Minor Gravitational Interactions as Contributors to Supermassive Black Hole Growth

Jessica Craig

Physics

Department of Physics

Lancaster University

January 2020

A dissertation submitted to Lancaster University for the degree of Master of Science by Research in the Faculty of Science and Technology

Supervised by Dr Brooke Simmons

Abstract

Galaxies have existed for the majority of the lifetime of the Universe, having first formed 400-1000 Myr after the Big Bang, and it is known that most galaxies contain supermassive black holes (SMBHs) at their centres. In recent years, correlations have been found between the evolution of galaxies and that of their central SMBHs. Therefore, understanding the evolution of SMBHs is key to our understanding of the evolution of their host galaxies and that of the Universe. Early studies have shown a link between merger activity in galaxies and the growth of their SMBHs; however, more recent studies have shown that merger-free processes and very minor interactions provide an important additional pathway to SMBH growth. We investigate a relatively unexplored possible pathway to SMBH growth: minor gravitational interactions (short of mergers) between galaxies. We compare the environments of a sample of nearby (z∼0.15) disc-dominated AGN host galaxies, which have no recent history of mergers (the 'merger-free' sample), to a control sample of AGN host galaxies at similar redshifts. We employ three main methods: cylinder searches for potential companion galaxies to the sample galaxies, a search for companion galaxies to the sample galaxies in a well-studied group catalogue and calculation of environment coefficients for each of the galaxies in each sample. We found no significant difference between the environments of galaxies each sample from our cylinder searches with depths of 500 kpc (0.6σ) and 1 Mpc (0.1σ) and only a marginally significant difference between the environments of galaxies from our cylinder search with a depth of 5 Mpc (2.1σ), with the merger-free sample having an increased number of potential companion galaxies found at projected

distances close to 500 kpc compared to projected distances ≤ 350 kpc. This effect was not seen as prominently in the control sample. These results may be interpreted to suggest that whilst AGN host galaxies with no recent history of mergers may have less nearby companion galaxies than control AGN host galaxies, they may have a greater number of galaxies gravitationally interacting with them at greater projected distances and with high peculiar velocities. We suggest that SMBH growth in galaxies with no recent history of mergers may be influenced by passing galaxies with high peculiar velocities, which gravitationally interact with galaxies, but do not become gravitationally bound to them. Such interactions may provide enough kinetic energy to gas in the centre of galaxies to fuel SMBH growth. We found no significant difference (1.1σ) between the environments of galaxies in each sample from our analysis of the results of a well-studied group catalogue. This suggests there may be no difference between the environments of bulgeless AGN host galaxies and control AGN host galaxies.Our calculation of environment coefficients for each of the galaxies in each sample found a marginally significant (2.6σ) difference between the environments of each sample. Overall, some of our results suggest that AGN host galaxies with no recent history of mergers may reside in denser environments than control AGN host galaxies, but further work would be required to ascertain whether or not these results are caused by a true difference in environment or by factors such as fibre collisions or redshift differences between samples.

For my parents: Helen and Ian

Acknowledgements

Declaration

This dissertation is my own work and no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification at this or any other institute of learning. This work was carried out within the Observational Astrophysics Group at Lancaster University between October 2018 and September 2019, with minor corrections having been made to this dissertation until January 2020.

I hereby grant the University of Lancaster the right to retain copies of this dissertation, either physically or electronically.

This dissertation does not exceed 35,000 words.

"The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when one contemplates the mysteries of eternity, of life, of the marvelous structure of reality. It is enough if one tries merely to comprehend only a little of this mystery every day."- Albert Einstein as quoted from the memoirs of William Miller, *Life* magazine, May 2, 1955. Reproduced with permission of the Hebrew University of Jerusalem.

Relevant Publications by the Author

<u>Chapter 1</u>

- Ana Paulino-Afonso, David Sobral, Behnam Darvish, Bruno Ribeiro, Arjen van der Wel, John Stott, Fernando Buitrago, Philip Best, Andra Stroe, and Jessica E. M. Craig. VIS3COS. II. Nature and nurture in galaxy structure and morphology. A&A, 630:A57, Oct 2019.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Galaxies are an integral part of the structure of the Universe. The first galaxies are estimated to have formed 400–1000 Myr after the Big Bang (which occurred 13.8 Gyr (3sf) ago [7],[8]), with candidates for galaxies with redshifts of $z > 8.5$ having been identified [9],[10] and galaxies with redshifts of $z \sim 7$ having been spectroscopically confirmed [11]. Since galaxies have existed for most of the lifetime of the Universe, studying their evolution over time can provide us with a way to understand the history of the Universe as well as predict its future.

The vast majority of galaxies, including our own Milky Way, contain supermassive black holes (SMBHs) at their centres [12],[13],[14]. These are black holes with masses of $\gtrsim 10^6 M_\odot$ (for examples of this see Figure 2.2). As such, it is important to understand the evolution of supermassive black holes and their interaction with their host galaxies as part of understanding the evolution of the host galaxies themselves.

In this dissertation, we investigate one possible consequence of interaction between a galaxies and their central SMBHs— the potential effect of a host galaxy's external environment on the growth of its SMBH.

1.1 Galaxy Evolution

1.1.1 Stages of Evolution and Morphology

Over the course of a galaxy's lifetime, its structure and composition evolves. This results in the existence of galaxies with varied morphologies and structures. Almost a century ago, Hubble first classified [15] 'extra-galactic nebulae' (later found to be galaxies) based on their morphology from photographic images. This classification was later formalised [16],[17] to the Hubble classification scheme that is well-known today. Hubble sorted 'extra-galactic nebulae' into three broad categories— elliptical, spirals and irregular. The 'spirals' category was further subdivided into 'normal spirals' and 'barred spirals', and these were both subdivided into 'early', 'intermediate' and 'late' types based on the size of their nucleus and unwinding of their spiral arms. Hubble hypothesised that elliptical galaxies morphed into spiral galaxies, whose spiral arms unwound as the galaxy evolved, hence Hubble's 'early' type spirals were those spiral galaxies that most closely resembled elliptical galaxies and 'late' type spirals were those with more clearly defined spiral arms.

Spiral galaxies (also referred to as disc galaxies) are characterised by their spiral arms, which extend from the centre of the galaxy. They are usually actively star-forming and therefore are bluer in colour than elliptical galaxies. Elliptical galaxies are mostly featureless and continuous. They are usually redder in colour than spiral galaxies due to having ceased star formation. Irregular galaxies are those which do not have a typical spiral or elliptical shape; they may have features such as trails of gas extending from core mass of the galaxy, or multiple 'clumps' (bright regions that are not at the galaxy's centre).

Since Hubble's classification, the morphology of galaxies has been studied extensively and linked to physical processes occurring during galaxy evolution [18],[19]. This has led to the understanding that spiral galaxies evolve into elliptical galaxies— the opposite of Hubble's hypothesis; however, the term 'early-type' is still regularly used to refer to elliptical galaxies.

1.1.2 Processes Affecting Galaxy Evolution

The main processes which occur during the transformation from a star-forming spiral galaxy to a quiescent elliptical galaxy are mergers between galaxies and quenching (the halting of star formation). Mergers can cause a galaxy's shape to become more similar to that of an elliptical galaxy [20], whilst quenching— which can occur via environmental processes [21]— causes a galaxy's colour to change. Galaxies that actively star-forming tend to be 'bluer' in colour than those that have ceased star-formation (quiescent galaxies).

Correlations have been found between the environment in which galaxies reside and their morphology [22],[23],[24],[25],[26],[27],[28],[29],[30],[31] and also between stellar mass and morphology [25],[26],[31]. Bamford et al (2009) [26] also relate colour to environment. These correlations further indicate that processes that may be related to a galaxy's environment, such as mergers and quenching, are responsible for changes in morphology and colour in galaxies.

Secular processes— which, for the purposes of this dissertation, we define as any process occurring within a galaxy and not involving external factors (sometimes referred to as 'in-situ' processes)— also affect the morphology and evolution of galaxies. For example, disc instabilities can lead to the formation of bars [32]. They may also affect the formation and/or growth of bulges, as discussed in Section 1.1.3.

1.1.3 Bulge Formation and Growth

Most spiral galaxies contain bulges at their centres. These can be generally defined as an area in the centre of a galaxy with a higher density of stars and a greater luminosity than the rest of the galaxy (the disc) [33]. Bulges can be classified into multiple categories [34]. Classical bulges are those that protrude from the galaxy when viewed edge on and display a rounded shape; boxy (sometimes called peanut) bulges also protrude from the galaxy and display a rectangular shape, usually the result of a bar in the galaxy's centre [34]. Areas that are more luminous than the rest of the galaxy but do not protrude (or protrude very little) from the disc also exist [35]— these are referred to as pseudo-bulges (sometimes called disc-like bulges), since they appear the same as a classical bulge from some

angles but not when viewed edge on, and therefore may easily be mistaken for classical bulges.

Classical bulges form when galaxies merge [35], and can also be formed by secular processes, including calm star formation processes [36]. Despite it being possible for bulges to form without mergers, it remains clear that galaxies without a classical bulge present would not have had any significant merger activity, since this would have caused a classical bulge to form.

Pseudo-bulges, on the other hand, are found in galaxies with a quiet merger history [37] and form only via secular processes [35]. Therefore, relevantly to this dissertation, the presence of a pseudo-bulge does not indicate any significant merger history in a galaxy.

