II). While in the latter objects the power source is more extended, and thus more likely to be observed at larger distances, in the accretion dominated nuclei the ionization originates in more centrally concentrated regions and thus is less likely to be included in the emission-line sample at large z where the signal-to-noise
is lower.

We find that the different spectral classes of objects exhibit distinct behavior in several physical characteristics. Figure 6 shows the distributions in log $L_{\text{OIII}}$ and log $L_{\text{OIII}}$, the intrinsic column density $N_H$, and the electron density $n_e$, separately for the Seyferts, LINERs, H II galaxy nuclei, and the whole sample of 6-line emitting galaxies. Line luminosities $L_{\text{OIII}}$ and $L_{\text{OII}}$ are calculated only for objects that show flux measurement uncertainties of $< 50\%$ in these lines and in the Balmer decrement $H\alpha/H\beta$, which is used to correct for intrinsic extinction. We assume a dust-free case-B recombination value of 2.86 for the standard, not absorbed $H\alpha/H\beta$, and, for the sake of simplicity and to ease comparison with previous calculations by [Kauffmann et al. 2003], an attenuation law of the form $\tau \propto \lambda^{-0.7}$ [Charlot & Fall 2000]. We use the same method for calculating the $N_H$ values. To estimate $n_e$ we employ the $[\text{S II}]\lambda 6716/\lambda 6731$ line-flux ratio [e.g., Osterbrock 1989]. The $[\text{O II}]\lambda 3729,3727$ feature could also be used as a density indicator, as it is available in the majority of the SDSS galaxy spectra, and the emission-line catalog that we use here includes measurements of the individual components. However, the small separation of only 3Å between the two lines may cause large uncertainties in their fluxes, and therefore in the $n_e$ estimates as well.

Investigation of the $L_{\text{OIII}}$-dependent behavior of different types of galaxies is important because this parameter is claimed to trace the AGN activity, in the sense that highly $[\text{O III}]$ luminous objects are more likely to be powered by an accretion-type nucleus, and that larger $L_{\text{OIII}}$ may correspond to larger accretion rate as well [Kauffmann et al. 2003]. The recent findings of Constantin et al. [2006] suggest that we also closely examine the role played by $L_{\text{OII}}$ in distinguishing among different spectral types and in their spatial clustering properties. Constantin et al. [2006] indicate that the presence and strength of the $[\text{O I}]$ emission may be linked to the presence and strength of the broad $H\alpha$ emission, and thus, to the dominance of the accretion-type ionization in the nuclear spectra of nearby emission-line galaxies. Note that LINERs stand out as the objects with the lowest $L_{\text{OIII}}$ and $L_{\text{OII}}$ relative to other spectral classes, rarely exhibiting values of log $L_{\text{OIII}}> 40$, with $L_{\text{OIII}}$ in units of erg s$^{-1}$. Seyferts, on the other hand, are the galaxies that are most likely to be luminous in these emission lines; the median values for log $L_{\text{OIII}}$ are $\sim 38.6$ for LINERs, and $\sim 39.8$ for Seyferts. Although the lack of LINERs with high $L_{\text{OIII}}$ can be explained by their definition as low ionization emission regions, which explicitly selects lower $[\text{O III}]$ fluxes, the absence of high $L_{\text{OIII}}$ LINERs is intriguing. Because the redshift and magnitude distributions of Seyferts and LINERs are very similar, selection effects cannot be attributed to this trend, thus their low $L_{\text{OIII}}$ and low $L_{\text{OII}}$ must be intrinsic to these systems.

Dust obscuration is another variable that could significantly affect the detection of AGN-like activity. In terms of intrinsic attenuation, as measured by $N_H$, LINERs are again at one extreme, showing low levels of extinction relative to Seyferts and H II nuclei, with the latter being the most obscured of the emission-line objects. These trends are to be expected, as star formation is generally accompanied by significant amount of obscuring dust. However, it may also be the case that in such a dusty star-forming environment, a potentially present weak AGN-like nuclear emission is suppressed, because it lies buried in surrounding H II-like features.

Based on the $[\text{S II}]$ line flux ratio estimates for the electron density, Seyferts show the highest $n_e$, as expected for sources in which the dominant ionization mechanism is accretion of matter onto a black hole. The LINERs show lower $n_e$ densities, while the starforming systems clearly populate the lowest end in the $n_e$ distribution; the median $n_e$ is $\approx 5 \times 10^2$ cm$^{-3}$, $\approx 1.5 \times 10^2$ cm$^{-3}$, and $\approx 5 \times 10^3$ cm$^{-3}$ for the Seyferts, LINERs and H II nuclei respectively. A Kolmogorov-Smirnov test indicates an almost null probability that any pair of these subsamples draw their distributions in $[\text{S II}]$ ratio from the same parent distribution; the KS probability that Seyferts and the whole sample of emission-line emitters belong to the same parent population is $K_{\text{Seyffers emiss}} = 4.24 \times 10^{-36}$, while for the comparison between the line emitters and the H II’s, $K_{\text{emiss/HII}} = 2.42 \times 10^{-23}$. There are roughly 3000 galaxy nuclei that manifest emission-line activity with $n_e < 10$ cm$^{-3}$, which is below the theoretical lower limit for which $n_e$ can be estimated based on $[\text{S II}]$ emission; about half of these objects are H II galaxies.

Median and average values of the parameters for which the distributions are compared among different spectral types of objects are recorded in Table 4.

3. THE CORRELATION FUNCTION

To quantify clustering in the observed galaxy distribution we estimate the 2-point correlation function (CF) $\xi(r)$, which measures the excess probability over random to find an object pair separated by comoving distance $r$. The separation $r$ is measured in redshift-space, thus it is important to note that the shape and amplitude of this measure of spatial clustering are determined not only by the spatial configuration but also by peculiar velocities. Small scale velocity dispersion, e.g., in cores of galaxy clusters, damps small scale clustering, while linear infall onto dense regions amplifies clustering on large scales.

Accurate estimation of the correlation function requires careful accounting for both angular and radial variation in the expected galaxy density. The former requires knowledge of the complicated geometry of the angular mask that describes the regions on the sky in which data were obtained and the sampling of galaxies in the observed regions, while the latter variation may depend on physical characteristics of the samples analysed, which can introduce important selection biases with distance.

To account for the survey geometry we employ random catalogs that describe the SDSS DR2 survey angular selection function. These random masks are generously made publicly available in the NYU Value-Added Galaxy Catalog and described in [Blanton et al. 2004]. The random samples we employ in all the $\xi(r)$ measurements are at least 10 times as large as the datasets involved. As discussed in Section 2.1.1, the $z$ distributions for the volume limited samples of the different categories of objects are not uniform, and are different from sample to sample. We therefore build individual radial selection functions for the random masks corresponding to each sample to be studied; we randomly assign redshifts that
follow distributions constructed based on 2nd order polynomial fits to the observed object z distributions. We have explicitly tested the CF $\xi(r)$ estimates for different parametrizations of the redshift distributions $N(z)$, and find that they do not significantly affect the results; the power-law fits, described below, show slopes and amplitudes that remain consistent within the value range bordered by uncertainties.