Boxy/peanut bulges form as part of a galaxy's bar [32],[34]. Bar formation is a secular process (by the definition we provide in Section 1.1.2), and evidence shows that boxy bulges are formed from disc material [38],[35]; therefore, boxy bulges, like pseudo-bulges, form via secular processes and therefore a barred galaxy which does not also display a classical bulge would not indicate a significant merger history.

1.2 Supermassive Black Holes

As noted at the beginning of this chapter, most, if not all, galaxies in the Universe have a central supermassive black hole (SMBH) [12],[13],[14]. Various scenarios have been suggested for their formation. Some models suggest that very massive, early (Population III) stars collapsed, leaving the beginnings of SMBHs in early galaxies [39],[40],[41]. Others suggest that SMBHs may be formed from the collapse of gas with low angular momentum in the centres of early galaxies [42],[43],[44],[45],[46],[47],[41], or via a mechanism of 'bars within bars', which would also cause gas to collapse into a SMBH [48],[49],[41]. Yet another 'family' of models suggest that SMBHs formed via stellar-dynamical interactions in dense systems of non-Population III stars after the first metals were formed [50],[51],[41].

1.2.1 Active Galactic Nuclei

Active galactic nuclei (AGN) are extremely luminous areas in the centre of galaxies, whose luminosity, unlike that of bulges, does not originate from stellar activity. AGN have been observed in various electromagnetic frequency bands [52], including optical [53],[54],[55] and radio [56],[57]. They are believed to form when gas becomes concentrated around a SMBH, forming an accretion disc. Gas flows into the SMBH and causes the disc to gain angular momentum, which heats the disc, increasing its luminosity [58],[59],[60],[61],[62],[63]. This theory is supported by observations from the Event Horizon Telescope [64].

AGN can be broadly classified as either radio-quiet or radio-loud. Radio-quiet AGN include those found in Seyfert galaxies, low-ionisation nuclear emission line regions (LINERs) and radio-quiet quasars; radio-loud AGN include radio galaxies, blazars and radio-loud quasars [63]. Radio-loud quasars are those that have relativistic jets; these are beams of ionised matter travelling at close to the speed of light away from the AGN. They are associated with magnetic flux around the centre of the disc [65],[61],[66].

Despite many types of AGN existing, it is theorised that all AGN are similar in structure— this is referred to as AGN unification [67],[68],[57].

The work in this dissertation only considers unobscured, radio-quiet AGN;

that is, AGN that are at such an angle that the broad line region is visible to observers and do not have jets. This type of AGN allows the black hole mass to be measured by using relations between emission line width and flux and black hole mass [69],[55].

1.3   Galaxy-Black Hole Coevolution

Correlations have been observed between the properties of supermassive black holes, most notably black hole mass, and the properties of their corresponding host galaxies [14],[70]. This implies that the evolution of supermassive black holes could be linked to that of their host galaxies.

Examples of correlations found between the properties of SMBHs and their host galaxies include correlations between black hole growth and star formation [71],[72]; AGN activity and total stellar mass [73]; AGN incidence and star formation rate [74]; black hole mass and total stellar mass [13],[75],[76]; black hole mass and bulge stellar mass [77],[78],[79],[80],[81]; black hole mass and velocity dispersion [12],[82],[83],[84],[85].

It has been established that supermassive black hole growth is connected to merger activity between galaxies. Such works as [86] (theoretical) and [87],[74] (observational) suggest a connection between major mergers and AGN triggering. Hopkins & Hernquist [88], for example, predict, using an analytic model, that gas is driven into the centre of galaxies during major mergers, thereby triggering AGN activity. Works such as [89] and [90] link the correlations between supermassive black hole growth and bulge growth to both processes being affected by major mergers.

However, it has been shown that supermassive black hole growth also occurs in galaxies without a history of major mergers such as bulgeless galaxies (see Section 1.1 for a discussion of why bulgeless galaxies are unlikely to have a history of major mergers) [91],[54],[55] and Ellison et al [92] suggest that recent mergers may not be the main trigger of AGN activity in nearby AGN host galaxies (z∼0). There is some evidence that, apart from mergers, secular processes such as calm star formation processes may contribute to supermassive black hole growth [93],[94],[95]. Disc instabilities, such as galactic bars, which we also classify as secular processes in this dissertation, are found to contribute very little or not at all to black hole growth and AGN triggering [96],[97],[98],[99]. However, a relatively unexplored factor that may affect supermassive black hole growth is gravitational interactions short of mergers (major or minor). These interactions

may influence the motion of gas in galaxies and therefore contribute to the accretion of gas in the centre of galaxies and therefore the growth and/or formation of supermassive black holes.

Chapter 2

Data

2.1 Galaxy Samples

2.1.1 Merger-Free Sample

In order to investigate supermassive black hole growth in galaxies without a history of mergers, we must analyse a sample of galaxies which are both disc-dominated (and hence have no history of mergers) and host growing black holes. We make use of the Simmons et al (2017) [55] sample of disc-dominated AGN host galaxies for this purpose. The following sections (2.1.1.1, 2.1.1.2) provide a brief description of the sample's properties and the selection processes applied by Simmons et al.

2.1.1.1 AGN Selection

Galaxies hosting unobscured AGN with broad emission lines were selected for this sample, since these allow for the black hole mass to be measured by using relations between emission line width and flux and black hole mass. This allows a comparison to be made between the black hole masses and fluxes of this sample and that of black holes from other samples, such as the control sample described in Section 2.1.2. These emission lines can be classified as broadened based on their width compared to the narrow line component of the spectrum, in this case

the primary component of [OIII] emission. In general, an emission line would be classified as broadened if its full width at half maximum (FWHM) is $\gtrsim 2$ times that of the narrow line component or if its FWHM is > 1000 kms^{-1}. Given that the peak primary [OIII] emission line FWHM for the merger-free sample is ~ 350 kms^{-1}, many of the emission lines for this sample would require a FWHM of $\gtrsim 700$ kms^{-1} to be considered broadened.

The AGN in this sample were selected using multi-wavelength imaging. Un-obscured AGN have characteristic colours in the ultraviolet [100], X-ray [101] and infrared [102] wavelength bands. By combining these bands, uncertainties in each band can be accounted for. Since all-sky surveys exist for each of these wavelength bands, it is possible to select unobscured AGN photometrically using a combination of wavelength bands. Simmons et al select an initial sample of un-obscured AGN using the W2R (WISE, 2MASS and ROSAT) sample [103], which comprises 4316 unobscured AGN sources from the Wide-field Infrared Survey Explorer (WISE) [104], the Two-Micron All-Sky Survey (2MASS) [105] and the ROSAT All-Sky Survey (RASS) [106]. W2R [103] selects unobscured AGN with a confidence level of $\gtrsim 95\%$ by combining infrared colours with X-ray information.

2.1.1.2 Morphology

Following the AGN selection described in Section 2.1.1.1, a subset of galaxies in this sample were selected due to their bulgeless morphology. This means that the galaxies are disc galaxies without a central bulge or with a very small central bulge. The morphology of these galaxies was defined as such by an expert classifier (a classifier who is considerably more experienced in the visual analysis of the morphology of galaxies than members of the general public, in this case Dr Brooke Simmons) who, following AGN selection, visually identified 137 galaxies that either displayed features typical of disc galaxies, such as spiral arms or bars, but which did not show an obvious bulge or displayed an edge-on disc morphology and similarly did not show an obvious bulge [55]. The galaxies may contain pseudo-bulges but none can be deemed to have classical bulges. This classification was performed using colour images from the Sloan Digital Sky Survey [107]. The

sample size was later reduced to 101 galaxies because only these galaxies had spectra available.

The morphology of the galaxies included in this sample indicates that they are unlikely to have undergone major mergers in their recent history—this is demonstrated theoretically by simulation in [108], in which none of the simulated galaxies with a bulge-to-total mass ratio of less than 0.3 have undergone a major merger at redshifts less than 1.

2.1.1.3 Spectroscopic Confirmation

The sources in the sample were spectroscopically confirmed to be unobscured AGN. Spectra for 96 galaxies were obtained from SDSS DR9 [107],[109] (encompassing SDSS DR7 [107],[110]) and the spectra for 5 more galaxies were obtained from observations taken by Simmons et al using the Intermediate Dispersion Spectrograph on the Isaac Newton Telescope between 21-25 May 2014 [55]. The sample galaxies all show broadened Hα line emission, which confirms that they do indeed host unobscured AGN. These spectra were also used to determine the redshift of the galaxies in the Simmons et al (2017) [55] sample, which are used in throughout this dissertation.

2.1.1.4 Removal of Three Galaxies from the Sample

Three galaxies were later removed from the merger-free sample following inspection of Hubble Space Telescope (HST) images processed whilst the research presented in this dissertation was ongoing. These galaxies appeared to be either elliptical galaxies or currently undergoing a merger, although a full morphological analysis of these images, which would confirm or disconfirm this, has yet to be conducted and is outside the scope of this dissertation.

These galaxies remain included in all results and analysis in this dissertation except for the calculation and analysis of environment coefficients, described in Sections 3.4 and 4.4. The exclusion of the three galaxies from these results is discussed in Section 4.4.1 and the environment coefficients that are yielded when these galaxies are not removed are shown in Figure 4.14. It was found that the effect of removing these galaxies on the results was minimal, and therefore all

of the results presented throughout this dissertation remain comparable to one another. For a further discussion of this, please see Section 4.4.1.

The HST band used to observe all three of the removed galaxies was F850LP. This produced red optical images of the removed galaxies. Using this band meant that the AGN at the centre of the galaxies did not appear as bright as they might have done in other bands/wavelengths and therefore the morphology of the galaxies was displayed clearly as opposed to the images being dominated by AGN emission. This also reduced the likelihood of AGN emission either dominating a galaxy's bulge or AGN emission appearing as a pseudobulge. Whilst using this band caused the images produced to be useful for morphological analysis, the same images would not be as useful for analysing the properties of the AGN themselves because the AGN emit little radiation in this band. The same images would also be less useful for locating potential companion galaxies because faint potential companion galaxies (which are at redshifts similar to those of the merger-free and control samples) are not likely to appear clearly in red optical images.

2.1.2 Control Sample

We use a control sample consisting of 101 galaxies from the Shen et al (2011) [53] sample of AGN host galaxies. We refer to this subset hereafter as the 'control' sample. This sample is identical to the 'QSOCONTROL' sample described and used in [55]. It provides examples of more 'typical' quasar-like objects with which to compare the Simmons et al (2017) [55] sample.

This is a suitable control sample to compare to the Simmons et al (2017) [55] (merger-free) sample because the selected galaxies are at similar redshifts to those in the merger-free sample and therefore exist at a similar epoch (see Figure 2.1 for a comparison of redshift distributions between the two samples). Due to this redshift selection, the control sample we use is reduced to a very small subset of the original Shen et al (2011) [53] sample, which contains 105783 galaxies at a wide range of redshifts up to $z = 4.95$ and was selected based on various different broad (with a full width at half maximum of >1000 kms^{-1}) emission lines depending on redshift. All spectra for the original Shen et al (2011) [53] sample

originate from the Schneider et al (2010) [111] compilation of the spectroscopic quasar catalogue from the seventh data release of the Sloan Digital Sky Survey [110]. These spectra were also used to determine the redshift of the galaxies in the control sample, which are used in throughout this dissertation. The 101 galaxies selected for the control sample are those with broadened (with a full width at half maximum of >1000 kms^{-1}) Hα emission lines (indicating unobscured AGN) which are within the redshift range of the Simmons et al (2017) [55] sample.

As noted in [55], the control sample galaxies that have morphological classifications available from Galaxy Zoo [112],[113] are, in general, not found to be disc galaxies, and therefore also cannot be bulgeless disc galaxies. This is expected for AGN host galaxies that are not selected based on their morphology, since as discussed in Section 1.3, supermassive black holes often form via mergers, which cause bulges to form.

2.1.3   Redshift

Figure 2.1 shows the redshift[1] distribution of the Simmons et al (2017) [55] (merger-free) and control (a subset of Shen et al (2011) [53]) samples of AGN host galaxies in terms of absolute numbers of galaxies at each redshift (in redshift bins of size 0.025). There is a tendency towards lower redshifts in the merger-free sample compared to the control sample. This occurs due to overlap between the Shen et al (2011) [53] sample and the Simmons et al (2017) [55] at lower redshifts. In the case of a galaxy being present in both samples, the galaxy does not form part of the 'control' subset of Shen et al (2011) sample [53] that we use. Therefore, the control sample has a greater average redshift than the merger-free sample. The average redshift of the merger-free sample is $\langle z \rangle = 0.132 \pm 0.053$ (3 sf) and the average redshift of the control sample is $\langle z \rangle = 0.141 \pm 0.045$ (3 sf)[2]. It must also be noted that the redshift distributions of the merger-free and control samples are not perfectly symmetrical Gaussian distributions. In both cases, the average redshift is lower than the centre of the redshift range that the

---

[1]All redshifts for galaxies in the merger-free and control samples are spectroscopic. See Sections 2.1.1.3 and 2.1.2 for the sources of these spectra.

[2]The errors in the average redshifts shown are the standard deviations of each redshift distribution around the mean average value for each sample.

distribution covers. This shows that there is a general tendency towards lower redshifts in both distributions.
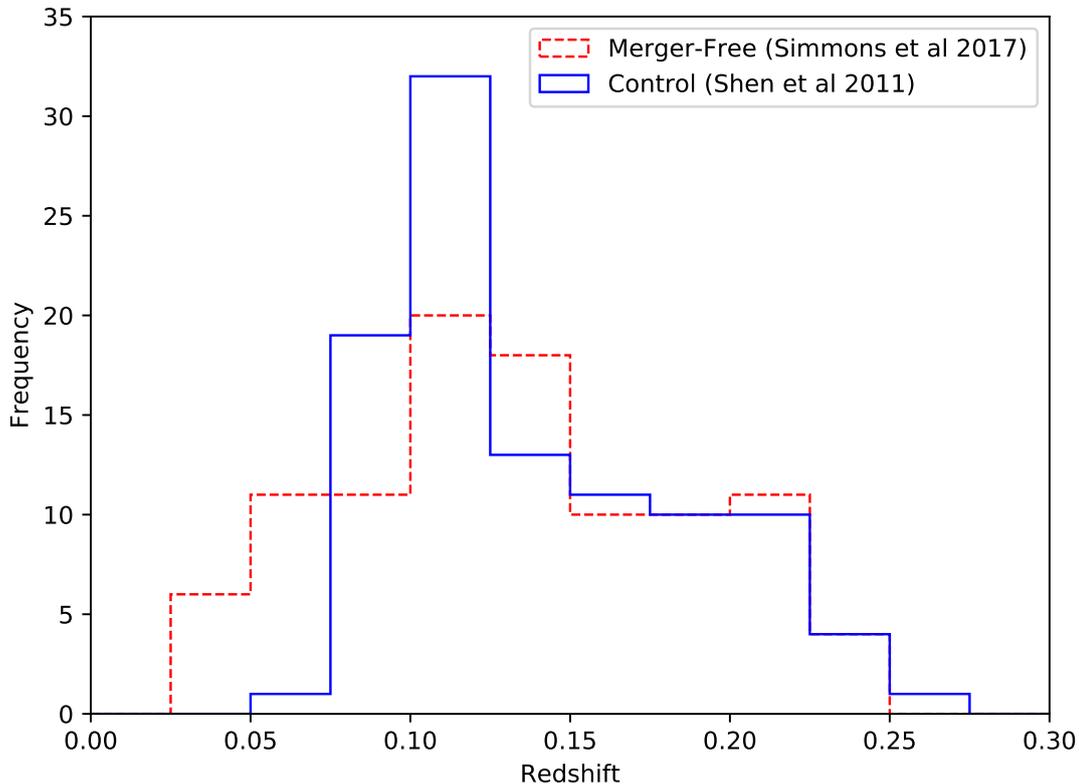
Figure 2.1: A frequency histogram showing the redshift distribution of the Simmons et al (2017) (merger-free) and control (a subset of Shen et al (2011)) samples of AGN host galaxies in terms of absolute numbers of galaxies at each redshift (in redshift bins of size 0.025)

A K-S statistical test (see Section 3.5 for a brief description of this method) comparing the redshift distributions of these samples gives a p-value of 0.102, which corresponds to a difference between distributions at the 1.6σ significance level. This indicates that there is a marginal difference in redshift distribution between the samples, which may be explained by overlap between the samples as previously discussed in this section, but the difference is not statistically significant, therefore the null hypothesis (that the redshift distributions of the two samples could be drawn from the same distribution) cannot be rejected. The two samples that we use may be therefore be considered to be from the same

evolutionary epoch.

The lowest redshift of any galaxy in the merger-free sample is $z = 0.0312$ (3 sf) and the highest redshift of any galaxy in the merger-free sample is $z = 0.244$ (3 sf). Applying the cosmological parameters $H_0 = 70$ kms^{-1}Mpc^{-1}, $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$, as described fully in Section 3.1, this gives a look-back time[1] of 0.426 Gyr for the lowest redshift galaxy in the merger-free sample and 2.899 Gyr for the highest redshift galaxy in the merger-free sample, thereby giving a look-back time difference of 2.473 Gyr between the galaxy with the lowest redshift and the galaxy with the highest redshift in the merger-free sample.

Similarly, the lowest redshift of any galaxy in the control sample was $z = 0.0645$ and the highest redshift of any galaxy in the control sample was $z = 0.253$. Applying the same cosmology as was applied to the merger-free sample in the previous paragraph and using the same method, the look-back time for the lowest redshift galaxy in the control sample is 0.862 Gyr and the look-back time for the highest redshift galaxy in the control sample is 2.987 Gyr. This gives a difference in look-back time of 2.125 Gyr over the entire control sample.

Table 2.1 summarises the results described in the previous two paragraphs.

Sample	Merger-Free	Control
Lowest Redshift	0.0312	0.0645
Highest Redshift	0.244	0.253
Minimum Look-back Time (Gyr)	0.426	0.862
Maximum Look-back Time (Gyr)	2.899	2.987
Difference in Look-back Time Over Entire Sample (Gyr)	2.473	2.125

Table 2.1: The lowest and highest redshifts of galaxies in the merger-free and control samples, along with their corresponding look-back times and the difference in look-back time between the highest and lowest redshift galaxies in each sample.

---

[1]All look-back times were calculated from cosmological parameters using the cosmological calculator developed by Wright et al (2006) [114].