As discussed above, we do not account for the potential biases on small scales introduced by the fiber collisions, because we have no reasonable means of doing so. In other estimates of galaxy clustering properties, such biases have been corrected for by assigning each pair member whose redshift was not obtained because of the fiber collision the same z as the pair member whose redshift was measured. However, for the purpose of studying the dependence of clustering on spectral characteristics we cannot apply such correction because there is no physical motivation for assigning to the unobserved pair member the spectral properties found to characterize the object whose spectrum is available. Therefore we treat the fiber collisions as a contribution to incompleteness in the survey; this variation of the completeness with angular position of the is accounted for in the random catalogs described above. This procedure leaves a residual effect due to fiber collisions that is significant at very small scale, i.e., at comoving transverse separations of $\sim 0.14h^{-1}$Mpc (the corresponding linear scale for 55″, at the outer edge of our samples $cz = 50,100$ km/s^-1). We restrict our measurements to separations > $0.14h^{-1}$Mpc and note that the incompleteness biases due to fiber collisions should similarly affect all of the subsamples we analyse, and thus are not a significant concern for this study because our focus is on the relative differences between the CF characteristics.

We calculate the correlation functions using the estimator,  
$$\xi(r) = 1 + \left(\frac{N_{rd}}{N}\right)^2 \frac{DD(r)}{RR(r)} - 2 \frac{N_{rd} DR(r)}{N RR(r)}, $$

using bins in pair separation r that are logarithmically spaced with a width of 0.33 in log(r/h^-1 Mpc), starting from 0.147h^-1 Mpc. In this formula, $DD(r)$ and $RR(r)$ are the number of pairs in the data and random catalogs respectively, while $DR(r)$ is the number of pairs between the data and the random samples. Pair counts are estimated using the computationally efficient “npt” algorithm described in [Moore et al. (2006)]. We estimate the covariance matrices of the uncertainties using the statistical jackknife method [e.g., Lupton (1993)] for which we divide our samples into 75 separate regions on the sky, of approximately equal area, $\leq 10$ deg $\times 5$ deg, and estimate the CF 75 times, each time leaving out one of these regions (see [Zehavi et al. (2002)] for a more detailed description of this method and for a discussion on its robustness in estimating the uncertainties).

We quantify the clustering properties of a given sample by fitting the measured CF with a power law of the form  
$$\xi(r) = (r/r_0)^{-\gamma},$$

where r denotes the comoving separation between the object pairs, $r_0$ is the correlation length used to express the amplitude of clustering, and the slope $\gamma$ quantifies the ratio of small to large scale clustering. In the fitting process we employ separations $3h^{-1}$ Mpc $< r < 10h^{-1}$ Mpc, in order to avoid separations on which the CFs are derived from fewer than 20 data-data pairs.

The CF of the full volume limited “Best” galaxy sample is characterized by a correlation amplitude $r_0 = 7.8h^{-1}$ Mpc, and a slope $\gamma \approx 1.2$. These values are consistent with previous calculations of the redshift-space clustering of the SDSS main galaxy sample, both for magnitude and volume limited samples similarly bounded in luminosity [Zehavi et al. (2002, 2003)].

3.1. Clustering Dependence on Spectral Properties

We examine here the redshift-space CF for various subsamples constructed based on sets of properties that reflect differences in the dominant ionization mechanism, i.e., accretion vs. star-forming type, or provide clues to the physical conditions of the line-emitting regions. We compare correlation functions of different spectral classes, and investigate the degree to which other emission-line related parameters, e.g., [O III] and [O I] luminosities, intrinsic neutral hydrogen column densities ($N_H$), and electron densities $n_e$, may influence the results.

The results of power-law fits to the CFs estimated for various subsamples of the parent “Best” magnitude limited galaxy sample considered in this analysis are recorded in Table 2 together with results for the CF of all galaxies in the sample. Object sample definitions as a function of $n_e$ and $N_H$ are described in Tables 3 and 4. It is clear that the clustering properties of subsamples of objects governed by distinct spectroscopic characteristics are different. In the following subsections we discuss these results in detail.

3.1.1. Spectral Class

The CFs estimated for emission-line objects, passive objects, and the three distinct line-emitting subsets of Seyferts, LINERs and H II nuclei, are illustrated in the upper and lower panels of Figure 7. We present and compare the shape and amplitude of power-law fits to the CFs for samples whose definition both include and exclude constraints on [O I] and [O I]/Hα, in the right and left panels respectively. The CF of the volume limited sample of all galaxies is also shown for comparison in all panels. The error bars denote the 1-sigma uncertainties. The upper panels include insets that compare the 1-sigma, 90%, and 99% confidence contour levels for the likelihood functions obtained by fitting power-law models to the correlation functions of actively line-emitting galaxies and the passive ones.

The passive galaxies have a substantially higher amplitude and steeper $\xi(r)$ than the emission-line galaxies, with a correlation length $r_0 \approx 11.4h^{-1}$ Mpc, compared to $r_0 \approx 5.6h^{-1}$ Mpc for the actively line-emitting sources. This trend is clearly present whether or not the selection of these objects employ the [O I] feature, although the datasets are smaller when [O I] is required in defining the sample and, therefore, the uncertainties are correspondingly larger.

Among the emission-line sources, LINERs exhibit the highest clustering amplitude, while H II’s show the weakest clustering, at both small and large scales (see lower panels of Figure 7). Seyferts distinguish themselves from
the other emission-line galaxies by the relatively steep slope of their correlation function, \( \gamma > 1.4 \). This effect might arise if the environment of Seyferts causes an excess (relative to a flatter power law) of AGN activity on a scale larger than clusters or groups; the power-law fits are over separations 3 – 10h\(^{-1}\)Mpc, so the excess small-scale clustering is not caused by very close pairs of Seyferts. Overall, Seyferts’ clustering amplitude remains intermediate between that of LINERs and HI’s.

A more quantitative comparison between the clustering properties of Seyferts, LINERs, HI’s, and of the whole galaxy sample is presented in Figure 8. The \( \chi^2 \) confidence contours indicate that LINERs’ clustering strength and properties are very similar to those of the galaxy sample as a whole, while Seyferts and star-forming galaxies are substantially less clustered.

That Seyferts are found to be less clustered than the average galaxies is an important result, and its statistical significance require some comment. Table 2 shows that the clustering amplitudes for Seyferts and for all galaxies in the main sample are \( r_0 = 6.00 \pm 0.64 \) and \( 7.80 \pm 0.49 \) respectively, which translates into a difference of order 3\( \sigma \). In the upper panel of Figure 8 one sees that the 2-dimensional “1-\( \sigma \)” limits just touch, and the 90\% contours overlap. This overlap arises because the 2-dimensional contours enclose the labeled percentage of the probability in the joint distribution of both \( r_0 \) and \( \gamma \), and these elliptical contours extend beyond the confidence limits that are computed in one parameter at a time (by marginalizing over the other parameter). The natural confidence regions in the 1-dimensional parameter space, say \( r_0 \), are the projections of the 2-dimensional regions defined by fixed \( \Delta \chi^2 \) into this 1-dimensional space of interest. To enclose the same probability, i.e., fraction of points, say 68.3\% (or \( \Delta \chi^2 = 1 \)), as the two parameters \( r_0 \) and \( \gamma \), the corresponding ellipse will correspond to \( \Delta \chi^2 = 2.3 \) as it must necessarily extend outside of both of them.\(^1\) The resulting constraints on \( r_0 \) and \( \gamma \) are reported in Table 2.

We find a slight tendency toward steeper CFs in the sub-samples, as compared to the CF of “all galaxies.” Note also that the correlation functions of the sub-samples do not necessarily average to the correlation function of the whole sample; the CF of the whole sample is a weighted sum over the correlations of the sub-samples and the cross-correlations between the sub-samples. It could be the case that the CFs of the sub-samples are steeper because certain environments favor the formation of particular species of AGN. In an extreme example, if Seyferts, LINERs, and HI galaxies came in clumps of one type only, they would be strongly clustered on small scales, still clustered on larger scales, but have a relatively steeper correlation function because of the extra small-scale clustering. In contrast, the cross-correlations between sub-samples, e.g. between Seyferts and HI galaxies, would have very little small scale clustering because they would be mutually exclusive on the scale of the clumps, thus the cross-correlation functions would be shallower than that of the whole sample.