2.1.4   Black Hole Mass and Bolometric Luminosity

Figure 2.2 shows the relationship between black hole mass and bolometric luminosity for both samples.
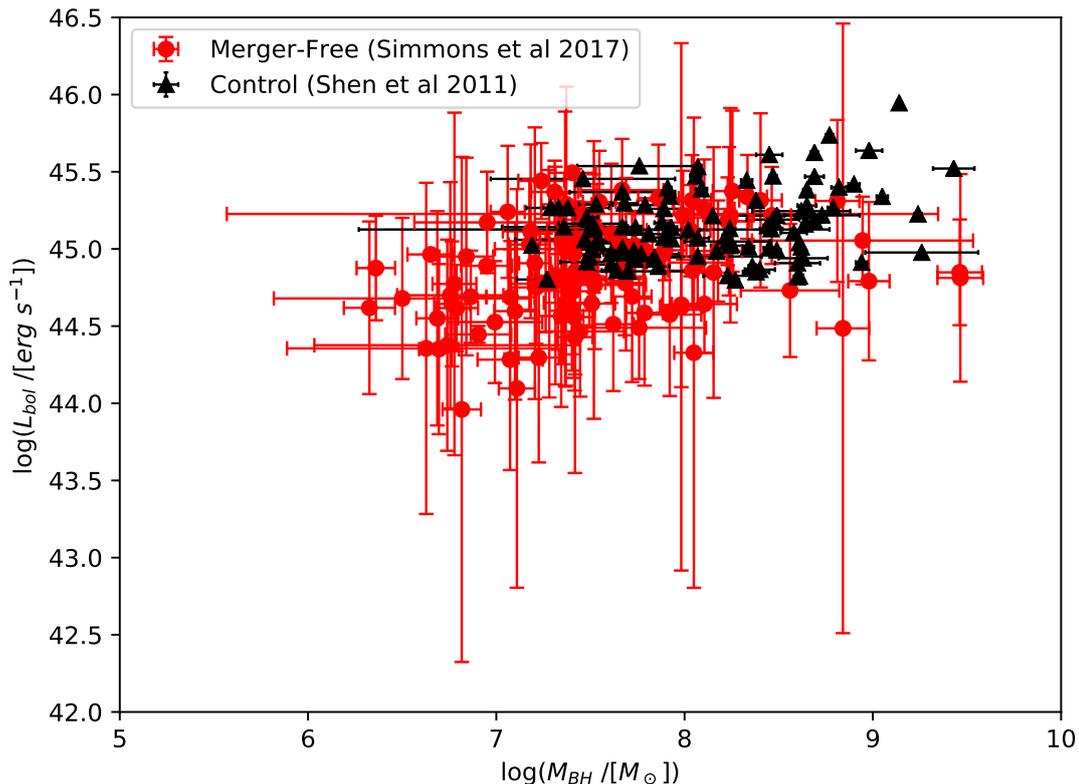
Figure 2.2: A logarithmic plot showing black hole mass (M_{BH}, x axis) plotted against bolometric luminosity (L_{bol}, y axis) for both samples of galaxies considered (Simmons et al (2017) and the 'control' subset of Shen et al (2011)).

Galaxies in the control sample have greater median black hole masses and greater median bolometric luminosities than those in the Simmons et al (2017) [55] sample. However, there is a large overlap between the samples on both axes. Whilst the error in some measurements may be large, the errors of both samples may be assumed to be heteroscedastic, and hence the median measurements of both black hole mass and luminosity are comparable to one another despite the presence of relatively large uncertainties. Therefore, the black holes in the control sample can be stated to be more luminous and more massive than black holes in

the merger-free sample, despite the large errors in black hole mass and bolometric luminosity. This could be an effect of the difference in redshift between the merger-free and control samples described in Section 2.1.3, since AGN at higher redshifts would require a greater luminosity to be detected than those at lower redshifts and due to the relationship between supermassive black hole mass and AGN luminosity, AGN with greater masses tend to also have greater luminosities. The galaxies in the control sample, on average, are at higher redshifts than those in the merger-free sample. Therefore, it would be plausible for galaxies in the control sample to have greater luminosities and black hole masses than galaxies in the merger-free sample for this reason.

2.2 Group Catalogue

2.2.1 The Yang et al Group Finders

We analyse the environment in which the samples of galaxies that we consider ([55],[53], see Sections 2.1.1 and 2.1.2) reside by considering the results of Yang et al (2007) group finder [115] when applied to the seventh data release of the New York University Value Added Galaxy Catalogue (NYU-VAGC DR7)[116] (based on the seventh data release of the Sloan Digital Sky Survey (SDSS DR7) [107],[110],[117]). For more information on our analysis see Section 3.3. In this section, we will discuss the properties of the Yang et al (2007) [115] group finder and its results when applied to NYU-VAGC DR7, as mentioned in [118].

The Yang et al (2007) group finder [115], hereafter Y07, is a slightly modified version of the Yang et al (2005) [119] group finder, hereafter Y05, which in turn improves upon on the 'friends of friends' (FOF) galaxy group finder algorithms that preceded it (for example, Davis et al (1985) [120]). 'Friends of friends' algorithms are those that link objects into groups in which, if a object is linked to one other objects in the group, it is also linked to all other objects in the group. Two objects in the a given group are not necessarily directly linked or connected to the same object, but will be associated via a 'chain' of connections between objects [121],[122]. In the case of galaxy group finders, this involves linking galaxies that are within a certain distance (defined by a 'linking length') of each other [121].

The key difference between Y05 and previous FOF algorithms is its consideration of dark matter haloes in the grouping of galaxies. This provides a physical basis upon which to group galaxies and, importantly for our purposes, to base claims that galaxies in the same group are likely to be undergoing gravitational interactions with one another, either directly or indirectly. The aim of Y05 is to group galaxies whose dark matter haloes overlap, and hence form one common halo; however, since dark matter haloes cannot be directly observed, this requires modelling the halo sizes and masses based on the luminosity of the galaxies that the haloes are associated with. The Y05 algorithm begins by using a 'traditional' FOF algorithm to link galaxies within 0.3 times the mean distance between galaxies in the sample being considered in the direction along the line of sight and 0.05

times the mean distance between galaxies in the sample being considered in the transverse direction, thereby obtaining initial group centres. Galaxies that are the brightest within the volume of a cylinder of radius 1 h^{-1} Mpc [1] and a depth (in terms of the relative velocities of galaxies around the sample galaxy) of ± 500 kms^{-1} around themselves are also considered as initial group centres. The total luminosity of each initial group is determined based on the luminosities of galaxies in the group. This group luminosity is then used to estimate the size, mass and velocity dispersion of the group (including the dark matter halo). This then allows more galaxies to be added to the groups based on the properties of the dark matter haloes associated with each group. The group centres are then recomputed and the process is repeated from the point of computing the total luminosity of each group until there are no changes in group membership between further iterations. Y05 also differs from previous FOF algorithms in that it allows groups to be of size one (with only one galaxy in the group) rather than requiring a group to consist of multiple galaxies.

The Y07 group finder differs from Y05 in that it considers all galaxies as potential group centres instead of only considering galaxies that are the brightest within a specific cylindrical volume. This means that every galaxy in the sample that the algorithm is run on will be assigned to a group, even if the group consists of only one galaxy.

The group catalogue we use is that obtained by applying the Y07 group finder to the New York University Value Added Galaxy Catalogue (NYU-VAGC) Data Release 7 [116], which is based on SDSS DR7 [107],[110],[117] as mentioned in Yang et al (2012) [118]. Specifically, we use the 'PetroA' galaxy and group catalogues. These result from the use of Petrosian [124] magnitudes and galaxies in 'Sample I' (as described in Y07) when running the group finder. Sample I is the sample of galaxies in the NYU-VAGC [116] with "reliable r-band magnitudes and measured redshifts from the SDSS"[115][2] and redshifts between 0.01 and 0.20 (inclusive).

---

[1] The distance this corresponds to depends on the cosmology applied. In Y07 the cosmological parameters used are those provided by the three-year data release of the WMAP mission [123], therefore $h = 0.73$ and the cylinder's radius is $\frac{1}{0.73}$Mpc (1.37 Mpc to three significant figures).

[2] Specifically, these galaxies have spectroscopic redshifts drawn from SDSS DR7 [110].

We acknowledge that Sample I suffers from incompleteness due to fibre collisions. As described in Yang et al (2007), "$\sim$ 7% of all galaxies eligible for spectroscopy do not have a measured redshift" [115] due to fibre collisions. This is likely to lead to some potential companion galaxies being excluded from the results of our analysis of the data from the Yang et al (2007) [115] group catalogue. Yang et al (2007) do provide a sample (sample III) in which corrections are made for fibre collisions by assuming that galaxies which are within 55 arcsecs of another galaxy (and therefore are affected by fibre collisions) are assigned the same redshift as the galaxy which they collide with, except for in cases where doing so would cause a galaxy to have an excessively large implied absolute magnitude. We do not use sample III in our analysis because, although the assumptions of Yang et al (2007) are reasonable and would eliminate the issue of galaxies being excluded from the data due to the lack of an available redshift, these fibre collision corrections would potentially introduce some error in the redshifts of potential companion galaxies, and hence in exact group sizes, since these redshifts are assumed rather than measured.

Chapter 3

Methods

3.1 Calculation of Relative Distances

Right ascension (RA) and declination (dec) co-ordinates and redshifts are available for both the sample galaxies and surrounding galaxies (see Chapter 2), therefore distances between galaxies can be derived from this information.

Initially, the angular diameter distance to each galaxy was calculated as a function of redshift, z, the Hubble constant, H_0, the mass density parameter of matter in the universe, Ω_m and the effective mass density parameter of dark energy in the universe, Ω_Λ using the following cosmology: $H_0 = 70$ kms^{-1}Mpc^{-1}, $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$— this cosmology is consistent to 2 significant figures and within the stated error of the results of the three-year data release of the Wilkinson Microwave Anisotropy Probe (WMAP) [123]. Spectroscopic redshifts were used for this calculation when available, otherwise photometric redshifts were used. Using photometric redshifts when spectroscopic redshifts were not available ensured greater completeness— at the cost of increased uncertainty— because few galaxies had spectroscopic redshifts available.

Both the photometric redshifts and the spectroscopic redshifts for potential companion galaxies were obtained from the thirteenth data release of the Sloan Digital Sky Survey (SDSS DR13)[107],[125], which also encompasses data from all previous Sloan Digital Sky Survey data releases. The photometric redshifts in

SDSS DR13 originated from SDSS DR12 [107],[126] and were obtained by Beck et al (2016) [127]. Beck et al (2016) employed a machine learning technique known as a kd-tree nearest neighbour fit (KF) in order to estimate redshifts for photometric observations in SDSS DR12 that have been tagged as galaxies, based on a training set of spectroscopic and photometric observations (including the entire SDSS DR12 spectroscopic catalogue [128],[129] and data from nine external surveys[1]), allowing a prediction of spectroscopic redshift to be made based on photometric observations for each galaxy.

Of the galaxies in the ninth data release of the Sloan Digital Sky Survey (SDSS DR9) [107],[109] that were within 10 arcmin of any galaxy in the merger-free sample [55], 3% had spectroscopic redshifts available, whereas 38% had photometric redshifts available. For the control sample [53], these were 4% and 36% respectively. For both samples, less than 1% of galaxies have both a spectroscopic redshift and a photometric redshift available, therefore in the merger-free [55] sample 41% of galaxies have a redshift available and in the control [53] sample 40% of galaxies have a redshift available— therefore only 40–41% of galaxies in each sample have a calculable three-dimensional distance. The information in this paragraph is summarised in Table 3.1.

---

[1]External surveys: 2dF [130],[131], 6dF [132],[133], DEEP2 [134],[135], GAMA [136],[137], PRIMUS [138],[139], VIPERS [140],[141], VVDS [142],[143], WiggleZ [144],[145] and zCOSMOS [146],[147].

Sample	Merger-Free	Control
Total number of galaxies within 10 arcmin of a sample galaxy	370268	383495
Number of galaxies within 10 arcmin of a sample galaxy with available photometric redshifts	141198	139370
Percentage of galaxies within 10 arcmin of a sample galaxy with available spectroscopic redshifts	12758	15943
Number of galaxies with both photometric and spectroscopic redshifts available	2470	2640
Total number of galaxies with available redshifts	151486	152643
Percentage of galaxies within 10 arcmin of a sample galaxy with available photometric redshifts (2sf)	38%	36%
Percentage of galaxies within 10 arcmin of a sample galaxy with available spectroscopic redshifts (2sf)	3.4%	4.2%
Percentage of galaxies with both photometric and spectroscopic redshifts available (2sf)	0.67%	0.69%
Total percentage of galaxies with available redshifts (2sf)	41%	40%

Table 3.1: A summary of the percentages of galaxies that were included in the ninth data release of the Sloan Digital Sky Survey that are within 10 arcmin of galaxies in the merger-free and control samples and also have photometric redshifts available from Beck et al (2016) and/or spectroscopic redshifts available from the thirteenth data release of the Sloan Digital Sky Survey.

Figures 3.1 and 3.2 compare the error in photometric redshift to the photometric redshift measurement for all galaxies within a 500 kpc projected distance of each sample, regardless of redshift. In both figures, it is clear that there is some dependence of the error in redshift on the redshift itself— greater redshifts have greater associated errors. The fact that the minimum error appears to also increase with redshift in both figures indicates that there may be a minimum percentage error in each photometric redshift measurement. The average error in photometric redshift is 0.0955 (3 sf) for the galaxies shown in Figure 3.1 and 0.0950 (3 sf) for galaxies shown in Figure 3.2, whilst the average redshifts of
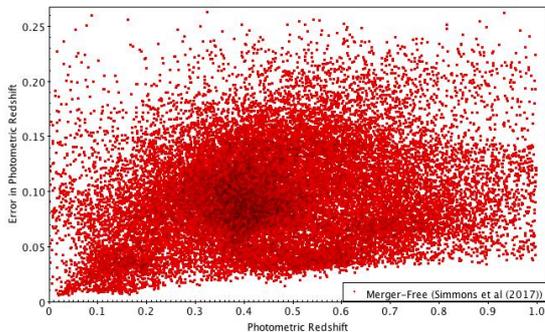
Figure 3.1: A comparison of the photometric redshifts for all galaxies within a 500 kpc projected distance of any galaxy in the merger-free sample (regardless of redshift) and the error in those redshifts.
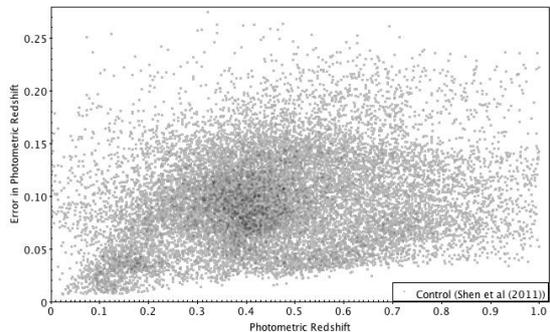
Figure 3.2: A comparison of the photometric redshifts for all galaxies within a 500 kpc projected distance of any galaxy in the control sample (regardless of redshift) and the error in those redshifts.

the same galaxies are 0.460 (3 sf) (Figure 3.1) and 0.451 (3 sf) (Figure 3.2) respectively. This indicates the photometric redshifts that we use have an error of $\sim 21\%$. In contrast to the spectroscopic redshifts we use, whose average errors are of the order 10^{-5}, these errors are very large and therefore spectroscopic redshifts would be preferable. However, as previously mentioned, our use of photometric redshifts is necessary for completeness.

Spectroscopic redshifts were available for all of the galaxies in both the control and merger-free samples because a spectrum had been taken for each galaxy as part of the sampling process [55],[53], therefore the angular diameter distance for these galaxies was always calculated using a spectroscopic redshift.

The standard trigonometric formulae for calculating the projected separation between two points on a projected sphere (in this case the sample galaxy and its neighbour), as shown in Equations 3.1 and 3.3, were used. In this case, the average declination of the two galaxies was used when calculating the RA distance (see Equation 3.2).

$$\Delta d_{\mathrm{RA}} = |d_z \tan((\mathrm{RA}_1 - \mathrm{RA}_2)\cos(\mathrm{dec}_{av}))| \tag{3.1}$$

$$\mathrm{dec}_{\mathrm{av}} = \frac{\mathrm{dec}_1 + \mathrm{dec}_2}{2} \tag{3.2}$$

$$\Delta d_{\mathrm{dec}} = |d_z \tan(\mathrm{dec}_1 - \mathrm{dec}_2)| \tag{3.3}$$

where Δd_{RA} represents the projected RA distance between a sample galaxy and its neighbour, RA_1 represents the RA co-ordinate of a sample galaxy, RA_2 represents the RA co-ordinate of the same sample galaxy's neighbouring galaxy, d_z represents the angular diameter distance to the sample galaxy, dec_1 represents the dec co-ordinate of a sample galaxy and dec_2 represents the dec co-ordinate of the same sample galaxy's neighbouring galaxy.

The overall three-dimensional distance between galaxies was then calculated using the three-dimensional pythagorean theorem shown in Equation 3.4.

$$d_{\mathrm{overall}} = \sqrt{(\Delta d_z)^2 + (\Delta d_{\mathrm{RA}})^2 + (\Delta d_{\mathrm{dec}})^2} \tag{3.4}$$

where d_{overall} represents the overall three-dimensional distance between a sample galaxy and its neighbouring galaxy, Δd_z represents the absolute value of the difference between the angular diameter distance to the sample galaxy and the angular diameter distance to the neighbouring galaxy, Δd_{RA} represents the RA distance between the sample galaxy and its neighbouring galaxy and Δd_{dec} represents the dec distance between the sample galaxy and its neighbouring galaxy.

3.2 Cylindrical Search for Potential Companion Galaxies

One method by which we are able to quantitatively describe the environment in which galaxies reside is by counting neighbouring galaxies within a defined search area. In order to achieve this, we use a cylindrical search area with the sample galaxy in the centre as shown in Figure 3.3.
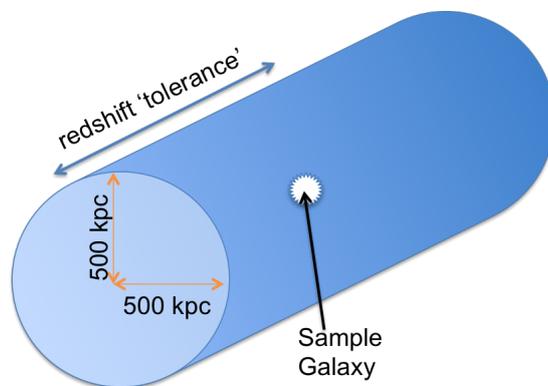
Figure 3.3: A labelled diagram showing the volume around each sample galaxy covered by the cylinder search.

The radius of the cylinder is 500 kpc at the redshift of the sample galaxy. This was achieved by performing a search for galaxies in the ninth data release of the Sloan Digital Sky Survey [107],[109], searching within the angular radius corresponding to a 500 kpc projected distance from the sample galaxy at the redshift of the sample galaxy.

An angular diameter 'redshift distance' between each galaxy found in the search and its corresponding sample galaxy was then calculated in the manner described in Section 3.1. The 'redshift distance' was used to determine the length of the cylindrical search area. Three 'redshift distance tolerances' were considered— 500 kpc, 1 Mpc and 5 Mpc.