In order to examine our results in the context of previous work on clustering of AGN observed by the SDSS, we also examine the CF of the narrow-line AGN vs. non-AGN (i.e., star-forming sources) using the classifications employed by Kauffmann et al., and Miller et al. Figure 9 shows the CFs estimated for such subsamples of the “Best” volume limited galaxy sample and comparisons of the confidence contours of power-law fits to these CFs. The galaxy CF is also shown for comparison. The fitted power-law parameters and number of objects in each sample are also listed in Table 2. Here we find that AGN samples selected using the criteria applied by Kauffmann et al. and Miller et al. show very similar clustering amplitudes to that of all galaxies, as claimed in previous studies based on estimates of population fractions in cores of galaxy clusters and intra-cluster fields (Miller et al. 2003), or measurements of CFs (Wake et al. 2004). This result is not surprising. Above in Section 2.1 we find that AGN samples selected using these criteria are dominated by LINERs. We also find that LINERs, as a spectroscopically-distinct population, are indeed clustered similarly to normal galaxies. Thus, the resemblance between the spatial clustering properties of such AGN samples and normal galaxies reflects dominance of LINERs in these AGN definitions, and it does not apply to AGN in general, particularly not to Seyferts as a separate class.

The results presented in this Section indicate clearly that bona-fide narrow-line low luminosity AGN, i.e., the Seyfert 2s, are less spatially clustered than typical galaxies. These findings have important implications for the environment of the host galaxies of AGN. The statistics of peaks in a Gaussian random field suggest (and results of fully non-linear N-body simulations demonstrate) that a large clustering amplitude is indicative of a population of objects that inhabit high peaks (dense regions) and, conversely, that a low clustering amplitude implies a sample that probes lower peaks of the field and less dense environments. Thus, we infer that Seyferts and HI galaxies tend to prefer less crowded environments than typical galaxies, while LINERs seem to follow the galaxy spatial distribution, as their correlation function is, within the uncertainties, identical to that of the whole galaxy sample. The cause of the relatively steeper slope of the CF of Seyferts is unclear, but it may be further evidence of a special environment for promoting their activity. Our analyses seem to confirm the trends reported by Miller et al. (2003), which show a the predilection of the star-forming galaxies to reside in more sparsely populated regions, and of the passive, non-line-emitting galaxies to live in more dense environments.

3.1.2. \([\text{O III}], [\text{O I}]\) line luminosities

Figure 10 shows the CF of subsamples selected by luminosities of the forbidden narrow emission lines \([\text{O III}]\) and \([\text{O I}]\). We refer to objects that have log \( L_{[\text{O III}]} > 40 \) and log \( L_{[\text{O III}]} < 37 \) as the high and low \( L_{[\text{O III}]} \) samples respectively. Line luminosities are in units of erg s\(^{-1}\). Conforming with the definitions listed in Table 2, a similar sample name designation is used for the low and high \( L_{[\text{O I}]} \) object samples. Our calculations show that objects that are bright in \([\text{O III}]\) and/or \([\text{O I}]\) are significantly less clustered than sources that are less luminous in these lines. We confirm the dependence of the clustering amplitude on \( L_{[\text{O III}]} \), as seen by Wake et al.
For the first time, we detect an even stronger variation of clustering with $L_{\text{[OIII]}}$. 

$\text{[O III]}$ is, as also discussed by Kauffmann et al. (2003), only weakly affected by residual star-formation, and therefore capable of distinguishing line emission activity due to accretion from that originating in star formation. As illustrated in Figure 4, using our spectral classification, the high $L_{\text{[OIII]}}$ objects include most of the Seyferts. Interestingly, the subset of luminous emission-line galaxies includes almost no LINERs, which clearly dominate the low luminosity end of the distribution. This trend supports and justifies further the physics behind the correlation found by Kauffmann et al. (2003) between the ionization state of the AGN, as expressed by their position angle in the [N II] diagnostic diagram, and the AGN [O III] luminosity. It is important to note that, because the number of Seyferts remains small in general, H II’s remain the dominant type in the high $L_{\text{[OIII]}}$ subset, comprising $\sim 40\%$ of these objects; the small fraction of H II’s with high $L_{\text{[OIII]}}$ makes a large contribution to this subsample and so complicates interpretation of our estimates of the CFs.

The CF of the high $L_{\text{[OIII]}}$ systems shows a low clustering amplitude, as expected from the properties of the objects that comprise the sample. Both Seyferts and H II’s are weakly clustered, with amplitudes that are similar, within the uncertainties. The steep slope of this CF, $\gamma \approx 1.4$, suggests a strong influence in behavior from the Seyferts. The low $L_{\text{[OIII]}}$ sources exhibit a CF that is consistent in both the amplitude $r_0 \approx 8.8$ and the slope $\gamma \approx 1.3$ with what we measure for the LINER sample. Such strong correspondence may indicate that low nuclear activity, either originating in star formation or accretion, is predominantly present in highly clustered hosts, and thus more probable in crowded environments.

The variation in the clustering amplitude between the samples of low and high $L_{\text{[OIII]}}$ is very strong, showing a difference of $\sim 6\sigma$. As with selection by $L_{\text{[OIII]}}$, the low $L_{\text{[OIII]}}$ objects have large clustering amplitudes, while the more active sources remain weakly clustered. The dominant majority of low $L_{\text{[OIII]}}$ objects are LINERs, suggesting again that galaxy nuclei that manifest the lowest level of activity are the ones that are strongly clustered. LINERs remain very rare (less than 5%) in the high $L_{\text{[OIII]}}$ sample. The high $L_{\text{[OIII]}}$ subset is highly populated by H II’s (\sim 60\%), and, although this dataset includes the majority of Seyferts, such systems make up only $\sim 9\%$ of the objects. Consequently, the clustering properties of the high $L_{\text{[OIII]}}$ objects are close to those of Seyferts and star-forming galaxies, revealing a relatively small clustering amplitude.

As suggested in Constantin et al. (2004), the strength of the [O I] emission may be related to the strength of the broad emission, and therefore to that of the accretion process. Thus, the clustering dependence on $L_{\text{[OIII]}}$ may be another indication of the relationship between the level of AGN activity in galaxy nuclei and the properties of their host dark matter halos: the active nuclei that exhibit the lowest levels of activity due to accretion are the most clustered, and thus, the most massive ones. Any signature of star-formation activity is feeble, if present at all in these systems; although numerically dominant among the emission-line systems, the H II-like galaxies are significantly underrepresented among the low luminosity objects (see Figure 4). On the other hand, the most active sources, whether in terms of nuclear star-formation or accretion, are weakly clustered, thus we infer that these sources are hosted by relatively less massive systems.

### 3.1.3. Gas Density & Intrinsic Extinction

As Figure 4 illustrates, LINERs stand out among line-emitting objects in other aspects as well. Their electron densities are lower than those of Seyferts but generally higher than in H II’s, and they show little intrinsic obscuration. A clear segregation of LINERs and Seyferts as a function of these parameters, similar to that as a function of $L_{\text{[OII]}}$ and $L_{\text{[OIII]}}$, is possible only for very small subsamples, and thus CF estimates remain impractical. Nevertheless, an investigation of the degree to which the clustering amplitude depends on $n_e$ and $N_H$ remains of interest even if the subsamples characterized by low and high values of these parameters comprise a mix of spectral types. Such analysis offers additional insights into the origin of the clustering differences exhibited by sources with contrasting physical conditions, especially those that may characterize the availability of fuel for accretion or star-formation.