3.2.1 Free Fall Time

The free fall time is the time that would be taken for a small test particle to fall from one point to another given that only gravitational forces act upon the particle and the particle has an initial velocity of zero. In the case of minor gravitational interactions between galaxies, this represents the approximate time elapsed before one galaxy could have any affect on another when the galaxies are separated by distance R.

The free fall time at each of the distances considered in the cylinder search was calculated using Equation 3.5 [1].

$$t_{ff} = \sqrt{\frac{2R^3}{GM}} \tag{3.5}$$

Where, in this case, t_{ff} represents the free fall time, R represents the distance between the centre of the sample galaxy (the location of the AGN) and the edge another galaxy, G represents Newton's gravitational constant and M represents the total mass in the area between the two galaxy centres. We take M to be the mass of the Milky Way galaxy, which is similar to the sample galaxies we consider, and assume a vacuum in the space between the two galaxies. Given these assumptions, when $R = 500$ kpc, $t_{ff} = 7.4 \times 10^9$ yr. This is $0.53t_H$ (2 s.f.), where t_H represents the Hubble time (14×10^9 yr, approximately the age of the universe). It is therefore reasonable to conclude that galaxies at a distance of 500 kpc from one another would be exert a gravitational force over one another and would therefore be interacting since enough time has elapsed for such interactions to have occurred and to have had affected the centre of the sample galaxy.

It is possible to calculate the value of R at a given value of t_{ff} by rearranging Equation 3.5 to give Equation 3.6.

$$R = \sqrt[3]{\frac{t_{ff}^2 GM}{2}} \tag{3.6}$$

---

[1]See Appendix A for a derivation of this. We assume Newton's Law of Gravitation and Newton's Laws of Motion [148] are valid.

where all definitions are as in Equation 3.5. From this, we find that when $t_{ff} = t_H$, $R = 750$ kpc. Therefore, galaxies at distances of over 750 kpc from the sample galaxy will not have directly gravitationally interacted with the sample galaxy.

However, it should be considered that redshift distances between nearby galaxies may not be indicative the true distances between these galaxies because potential companion galaxies may have peculiar velocities, which are not considered in the calculation of an angular diameter distance as this assumes a galaxy's redshift only depends on the expansion of the universe.

$$v_H = H_0 d \tag{3.7}$$

where v_H represents the Hubble velocity (velocity due to the expansion of the universe) of a galaxy, H_0 represents the Hubble constant and d represents the distance along the line of sight- in this case the 'redshift distance'. Considering only the expansion of the universe and applying Equation 3.7 as an approximation of the calculation used to determine the 'redshift distance' and $H_0 = 70$ kms^{-1}Mpc^{-1} as was used in when calculating 'redshift distance' (see Section 3.1), a 'redshift distance' of 1 Mpc corresponds to a total velocity of 70 kms^{-1}. The same calculation performed for a 'redshift distance' of 750 kpc (the free fall distance corresponding to the age of the universe) gives a Hubble velocity of 52.5 kms^{-1}. Since a galaxy's total velocity is the sum of its Hubble velocity and its peculiar velocity, a galaxy at a distance of 750 kpc from another and a peculiar velocity of 17.5 kms^{-1} would appear to have a 'redshift distance' of 1 Mpc from the other galaxy.

There is evidence from previous studies that the large-scale environment of galaxies is relevant to their evolution (an effect known as conformity) [149],[150]. Therefore, we also continue to include the results of the cylinder search using a 5 Mpc 'redshift distance' both for completeness and to provide a measurement of the wider environment in which the sample galaxies reside despite many of the galaxies identified in this search being unable to directly interact with the sample galaxies.

3.2.2 Redshift Availability in the Cylinder Search and the Potential Effect of Fibre Collisions

At the highest redshift of any galaxy in either of the merger-free or control samples ($z = 0.253$), 500 kpc corresponds to an angular distance of 2.10 arcmin (3 sf)[1]. At this angular scale, fibre collisions are unlikely to affect our results [151]. However, since our study intended to detect galaxies within distances of less than 500 kpc from other galaxies, in cases where galaxies were at angular separations of less than 1 arcmin [151] from one another (and especially in cases where galaxies were separated by an angle of less than the SDSS fibre collision limit of 55 arcsec [107]), fibre collisions may have limited the availability of spectroscopic detections in the SDSS.

Similarly, at the lowest redshift of any galaxy in either of the merger-free or control samples ($z = 0.0312$), 500 kpc corresponds to an angular separation of 12.9 arcmin (3sf), a scale at which fibre collisions should not affect results. The same caveat— that galaxies separated by less than 1 arcmin, and especially galaxies separated by less than 55 arcsec, from another galaxy may not be spectroscopically detected in the SDSS— applies at all redshifts, but it is less likely that this will occur at lower redshifts since all distances correspond to larger angular scales than at higher redshifts.

This means that fibre collisions may partially account for the lack of availability of redshifts for all galaxies. Since this cylinder search includes photometric (imaging) detections of galaxies as well as spectroscopic detection, this effect does not necessarily extend to a lack of detections of potential companion galaxies at angular separations of less than 1 arcmin, but may lead to these galaxies not having available redshift measurements. If a potential galaxy does not have an available redshift measurement (either spectroscopic or photometric), then its three-dimensional distance from the galaxies in the merger-free and control samples cannot be calculated and therefore the galaxy cannot be presented in the results of the cylinder search. Especially for the highest redshift galaxies in the

---

[1]Applying the same cosmological parameters that we used to calculate the 'redshift distances' of potential companion galaxies in Section 3.1 ($H_0 = 70$ kms^{-1}Mpc^{-1}, $\Omega_m = 0.27$ and $\Omega_\Lambda = 0.73$).

merger-free and control samples, this could lead to an artificial scarcity of galaxies at small separations from the sample galaxies in the results of the cylinder search. This is discussed further in Section 4.1. If a potential companion galaxy has a photometric redshift measurement available but not a spectroscopic redshift measurement, then the galaxy will be included in our results, but the error in its redshift may be $\sim 21\%$, as discussed in Section 3.1, which would mean that its three-dimensional distance had a greater uncertainty than that of galaxies whose three-dimensional distances were calculated using spectroscopic redshifts.

Table 3.2 shows the availability of redshifts for all galaxies within a 500 kpc projected distance of any galaxy in the merger-free or control samples. 60–70% of these galaxies do not have an available redshift measurement. This may be due to fibre collisions or due to the galaxies appearing as very faint in imagery. It is likely that many of these galaxies are background galaxies, with much higher redshifts than those of the merger-free and control samples. However, without any measured redshift for these galaxies, it is not possible to state for certain whether or not these galaxies could be potential companions to galaxies in the merger-free or control samples.

Tables 3.3, 3.4 and 3.5 show the numbers and percentages of galaxies within each 'redshift distance tolerance' for which spectroscopic or photometric redshifts were used to calculate their three-dimensional distances. As described in Section 3.1, spectroscopic redshifts were preferred over photometric redshifts when calculating three-dimensional distances. For all tolerances, a greater percentage of potential companion galaxies to the merger-free sample had spectroscopic redshifts available than for the control sample. This may be related to the control sample having a greater average redshift than the merger-free sample (see Section 2.1). This would have given potential companion galaxies to the control sample a greater likelihood of being affected by fibre collisions and therefore could have limited the availability of spectroscopic redshifts for potential companion galaxies to the control sample.

Sample	Merger-Free	Control
Total Number of Galaxies	73696	60475
Number of Galaxies that do not have an available redshift measurement	44013	42557
Number of Galaxies that have Available Spectroscopic Redshifts	152	76
Number of Galaxies that have Available Photometric Redshifts	29683	17917
Number of Galaxies that have an Available Spectroscopic Redshift and an Available Photometric Redshift	152	75
Percentage of Galaxies that do not have an available redshift measurement (2 sf)	60%	70%
Percentage of Galaxies that have Available Spectroscopic Redshifts	0.21%	0.13%
Percentage of Galaxies that have Available Photometric Redshifts	40%	30%
Percentage of Galaxies that have an Available Spectroscopic Redshift and an Available Photometric Redshift	0.21%	0.12%

Table 3.2: The numbers and percentages of all galaxies within a 500 kpc projected distance of any galaxy the merger-free or control samples that have (and that do not have) spectroscopic and/or photometric redshifts available from SDSS DR13. This includes background galaxies with redshifts that are much greater that those of the merger-free and control samples.

Sample	Merger-Free	Control
Total Number of Galaxies	12	7
Number of Galaxies whose distance was calculated using a spectroscopic redshift	8	3
Number of Galaxies whose distance was calculated using a photometric redshift	4	4
Percentage of Galaxies whose distance was calculated using a spectroscopic redshift	67%	43%
Percentage of Galaxies whose distance was calculated using a photometric redshift	33%	57%

Table 3.3: The numbers and percentages of all galaxies within a 500 kpc projected distance and within a 500 kpc 'redshift distance' of any galaxy the merger-free or control samples whose three-dimensional distances were calculated using spectroscopic and/or photometric redshifts available from SDSS DR13.

Sample	Merger-Free	Control
Total Number of Galaxies	17	16
Number of Galaxies whose distance was calculated using a spectroscopic redshift	11	7
Number of Galaxies whose distance was calculated using a photometric redshift	6	9
Percentage of Galaxies whose distance was calculated using a spectroscopic redshift	65%	44%
Percentage of Galaxies whose distance was calculated using a photometric redshift	35%	56%

Table 3.4: The numbers and percentages of all galaxies within a 500 kpc projected distance and within a 1 Mpc 'redshift distance' of any galaxy the merger-free or control samples whose three-dimensional distances were calculated using spectroscopic and/or photometric redshifts available from SDSS DR13.