Figure 4 shows a comparison of the CFs estimated for subsamples characterized by low and high values of $n_e$ and $N_H$, with the sample definitions given in Tables 3 and 4. Consistent with the trends found so far, the low $n_e$ subsample, in which 70% of the objects are H II galaxies, has a significantly lower clustering amplitude. Interestingly, the high $n_e$ subsample shows a high clustering amplitude that is consistent with that of the LINERs and the average galaxy sample even though it comprises a generously mixed population of emission line systems: 40% LINERs, 18% Seyferts, 19% H II’s, and 23% unclassified line-emitters (that do not comply simultaneously with all the three diagnostic diagrams, see Section 3.2). It is striking that only a slight number dominance of the LINERs in this subsample of sources with relatively high $n_e$ leads to such results. This trend suggests that the moderately high $n_e$ unclassified emission-line objects have clustering properties more similar to LINERs than to Seyferts or H II’s.

The subsamples split on $N_H$ shows clustering trends that are consistent with the results presented so far. There is a slight suggestion that the less obscured sources ($N_H < 10 \times 10^{20} \text{ cm}^{-2}$), in which the LINERs are in majority, are more clustered, while the highly obscured systems, which are mostly star-forming objects, are less clustered. These differences reflect once more a potential link between the amount, nature, and availability of the fuel that powers the galaxy nuclei and their clustering characteristics: the dusty spirals that usually live in less crowded environments are more likely to harbor H II’s in their centers, while the nuclei of ellipticals would be generally dust-free due to ram-pressure stripping in their more crowded, cluster-like habitats, and thus would show LINER-like activity and physical properties.

### 3.2. CF Estimates and Host Galaxy Properties

Studies based on large redshift surveys, 2dF and SDSS in particular, have clearly shown that spatial clustering of galaxies depends on galaxy structure (or morphology),
luminosity, and internal star formation history, as measured by spectral type or colors (Norberg et al. 2002; Zehavi et al. 2002; Budavari et al. 2003; Madgwick et al. 2003). The sense of this dependence is that more luminous, redder, and earlier type galaxies, which also harbor older stellar populations, are generally more clustered than the rest of the galaxies.

The well-described dependence of clustering on galaxy properties might explain some of the difference we see between the clustering amplitude of H II nuclei, which seem to prefer bluer, later-type hosts, and that of AGN-like objects (i.e., Seyferts and LINERs), which are more likely to be hosted by red, early type galaxies. As illustrated in Section 2.4, AGN samples defined based on simple classification schemes are dominated by LINER-like behavior and are therefore more clustered than the H II systems (i.e., Figure 9). The median values of concentration indices, colors, and other parameters of the host galaxies of our spectroscopically-classified subsamples (see Figure 12 and Table 5) clearly illustrate the distinct preferences of the stellar and non-stellar ionized galaxy nuclei for different host types.

The variation of clustering among Seyferts, LINERs, and H II systems is not, however, so easily explained. When we use more detailed spectroscopic classification to separate Seyferts and LINERs, we find that they differ significantly in their clustering properties despite our finding, consistent with previous studies (e.g., Ho et al. 1997b), that the large-scale, global properties of their hosts are quite similar. Although there is a slight shift of the Seyferts toward bluer and later type hosts than LINERs, comparison using a Kolmogorov-Smirnov (KS) test shows that their distributions in concentration index, C, and color u−r are consistent with being withdrawn from the same parent population (KS \(C_L = 0.104, KS u−r_L = 0.617\)). LINERs' host morphology (C) and color (u−r) distributions are quite different from those of the whole galaxy sample, both in the shapes and median values of these distributions (KS \(C_{\text{passive}} = 4.8e-14, KS u−r_{\text{passive}} = 5.6e-11\)). Yet, the clustering amplitudes of LINERs and typical galaxies are very similar. Likewise, Seyferts and the passive galaxies are close to sharing the same parent distributions in host properties (KS \(C_{\text{passive}} = 0.108, KS u−r_{\text{passive}} = 4.2e-4\)), although their clustering properties are markedly different. Furthermore, the clustering amplitudes of Seyferts and H II's are quite similar, but their galactic hosts are, beyond any doubt, very different in type.

We can, therefore, safely conclude that the differences we find between the clustering amplitudes of different spectral classes of galaxy nuclei, and in particular Seyferts and LINERs, are not driven primarily by the morphology-density relation.

In the light of the effects the host characteristics have on clustering, the differences in \(r_0\) exhibited by the low and high \(L_{\text{[OIII]}}\) and \(L_{\text{[OIII]}}\) samples merit further attention as well. As it can be seen from Figure 13 these datasets are clearly distinguishable in their median concentration indices and u−r colors, with which the clustering amplitudes seem to correlate. The sources with low [O I] and [O III] luminosities are also the reddest u−r \(\approx 2.8\) and generally early in their morphological type, as revealed by their high median value of C \(\approx 3\), and present strong clustering. The most luminous sources in [O I] and [O III] are bluer (u−r \(\approx 2.3\)) and of later type morphologies (C \(\approx 2.6\)), and are weakly clustered. Thus, it may be apparent that the different clustering amplitudes for the high and low luminosity objects is at least partially driven by the morphology-density relation that galaxies usually obey. However, based on exhaustive studies of the color and host type dependence of the correlation function, Zehavi et al. (2002, 2003) reveal differences in \(r_0\) of only \(\lesssim 2h^{-1}\) Mpc, which are smaller than those between our low and high \(L_{\text{[OII]}}\) systems, and especially those we observed between the objects with high and low \(L_{\text{[OII]}}\) (see Table 2). Moreover, our subsamples span a significantly reduced range of values in their median C and u−r than those employed in the Zehavi et al. comparisons, suggesting that, the morphology-color effect on the discrepancy in \(r_0\) measured in our samples should be even smaller. We thus conclude that the differences in the clustering amplitudes of objects that differentiate themselves through their [O I] and [O III] luminosities are primarily constrained by these parameters, and only secondarily by their host properties. In the following section we discuss various scenarios that may explain such effects.

4. DISCUSSION

4.1. AGN Fueling & Life Cycle

Under simple assumptions, the amplitude of the correlation function can be used to estimate the typical mass of the dark matter halos in which the objects reside (e.g., Kaiser 1984; Grazian et al. 2004; Magliocchetti et al. 2004), and for AGN in particular, to estimate their typical lifetime (Martini & Weinberg 2001). If, as expected in the hierarchical picture of structure formation, the most massive systems are more clustered, our results suggest that Seyferts live in less massive halos than those that host LINERs. The empirical relation between black hole mass \(M_{\text{BH}}\) and galaxy mass \(M_{\text{DM}}\) (Ferrarese 2002; Baes et al. 2003; Ferrarese & Merritt 2000) is based on the \(M_{\text{BH}} - \sigma_{\text{gal}}\) relation which seems to hold similarly for both inactive galaxies (Gebhardt et al. 2000a; Ferrarese & Merritt 2000) and AGN hosts (Gebhardt et al. 2000b; Ferrarese et al. 2001; Nelson et al. 2004; Onken et al. 2004; Greene & Ho 2005), implies that the black holes harbored by these galaxy nuclei must scale in the same manner. Thus, Seyferts’ black holes must be smaller than those that live within LINERs.

Together with the different trends we see in the physical properties of galaxy nuclei with distinct spectral classification, this possible distinction in their BH mass offers the missing link in our understanding of their intrinsic nature. The high \(L_{\text{[OII]}}\) and \(L_{\text{[OIII]}}\) of Seyferts suggest that their smaller BHs are accreting at relatively higher rates or maybe more efficient than the larger BHs whose activity is classified as LINER-like. The low activity level of LINERs’ massive BHs may well be a consequence of the fact that these systems have less fuel available for accretion. Both the electron density and column density of surrounding (dusty) material are significantly lower in these objects, suggesting a scarcity of matter supply. Seyferts, on the other hand, show the highest values of \(n_e\) and a relatively large range in \(N_H\) suggesting that in these sources fuel can easily be provided at levels suffi-
cient to sustain accretion at modest levels.

In this scenario, Seyferts grow their BHs at a relatively fast pace, and this activity would be expected to last for short periods of time. Such an highly active phase would end with the exhaustion of surrounding accretion fuel, when the BH mass reaches levels similar to those of LINERs, after which weak (or inefficient) accretion can continue for much longer time. This picture is consistent with the relative numbers of Seyferts and LINERs. If Seyferts are simply the manifestation of the stronger but short accretion phase, then their numbers must be small, especially when compared to those of LINERs that subsist much longer at low activity levels and thus are more common among active galaxies. The results of our AGN classification, together with other previous studies discussed above (see section 2.1.1) indicate indeed that LINERs numerically dominate the AGN samples, as the LINER-type emission is found in at least three times more galaxies than Seyfert-like activity.

If LINERs harbor more massive black holes then, as suggested by the close connection between BH mass and bulge size in their host galaxies (e.g., Magorrian et al. 1998, Graham et al. 2001), LINERs’ hosts must be more bulge-dominated than those of Seyferts. Although we have seen that the Seyfert and LINER overall distributions in C and color are very similar, the median values show small differences (see Table 2) indicating slightly higher C and redder $u - r$ colors for LINERs, consistent with the expected trend. Future larger SDSS samples, together with more successful morphology classification schemes (Park & Choi 2005) should allow more detailed investigations of such feature.

4.2. The H II & the Transition Systems

Further, we may extrapolate this picture to galaxies whose nuclear activity is dominated by star-formation, the H II systems. Assuming that the same scaling relations hold for these sources, the weak clustering of H II’s implies that they harbor relatively small black holes in their centers. In general, they show moderately low, quiescent [O I] and [O III] luminosities. However, because they are very numerous, the high luminosity tail of their line-emission distribution dominates the statistics of the sub-sample of luminous actively line-emitting galaxy nuclei. There is a generally high level of obscuration in these galaxy nuclei, as shown by their large column densities $N_H$. The overall low $n_e$ values that characterize the emitting gas in H II’s suggest that fuel for eventual accretion is tenuous in these galaxies. Thus, if their BHs are accreting, the rates must not be high, or if the BHs are actively accreting, the the resulting radiation is highly obscured. In other words, even if the kinetic energy output from the central few tens of parsecs is as strong as that of the Seyferts, there is only a very small fraction that gets transferred to the thermal plasma to produce the enhancement in the optically collisionally excited lines, at the level seen in the accretion-powered Seyferts or LINERs. Consequently, any accretion signature due to robust activity would be sub-dominant in the H II’s spectra because it would be buried in the surrounding emission originating in stellar heating.

We now turn to consider the Transition objects using similar arguments. Figure 14 illustrates the CFs estimated for the Transition and Pure LINER samples (see definition in Table 1) along with a comparison of the distributions in $C$ and the $u - r$ color for these samples. Figure 15 illustrates how Transition systems and the Pure LINERs compare in terms of $L_{\text{[OIII]}}$, $L_{\text{[OIII]}}$, $N_H$, and $n_e$. The results of the power-law fitting of the correlation functions, that are also recorded in Table 2 show a great discrepancy in the clustering amplitudes of these two sub-categories of LINERs. The Pure LINERs, which constitute the majority of the LINER sample, remain strongly clustered with an amplitude similar to that measured for the whole LINER sample and consistent with that of the whole galaxy sample. The Transition objects are clearly less clustered than the Pure LINERs, and are comparable in their clustering amplitude to Seyferts and/or H II’s.

As expected, the intrinsic properties of the Transition objects suggest a tendency toward a more Seyfert or H II-like behavior. These nuclei show systematically larger $L_{\text{[OIII]}}$, $L_{\text{[OIII]}}$, and higher $N_H$ than the Pure LINERs. However, the nature of the Transition objects continues to remain ambiguous. Their lower $n_e$ makes them more similar to the star-forming systems than to Seyferts, while their host galaxies remain in general of early-type, in the range of those that usually harbor Seyferts, not H II’s. Their clustering amplitude is significantly lower than that of the typical red, early type galaxies. Thus, a possibility for the power source in these objects is a small black hole that is accreting at rates that are higher than those in LINERs but still low, seemingly due to the general fuel deficiency. Given that the obscuration remains relatively low in these sources, moderate activity originating from such a source could explain their intermediate behavior between that of traditional AGN and H II systems.

Our own Milky Way seems to provide the most nearby example of a typical H II or H II/LINER transition object, as outlined above. There is compelling evidence for the presence of a relatively low mass Galactic black hole (e.g., Schodel et al. 2003, Ghez 2001) that is accreting, as shown by radio (Sgr A* source, Balick & Brown 1974) and X-ray (Baganoff et al. 2003) emission, and which exhibits large circumnuclear extinction ($A_V \approx 31$ mag; Rieke, Rieke & Paul 1989). In terms of its nebular luminosity, surface brightness distribution, and ionization, Shields et al. (2005) reason that the Galaxy may be a typical example of a H II or maybe a transition system, although its Sbc Hubble type is most consistent with the former type of emission-line object.

4.3. Environment & AGN Activity

The importance of the role that interactions with other galaxies plays in triggering efficient gas fueling in AGN remains a matter of debate. Although there is some evidence that the frequency of occurrence of Seyferts in pairs and groups is higher than among isolated galaxies (Kelm et al. 1998, 2004), no clear relationships are found between the presence of AGN and detailed morphological properties that are signatures strong interactions (e.g., bars, see Keel 1996, Ho et al. 1997).

A problem with many previous investigations of the environments of AGN is that their definition as a class relies on relaxed definitions of such objects which result, as illustrated in Section 2.1.1 in samples dominated by LINERs. The influence of Seyfert-like behavior re-
mains negligible in such studies due to the small fraction of such objects. For this reason, results that infer that AGN environments are undistinguishable from those of the normal, inactive galaxies (Miller et al. 2003; Wake et al. 2004) apply for LINERs, but not necessarily for Seyferts. We find that LINERs and Seyferts differ not only in their intrinsic physical characteristics (ionization level, obscuration, density of their line-emitting gas) but also in their clustering properties. It is also important to understand that environmental properties may vary with spatial scale. For instance, the unusually steep slope of the Seyferts’ CF suggests that these sources may present extra clustering at small scales, despite a generally low clustering amplitude at scales $> 10 h^{-1}$ Mpc. Thus, it is possible that LINERs and Seyferts share similar small scale habitats, while their large scale environments are quite different.