Sample	Merger-Free	Control
Total Number of Galaxies	51	61
Number of Galaxies whose distance was calculated using a spectroscopic redshift	28	20
Number of Galaxies whose distance was calculated using a photometric redshift	23	41
Percentage of Galaxies whose distance was calculated using a spectroscopic redshift	55%	33%
Percentage of Galaxies whose distance was calculated using a photometric redshift	45%	67%

Table 3.5: The numbers and percentages of all galaxies within a 500 kpc projected distance and within a 5 Mpc 'redshift distance' of any galaxy the merger-free or control samples whose three-dimensional distances were calculated using spectroscopic and/or photometric redshifts available from SDSS DR13.

3.3 Analysis of Data from a Group Catalogue

We apply the analysis techniques described in this section to the data described in Section 2.2.1.

Initially, we matched the data in the galaxy catalogue (iPetroA as described in Section 2.2.1) to the single group catalogue detailing the galaxies contained in each group by matching catalogue entries with the same object ID in the NYU-VAGC DR7 catalogue [116]. This produced a single table for each sample (a total of two tables), containing enough information to place each of the sample galaxies into groups and find any galaxies that this group catalogue deems to be gravitationally interacting with the sample galaxies. We then cross-matched the resulting tables to the sample galaxy data for both samples described in Sections 2.1.1 and 2.1.2 based on the right ascension and declination (RA and dec) of each galaxy, with a tolerance of 1 arcsec, in order to locate any sample galaxies included in the iPetroA galaxy catalogue data. This provided 75 matches (74% of the sample) for the merger-free sample and 53 matches (52% of the sample) for the control sample.

Following this, we searched for all galaxies in the group catalogue that had the same group identifier as any of the sample galaxies, and therefore would be in the same group as at least one of the sample galaxies. This produced another table for each sample containing the sample galaxies and the galaxies found using this search. These tables contained 217 galaxies for the merger-free sample and 115 galaxies for the control sample. A further table was created by matching the tables containing both sample galaxies and galaxies in groups containing sample galaxies to the sample data, using RA and dec co-ordinates with a tolerance of 1 arcsec, and removing the sample galaxies to produce tables containing only galaxies in the same group as sample galaxies and not the sample galaxies themselves. This would be of use when comparing the distances from the sample galaxies at which such 'companion' galaxies reside (see Section 4.2). These tables contained 142 galaxies for the merger-free sample and 62 galaxies for the control sample. The information in the previous two paragraphs is summarised in Table 3.6.

Sample	Merger-Free	Control
Total Number of Galaxies in Sample	101	101
Number of Galaxies in the Sample that are also in the Yang et al (2007) Group Catalogue	75	53
Percentage of Galaxies in the Sample that are also in the Yang et al (2007) Group Catalogue	74%	52%
Total Number of Galaxies that are in the same group as a Galaxy in the Sample (including the sample galaxies)	217	115
Number of Potential Companion Galaxies	142	62

Table 3.6: A table summarising the results of matching the merger-free and control sample tables of galaxies to the Yang et al (2007) group catalogue.

The tables produced from this analysis were used to produce the plots and results shown and discussed in Section 4.2.

3.4   Environment Coefficients

In order to gain a single-number measure of the density of environment in which each galaxy resides, we make use of the method described in [152]. An environment coefficient, σ, was calculated for each galaxy according to Equations 3.8 and 3.9.

$$\sigma = \frac{\log_{10}(\Sigma_4) + \log_{10}(\Sigma_5)}{2} \tag{3.8}$$

$$\Sigma_N = \frac{N}{\pi d_N^2} \tag{3.9}$$

where N describes the order of galaxies in terms of distance from the sample galaxy, with the nearest galaxy having $N = 1$, the second nearest having $N = 2$, etc. In this case, the fourth and fifth nearest galaxies to the sample galaxy are used. d_N represents the three-dimensional distance between the sample galaxy and the Nth galaxy.

We perform this calculation for all galaxies in each of the samples that we consider, using the five nearest galaxies to each sample galaxy in the ninth data release of the Sloan Digital Sky Survey [107],[109] based on three-dimensional distances calculated using the method described in Section 3.1. Since, as noted in Section 3.1, not all galaxies have available redshifts and we are unable to calculate three-dimensional distances for those that do not have redshifts available, we are only able to consider galaxies for which redshifts (either spectroscopic or photometric) are available when calculating environment coefficients. This is 40–41% of all galaxies within 10 arcmin of galaxies in the merger-free and control samples[1]. This may affect the result of the calculation for some galaxies, since their nearest five neighbouring galaxies may include those for which we do not have a measured redshift. Also, since most galaxies do not have available spectroscopic redshifts, more uncertainty in our calculation of environment coefficients may be introduced by using photometric redshifts for these galaxies in our calculation of their three-dimensional distance from a sample galaxy; however, this impacts

---

[1]For a breakdown of the exact numerical and percentage availability of redshifts for all galaxies within 10 arcmin (projected distance) of galaxies in both the merger-free and control samples, which includes all galaxies used to calculate environment coefficients, see Table 3.1.

the results of our calculation of environment coefficients far less than the possible effect of only considering galaxies with available spectroscopic redshifts due to the lack of available spectroscopic redshifts.

3.5 Statistical Testing

3.5.1 Kolmogorov-Smirnov (K-S) Test

We make use of the Kolmogorov-Smirnov (K-S) [153],[154],[155] method of statistical testing in order to assess the level of significance of our results. When applied in two dimensions [156], this method compares two distributions of results by shape and determines the likelihood of the results having been drawn from the same distribution. We choose to use this test since we are comparing two distributions (those of the two galaxy samples and their associated environments) and we wish to determine whether or not the distributions could be drawn from different functions (e.g. different environments, different redshift distributions).

3.5.2 Pearson Correlation Coefficient

The Pearson product-moment (sometimes referred to as Pearson 'r') correlation coefficient [157],[158] is a quantity which describes the strength and direction (positive or negative) of linear regression between two continuous variables. A correlation coefficient of $r = 0$ would indicate that no linear correlation was present, whereas correlation coefficients of $r = 1$ or $r = -1$ would indicate that a line of best fit drawn on a scatter plot comparing the two variables would pass through every data point, with $r = 1$ corresponding to a perfect positive correlation and $r = -1$ corresponding to a perfect negative correlation. This can be used to easily determine whether or not linear correlation is present. We apply this method in Section 4.4.4 to quantitatively determine whether a clear correlation between redshift and environment coefficient is present.

Chapter 4

Results and Discussion

4.1 Cylinder Search

As discussed in Section 3.2, we performed a cylinder search for galaxies around the sample galaxies, applying three different 'redshift distance tolerances' and thereby producing three datasets containing galaxies found in each of the three searches. We obtained data for redshift distance tolerances of 500 kpc, 1 Mpc and 5 Mpc. It should be noted that these datasets are not mutually exclusive; by definition, the entirety of the 500 kpc dataset is contained within the 1 Mpc dataset and the entirety of the 1 Mpc dataset (and therefore also the 500 kpc dataset) is contained within the 5 Mpc dataset.

4.1.1 Cumulative Histograms Showing Projected Distances Between Sample Galaxies and Companions

Figures 4.1, 4.2 and 4.3 are cumulative frequency histograms showing the distances of potential companion galaxies from their associated sample galaxy. The histograms show the results of cylinder searches conducted using 'redshift tolerances' of 500 kpc (Figure 4.1), 1 Mpc (Figure 4.2) and 5 Mpc (Figure 4.3).
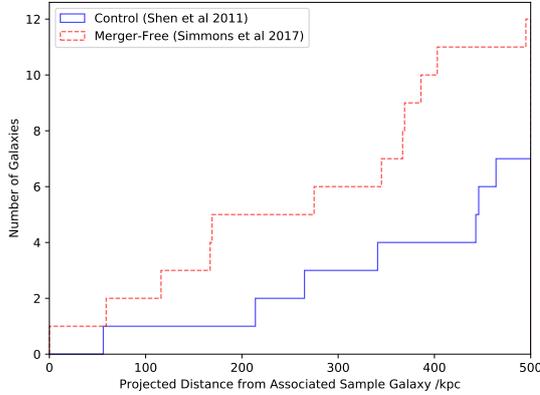
Figure 4.1: A cumulative histogram showing the results of a cylinder search for galaxies that are within 500 kpc projected distance of any sample galaxy and within 500 kpc 'redshift distance' of any sample galaxy. The projected distance (not considering redshift) between each galaxy found in the search and its associated sample galaxy (x) is plotted against cumulative frequency (y). A bin size of 1 kpc is used.
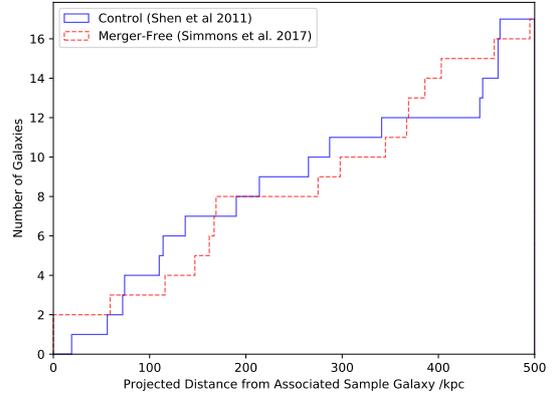
Figure 4.2: A cumulative histogram showing the results of a cylinder search for galaxies that are within 500 kpc projected distance of any sample galaxy and within 1 Mpc 'redshift distance' of any sample galaxy. The projected distance (not considering redshift) between each galaxy found in the search and its associated sample galaxy (x) is plotted against cumulative frequency (y). A bin size of 1 kpc is used.