To compare the small-scale environments of different classes of AGN, we examine local galaxy density on a scale $\sim 1 h^{-1}$ Mpc. For the spectroscopic samples we built and analyzed in this study, we employ Blanton et al. (2005) galaxy overdensities estimated by counting all galaxies around each object in cylinders $1 h^{-1}$ Mpc in radius and $1,600$ km/s long (in the redshift direction) centered on the source. The tracer galaxies are a volume-limited sample with $0.017 < z < 0.082$, and $-23.5 < M_r < -19.5$. We find that the median overdensities for Seyferts, LINERs and H II’s are 6.15, 6.13, and 4.17, respectively, suggesting more crowded environments for active galaxies with line emission dominated by star formation. The passive galaxies show the largest median overdensity, 9.39. For comparison, the median overdensity for the whole galaxy sample is 6.34. These results support our inferences from estimates of the CF that suggest a higher density environment for passive galaxies and a relatively lower local density around the H II’s. The similarity in galaxy densities around Seyferts and LINERs on this small scale may be a consequence of the Seyferts’ extra clustering at small scales; the density estimates characterize the environment only at scales $\sim 1 h^{-1}$ Mpc, while the clustering amplitudes, which show LINERs to be more strongly clustered, are obtained by fitting the CFs at scales $3 < r < 10 h^{-1}$ Mpc.

If the typical environments of LINERs are no different than those of galaxies, then major mergers, harassment (Lake, Katz & Moore 1998), or other cluster-related mechanisms are unlikely to be the trigger for fuelling nuclear activity in these objects. As discussed by Waskett et al. (2005), because the number of close companions of modest luminosity AGN (i.e., LINERs) is very similar to that of non-active galaxies (Waskett et al. 2005; Grogin et al. 2003), and because their spatial distribution is unbiased relative to that of the normal galaxy population (Miller et al. 2003, Wake et al. 2004, Section 3.12) of this study), a possible cause of activity in these AGN may be minor mergers.

Seyferts, on the other hand, might still be caused by, e.g., galaxy harassment, with these more luminous AGN being relatively recent additions to clusters, and thus situated either at the periphery or in the process of falling into dense regions. A clumpy small-scale spatial distribution would then be a natural consequence of such environments. Such migrants are also more likely to have large gas reservoirs and only slightly disturbed morphologies. We have shown that Seyferts are distinct from other emission-line nuclei in their high electron densities. So far, morphological studies of Seyfert hosts reach no clear, statistically significant conclusion on the degree of harassment in comparison with those of the normal galaxies.

5. SUMMARY & CONCLUSIONS

We investigate the spatial clustering of low-luminosity AGN in the nearby universe, using spectroscopic classification and measurement of related physical parameters to examine this clustering as a function of their nuclear emission properties. We estimate and compare the CFs for sub-samples selected on various properties, and analyze these results in light of the well-known dependence of clustering on halo mass, as well as the empirical relation between black hole mass and galaxy mass. This study reveals a strong connection between the galaxy clustering amplitudes, hence the properties of their dark matter halos, and the intrinsic physical properties of these objects’ line-emitting regions. A brief overview of our main results is given in Table 6, and we list our conclusions as follows.

(i) The amplitude of the two-point correlation function is strongly dependent on galactic nuclear emission properties. By employing detailed spectroscopic classification schemes we find that, contrary to previous claims, bona fide low luminosity AGN, spectroscopically classified as Seyfert galaxies, are clearly less clustered than the average galaxies, and therefore do not tracing the same underlying structures. LINERs, on the other hand, exhibit a clustering amplitude higher than that of Seyferts, and consistent with that of the whole galaxy sample.

(ii) AGN defined using only the [N II]/Hα vs. [O III]/Hβ diagnostic diagram are dominated by the LINER-like objects, and include a significant fraction of objects with vigorously strong nuclear star-formation activity. Two, or better, three-dimensional diagnostic diagrams that employ forbidden lines whose emissivity is less sensitive to abundance effects, like [S II] and/or [O I], are necessary to accurately separate sources powered by accretion from those in which star-forming activity is dominant.

It is, therefore, important to emphasize that more lenient AGN definitions are strongly biased toward LINER-like behavior, and consequently, do not represent the traditional AGN activity and physical characteristics. LINERs differ from Seyferts in many of their intrinsic characteristics. In LINERs, the emitting gas density is lower than in Seyferts, and the obscuration is generally at very low levels. As suggested by the large differences we observe in the [O I] and [O III] line luminosities, the accretion activity is significantly reduced in LINERs while in Seyferts is generally high.

(iii) Emission parameters such as $L_{[OII]}$ and, in particular, $L_{[OIII]}$, which are sensitive to the level of nuclear accretion activity, play an important role in differentiating systems of different clustering properties. Highly luminous and, therefore, more active objects are significantly less clustered than faint, less active ones; the latter types are strongly represented by LINERs.

(iv) The amount and nature of the fuel available for nuclear activity, whether accretion or star-formation, is also
connected to the spatial clustering characteristics. The less obscured systems and the nuclei with moderately high emitting gas densities are more clustered. Such sources are highly represented by LINERs.

(v) Host galaxy properties and, consequently, the morphology-density relation, do not strongly influence the difference in the spatial clustering found between Seyferts and LINERs, and play only a secondary role in the large differences between the clustering amplitudes of the low and high $L_{[OIII]}$ and/or $L_{[OII]}$ subsets.

(vi) Based on the strong relationship between host halo mass, as given by the clustering amplitude, and the spectral properties of galaxy nuclei, we speculate that Seyfert activity arises from active accretion onto small BHs that is abundantly fueled, while LINERs’ emission-line properties originate from ionization by slowly accreting massive BHs whose circumnuclear material supply is generally low. In this scenario, H II systems may be interpreted as harboring relatively small, weakly accreting BHs whose activity remains ensheathed within the surrounding star-forming activity. The similarity of the Milky Way nuclear system to that of a typical H II galaxy in this picture provides a means of integrating the Galactic Center into the larger context of galaxy nuclei.

Support for this work was provided by NASA through grant NAG5-12243. M. S. V. thanks the Department of Astrophysical Sciences at Princeton University and the Korea Institute for Advanced Study for their hospital-
Fig. 1.— Volume-limited sample definition in redshift-absolute magnitude space. The smooth curves delineate the SDSS main galaxy sample boundaries corresponding to our choice of apparent magnitude limits of $14.5 < m_r < 17.7$. The rectangle shows the limits in absolute magnitude of the parent sample of galaxies used in this analysis.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., & Flannery, B.P., 1992, Numerical Recipes in C (Second ed.; Cambridge: Cambridge Univ. Press)
Richards, G.T., et al., 2002, AJ, 123, 2945
Rieke, G.H., Rieke M.J., & Paul, A.E., 1989, ApJ, 336, 752
Scheib, R., Ott, T., Genzel, R., Eckart, A., Mouawad, N., & Alexander, T., 2003, ApJ, 596, 1015
Schlegel, D. J., Finkbeiner, D. P., & Davis, M., 1998, ApJ, 500, 525
Schneider, D.P., et al., 2002, AJ, 123, 567
Schneider, D.P., et al., 2003, AJ, 126, 2579
Shields, J.C., et al., 2005, ApJ, submitted.
Storchi-Bergmann, T., & Pastoriza, M.G., 1989, ApJ, 347, 195
Storchi-Bergmann, T., & Pastoriza, M.G., 1990, PASP, 102, 1359
Strateva, I., et al., 2001, AJ, 122, 1861
Strauss, M., et al., 2002, AJ, 124, 1810
Tegmark, M., et al., 2004, ApJ, 606, 702
Tremonti, C.A., et al., 2004, ApJ, 613, 898
Veilleux, S., & Osterbrock, D. E., 1987, ApJS, 63, 295
Wake, D., et al., 2004, ApJ, 610, 85
Waskett, T. J., Eales, S. A., Gear, W. K., McCracken, H. J., Lilly, S., & Brodin, M., 2005, MNRAS, 363, 801
Zehavi, I., et al., 2002, ApJ, 571, 172
Zehavi, I., et al., 2005, ApJ, 630, 1
York, D. G., et al., 2000, AJ, 120,1579
Fig. 2.— Emission-line diagnostic diagrams with separate density contours for the four subclasses of narrow emission line galaxies that have all 6 features (Hα, Hβ, [O III], [N II], [S II], and [O I]) detected in emission at 2σ significance. The transition objects are shown in red. Density contour lines correspond to factors of $n$ of the total number of objects in each class, where $n = 0.1, 0.2, 0.3, 0.5, 0.7, 0.9$ starting from outermost contour. The contours corresponding to the Transition objects are overplotted in red, and show the lines with $n = 0.1, 0.3, 0.6, 0.9$. The solid curve is the [Kewley et al. (2001)] theoretical separation between accretion and star-forming systems, and the dotted lines are drawn at ±0.1 dex of this prediction.
Fig. 3.— Emission-line diagnostic diagrams showing how the AGN defined using definitions of Miller et al. (left panels, “2-line” AGN in red, and “4-line” AGN in magenta and multicolored density contours), and Kauffmann et al. (right panels, red + blue points) distribute as a function of [S II]/Hα and [O I]/Hα line-flux ratios. For the latter case, the objects situated below the Kewley et al. curve in the [N II]/Hα vs. [O III]/Hβ diagram, which are most likely to include contamination from star-forming activity, are overplotted in blue in this diagram, and are shown in multicolored contours in the other two diagrams. The density contours presented in Figure 2 are shown in the background for comparison.

Fig. 4.— Redshift distributions for different subclasses of objects, in the absolute magnitude limited samples. The left panels correspond to spectral classification that does not include conditions on [O I] and [O I]/Hα while the right panels show objects whose classification includes these parameters (note the lower number of the emission-line systems, especially at higher redshifts). The histograms are re-normalized to have the total number of objects equal to that of the galaxy sample. The solid curves show the corresponding 2nd order polynomial fits for subclasses of objects in the absolute magnitude limited samples, while the dotted line shows the polynomial fit for the volume limited galaxy sample.
Fig. 5.— Distributions in apparent and absolute $r$ magnitude of the hosts of different spectroscopically-selected subclasses, all of which were drawn from a volume-limited parent galaxy sample. The histograms are re-normalized to have the total number of objects equal to that of the galaxy sample.

Fig. 6.— Distributions in log $L_{[O\ I]}$, log $L_{[O\ III]}$, intrinsic column density $N_H$, and $[S\ II] \lambda 6716/\lambda 6731$ line ratio, for different subclasses of objects, in the absolute magnitude limited samples. In the latter plot, the shaded areas indicate corresponding values of $n_e$ that are beyond the minimum and maximum theoretical limits, max $n_e = 10^5$, min $n_e = 10^1$ cm. Luminosities are in erg s$^{-1}$. The histograms are re-normalized to have the total number of objects equal to that of the emission-line galaxy sample.
Fig. 7.— Upper panels: Comparison of the CFs estimated for the whole galaxy volume limited sample, for the objects that show activity in the form of emission lines, and for the “passive” (i.e., relatively normal) galaxies. Lower panels: Comparison of the CFs for different spectral classes of objects. The insets in the upper plots show the $\chi^2$ contours indicating the 68.3%, 90%, and 99% confidence regions for the passive and the actively line-emitting galaxies. The left panels correspond to spectral classification that excludes [O I], while the right panels show the results for sample definition that includes this feature; here, the CF of the emission-line galaxies show lower signal-to-noise and the power-law fits have larger error contours. The passive galaxies are significantly more clustered than the “active” line-emitting galaxies. $\chi^2$ contours corresponding to the CF’s of Seyferts, LINERs and H ii’s are illustrated in Figure 8.
Fig. 8.— Confidence intervals of the power-law fit parameters, $r_0$ and $\Gamma$. Contours indicate the two-dimensional 68.3%, 90%, and 99% confidence regions for the CF of Seyferts, LINERs, and H II galaxies. The left panels show results for sample definitions that do not employ conditions on [O I], while for the right panels conditions on [O I] are used. Each plot shows, for comparison, the results for the volume limited galaxy sample; the axis range is the same in all plots.
Fig. 9.— Comparison of the CFs estimated for AGN and line emitting non-AGN classified using the definitions of Kauffmann et al. (left), and Miller et al. (right). When such AGN samples are analysed, the LINER behavior is dominant, thus, their clustering properties are similar to those of the galaxies.

Fig. 10.— The CFs for object samples characterized by different ranges of \([\text{O} III]\) (left) and \([\text{O} I]\) (right) emission-line luminosities. The galaxy correlation function is also shown for comparison in both panels. Sources with lower \(L_{\text{[OIII]}}\) and low \(L_{\text{[OI]}}\) show significantly stronger clustering. Comparisons of the \(\chi^2\) contours corresponding to the 68.3%, 90%, and 99% confidence regions are indicated in the inner upper right insets in each plot.
Fig. 11.— *left*: CFs for object samples that are characterized by different electron densities \((n_e)\) as obtained from [S II] line ratios (see text for details). *right*: CFs for object samples with high and low levels of intrinsic absorption, \(N_H\). The galaxy correlation function is also shown for comparison in each plot. Confidence limits for the power-law fits to the CF for the subsamples of low and high \(n_e\) and \(N_H\) respectively are compared and shown in the upper right inset of each panel.

Fig. 12.— Distribution in the concentration index \(C\) (left), and the \(u - r\) color (right) for different subclasses of objects, in the absolute magnitude limited samples. The histograms are re-normalized as in Figure 6. Note the similarity between the distributions in these parameters for the Seyfert and LINER samples.

Fig. 13.— Relationships between the clustering amplitude \(r_0\) and the median values of the concentration index \(C\) (left), and the \(u - r\) color (right), for the corresponding samples of different emission-line properties.
Fig. 14.— left: CFs corresponding to Pure LINERs and Transition objects: the transition systems are significantly less clustered than the Pure LINERs. right: Comparison of distributions in $C$ and $u - r$ for these two subsets of LINERs; the histograms are re-normalized to have the total number of objects equal to that of the whole LINER sample. Note the slight shift toward bluer and later type hosts for the Transition objects.

Fig. 15.— Comparison of distributions in $L_{[OIII]}$, $L_{[OI]}$, $N_H$, and $n_e$ for Transition objects and Pure LINERs. The histograms are re-normalized to have the total number of objects equal to that of the whole LINER sample. Transition object properties show a slight tendency toward a more H ii-like behavior.
### TABLE 1
OBJECT SAMPLE DEFINITION: SPECTRAL CLASSES.

| Sample           | \(N_{\text{obj}}\) | \([\text{O III}]/\text{H}\beta\) | \([\text{N II}]/\text{H}\alpha\) | \([\text{S II}]/\text{H}\alpha\) | \([\text{O I}]/\text{H}\alpha\) |
|------------------|---------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| Emission-line    | 54,410              | > 0                             | > 0                             | > 0                             | > 0                             |
| Seyferts         | 2,392               | > 3                             | < 0.6                           | < 0.4                           | > 0.08                          |
| LINERs           | 7,468               | < 3                             | > 0.6                           | > 0.4                           | > 0.08                          |
| ...pure liners   | 5,313               | < 3                             | > 0.6                           | > 0.4                           | > 0.17                          |
| ...transition    | 2,155               | < 3                             | > 0.6                           | > 0.4                           | > 0.08; < 0.17                  |
| H II nuclei      | 34,781              | > 0                             | < 0.6                           | < 0.4                           | < 0.08                          |

Note. — Columns (3) to (6) indicate the range of values in the line flux ratios.

### TABLE 2
ABSOLUTE MAGNITUDE LIMITED CORRELATION FUNCTION SAMPLES

| Sample                        | \(N_{\text{obj}}\) | \(r_0\)     | \(\gamma\) | \(\sigma_{r_0}\gamma/\sqrt{\sigma_{r_0}\sigma_{\gamma}}\) |
|-------------------------------|---------------------|-------------|-------------|-------------------------------------------------|
| all galaxies                  | 45,303              | 7.80±0.49   | 1.17±0.05   | -0.82                                          |
| passive                       | 6,383               | 11.42±1.17  | 1.28±0.07   | -0.82                                          |
| Emission-line\(^a\)          | 12,676              | 6.33±0.94   | 1.30±0.16   | -0.15                                          |
| Seyferts\(^a\)               | 829                 | 5.67±0.62   | 1.56±0.17   | -0.58                                          |
| LINERs\(^a\)                 | 3,263               | 7.82±0.64   | 1.39±0.09   | -0.53                                          |
| ...pure LINERs               | 2,424               | 7.39±0.67   | 1.33±0.11   | -0.46                                          |
| ...transition                | 839                 | 5.38±0.71   | 1.35±0.23   | -0.18                                          |
| H II nuclei\(^b\)            | 6,622               | 5.81±0.53   | 1.28±0.11   | 0.08                                           |
| Emission-line\(^b\)          | 21,497              | 6.45±0.24   | 1.24±0.04   | -0.67                                          |
| Seyferts\(^b\)               | 1,018               | 6.00±0.64   | 1.41±0.15   | -0.48                                          |
| LINERs\(^b\)                 | 5,326               | 7.26±0.61   | 1.29±0.08   | -0.73                                          |
| H II nuclei\(^b\)            | 12,649              | 5.73±0.24   | 1.22±0.06   | -0.46                                          |
| AGN (Kauffmann et al.)        | 4,321               | 7.44±0.37   | 1.25±0.07   | -0.44                                          |
| non-AGN (Kauffmann et al.)    | 13,630              | 6.26±0.69   | 1.24±0.11   | -0.52                                          |
| AGN (Miller et al.)          | 9,846               | 7.01±0.43   | 1.27±0.06   | -0.63                                          |
| non-AGN (Miller et al.)      | 15,178              | 5.02±0.59   | 1.24±0.12   | 0.03                                           |
| log \(L_{\text{II}}\) > 40   | 2,041               | 6.22±0.57   | 1.41±0.11   | -0.24                                          |
| 38 < log \(L_{\text{II}}\) < 39 | 10,464             | 7.17±0.44   | 1.21±0.05   | -0.63                                          |
| log \(L_{\text{II}}\) < 37   | 1,580               | 8.79±0.75   | 1.31±0.10   | -0.54                                          |
| log \(L_{\text{II}}\) > 39   | 4,561               | 6.11±0.27   | 1.34±0.07   | -0.36                                          |
| 38 < log \(L_{\text{II}}\) < 38 | 2,153             | 13.15±1.09  | 1.33±0.07   | -0.77                                          |
| high \(N_H\)                 | 2,575               | 6.91±0.57   | 1.25±0.09   | -0.19                                          |
| low \(N_H\)                  | 2,921               | 7.38±0.51   | 1.35±0.07   | -0.68                                          |
| high \(n_e\)                 | 788                 | 7.50±0.88   | 1.15±0.13   | -0.88                                          |
| low \(n_e\)                  | 3,435               | 5.51±0.45   | 1.43±0.11   | -0.11                                          |

Note. — All samples use \(14.5 < r < 17.7\), \(0.05 < z < 0.12\), and \(-21.6 < M_r < -20.2\). \(r_0\) is in units of \(h^{-1}\) Mpc. \(\sigma_{r_0}\gamma/\sqrt{\sigma_{r_0}\sigma_{\gamma}}\) is the normalized correlation coefficient between \(r_0\) and \(\gamma\).

\(^a\)definition criteria include \([\text{O I}], [\text{O II}]/\text{H}\alpha\)
\(^b\)definition criteria exclude \([\text{O I}], [\text{O I}]/\text{H}\alpha\)
\(^c\)the line luminosities are in units of erg s\(^{-1}\)

### TABLE 3
OBJECT SAMPLE DEFINITION AS A FUNCTION OF \(n_e\)

| Sample         | \(N_{\text{obj}}\) | \([\text{S II}]\lambda6716/\lambda6731\) | \(n_e\) (cm\(^{-3}\)) |
|----------------|---------------------|------------------------------------------|------------------------|
| high \(n_e\)   | 5,922               | \(\leq 1.05\)                            | \(\geq 5 \times 10^2\) |
| low \(n_e\)    | 27,116              | \(> 1.32\)                               | \(\leq 1 \times 10^2\) |
### TABLE 4

**Object Sample Definition as a Function of \( N_H \)**

| Sample       | \( N_{obj} \) | \( N_H \)    | \((E - B)_V\) |
|--------------|---------------|--------------|---------------|
| low \( N_H \) | 19,343        | < \( 1 \times 10^{20} \) cm\(^{-2}\) | < 0.20        |
| high \( N_H \) | 8,573         | > \( 2 \times 10^{20} \) cm\(^{-2}\) | > 0.45        |

### TABLE 5

**Median Values in Absolute Magnitude Limited Samples**

| Sample               | \( N_{obj} \) | \( C \) | \( u - r \) | \( \log L_{[O I]} \) | \( \log L_{[O III]} \) | \( N_H \) | \( n_e \) |
|----------------------|---------------|---------|-------------|----------------------|----------------------|---------|---------|
| emission-line        | 12,676        | 2.61    | 2.31        | 38.80                | 39.26                | 13.44   | 1.7 \( \times 10^1 \) |
| seyferts             | 829           | 2.82    | 2.63        | 38.93                | 39.81                | 11.65   | 1.9 \( \times 10^2 \) |
| liners               | 3,263         | 2.93    | 2.73        | 38.34                | 38.65                | 9.64    | 1.2 \( \times 10^2 \) |
| ...pure LINERs       | 2,424         | 2.99    | 2.79        | 38.22                | 38.48                | 8.19    | 1.2 \( \times 10^2 \) |
| ...transition        | 839           | 2.73    | 2.53        | 38.70                | 39.15                | 12.81   | 1.0 \( \times 10^2 \) |
| H II nuclei          | 6,622         | 2.38    | 2.01        | 38.90                | 39.37                | 15.92   | 1.6 \( \times 10^1 \) |

**Note.** — Col. (2): number of objects in the samples; Col. (3): concentration index; Col. (4): \( u - r \) color for the host galaxies; Cols. (5) and (6): logarithmic values of \([O I]\) and \([O III]\) luminosities in erg s\(^{-1}\); Col. (7): neutral hydrogen column density in units of \( 10^{20} \) cm\(^{-2}\); Col. (8): electron density in cm\(^{-3}\).

### TABLE 6

**Summary of Results & Interpretation**

| Properties | Sample | \( H II \) | Transition | Seyferts | Pure LINERs |
|------------|--------|------------|------------|----------|-------------|
| clustering (\( r_0 \)) | weak | weak | weak | strong \( -like galaxies \) |
| fueling rate \((L_{[O I]}, L_{[O III]})\) | average \( (not low) \) | moderately low | relatively high efficient? | \( (very) low inefficient? \) |
| fuel available \((n_e)\) | low | low | high | moderately high \( (wide range) \) |
| obscuration | high | low | wide range | low |
| host- morphology | blue-ish | red-ish | red-ish | red-ish |
| life-time | small | small | small | large |

- \( M_{BH} \) small
- \( \tau \) short
- \( \tau \) long