From Figures 4.1, 4.2 and 4.3, it appears that for the cylinder search with a 'redshift distance tolerance' of 500 kpc, a greater number of potential companions are found for the merger-free sample than for the control sample, however the inverse is shown in the histogram for the 5 Mpc 'redshift tolerance' search— the control sample appears to have a greater number of potential companions than the merger-free sample at projected distances greater than 50 kpc from a sample galaxy.

However, when the distributions are compared using a K-S test, the 500 kpc 'redshift distance tolerance' search results (Figure 4.1) differ between samples with a p-value of 0.568 (corresponding to the 0.6σ significance level); the 1 Mpc 'redshift distance tolerance' search results (Figure 4.2) differ between samples with a p-value of 0.930 (corresponding to the 0.1σ significance level) and the 5 Mpc 'redshift distance tolerance' search results (Figure 4.3) differ between samples with a p-value of 0.0363 (corresponding to the 2.1σ significance level). This
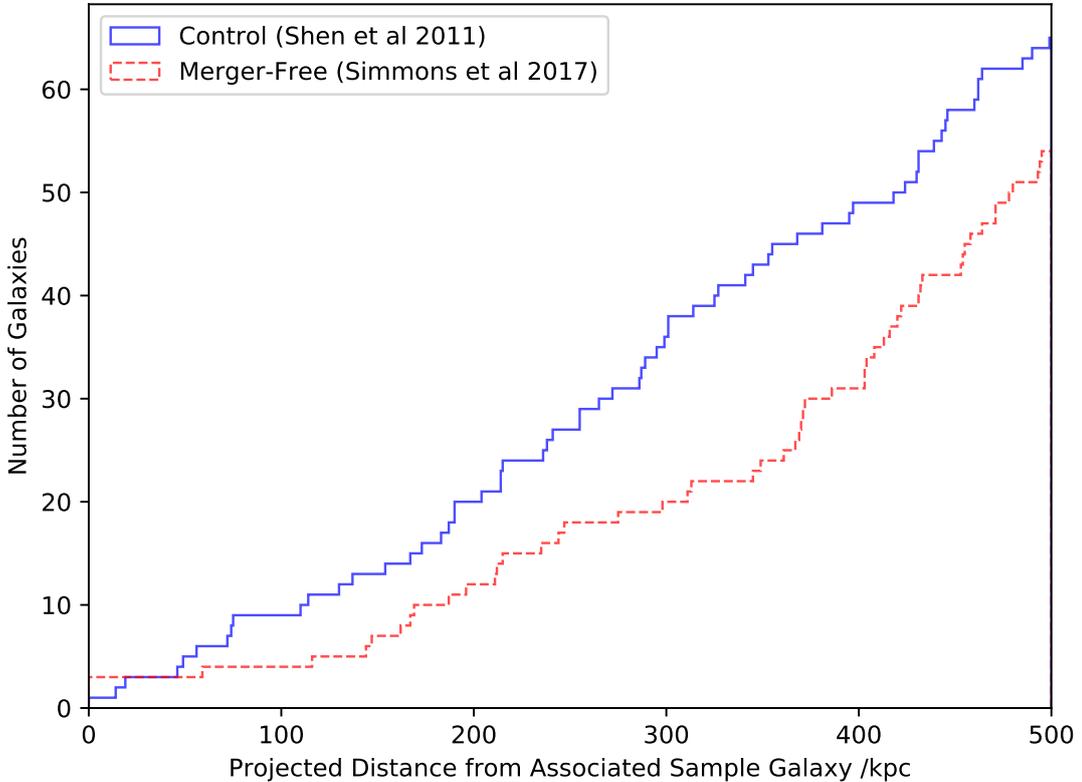
Figure 4.3: A cumulative histogram showing the results of a cylinder search for galaxies that are within 500 kpc projected distance of any sample galaxy and within 5 Mpc 'redshift distance' of any sample galaxy. The projected distance (not considering redshift) between each galaxy found in the search and its associated sample galaxy (x) is plotted against cumulative frequency (y). A bin size of 1 kpc is used.

suggests that there is no statistically significant difference in distribution of distances of potential companion galaxies between the merger-free and control samples when a cylinder search is conducted using a 500 kpc or a 1 Mpc 'redshift distance tolerance' and the distributions differ with marginal significance when a cylinder search is conducted using a 5 Mpc 'redshift distance tolerance'. Therefore, the null hypothesis that the distributions of distances of potential companion galaxies from their associated sample galaxies do not differ between the two samples that we consider cannot be rejected for any of the cylinder searches that we conducted.

Since the distribution of distances of galaxies from the sample galaxies is a measure of the environment in which the sample galaxies reside, the fact that we cannot reject the null hypothesis in this case indicates that the environments in which both samples reside, in terms of density of galaxies, are similar. However, the marginally significant difference in distributions for the cylinder search with a 5 Mpc 'redshift tolerance' may be indicative of a difference in environment either at further distances from the sample galaxies or in terms of galaxies with high peculiar velocities. This is discussed further in Section 4.1.2.

4.1.2 Non-Cumulative Histograms Showing Projected Distances Between Sample Galaxies and Companions

Through the use of non-cumulative histograms showing the projected distances of potential companion galaxies from their associated sample galaxies (Figures 4.4, 4.5 and 4.6), we are able to more intuitively compare environments in terms of the projected distance of potential companion galaxies from the sample galaxies and compare these between the two samples that we consider and between the three 'redshift distance tolerances' that we consider.

Since the same distributions of distances are being compared, the same statistical results apply as in Section 4.1.1; there is no statistically significant difference between the distributions of distances for the two samples we consider for the cylinder searches applying 'redshift distance tolerances' of 500 kpc and 1 Mpc (Figures 4.4 and 4.5). In the case of the cylinder search with a 500 kpc 'redshift distance tolerance', this may be due to the total number of potential companions found being too small to give a significant result. The difference in distributions for the cylinder search applying a 'redshift distance tolerance' of 5 Mpc (Figure 4.6) is of marginal significance. This difference is visible in Figure 4.6, as the peak frequency for the histogram for the merger-free sample occurs in the 450–500 kpc projected distance bin and a positive correlation is displayed between projected distance from a sample galaxy and frequency (number of galaxies at each projected distance). In contrast, the control sample, whilst displaying a slightly increased frequency of galaxies at projected distances greater than 150
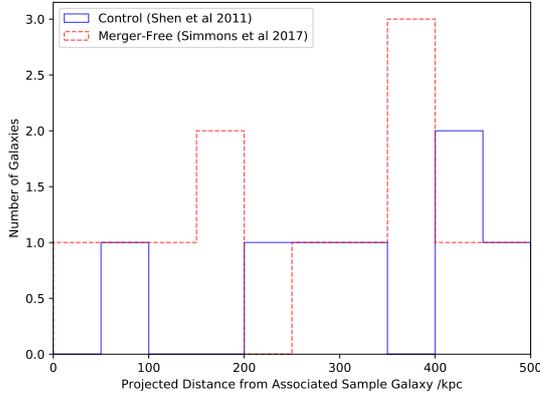
Figure 4.4: A non-cumulative histogram showing the distribution of distances of potential companion galaxies from their associated sample galaxies for galaxies within a 500 kpc 'redshift distance' of their associated sample galaxy and within a 500 kpc projected distance around their sample galaxy. We use distance bins of size 50 kpc.
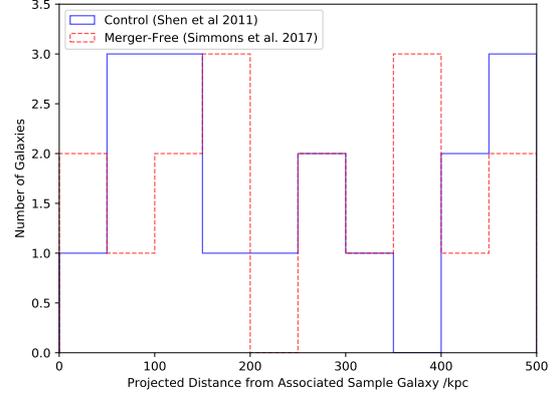
Figure 4.5: A non-cumulative histogram showing the distribution of distances of potential companion galaxies from their associated sample galaxies for galaxies within a 1 Mpc 'redshift distance' of their associated sample galaxy and within a 500 kpc projected distance around their sample galaxy. We use distance bins of size 50 kpc.

kpc, remains generally level, showing little correlation between projected distance from a sample galaxy and frequency. This is notable because this occurs only for the 5 Mpc 'redshift distance tolerance' search. As previously mentioned in Section 4.1.1, this could be interpreted to indicate a presence of galaxies with high peculiar velocities relative to the sample galaxies that only occurs for the merger-free sample. These galaxies may be travelling at sufficiently high velocities relative to the sample galaxies that they could pass a sample galaxy at a projected distance of less than 500 kpc and not become bound in orbit. We refer to such galaxies as 'fly-by' galaxies. Such an interaction may transfer enough kinetic energy to gas within a sample galaxy to affect the growth of its supermassive black hole.

The distributions shown throughout this section may also be subject to some biases due to the effect of fibre collisions or a lack of available redshifts for faint galaxies. As discussed in Section 3.2.2, fibre collisions and faint galaxies could lead to the decreased availability of redshifts for potential companion galaxies at higher redshifts, which could cause less galaxies to appear in the results of the
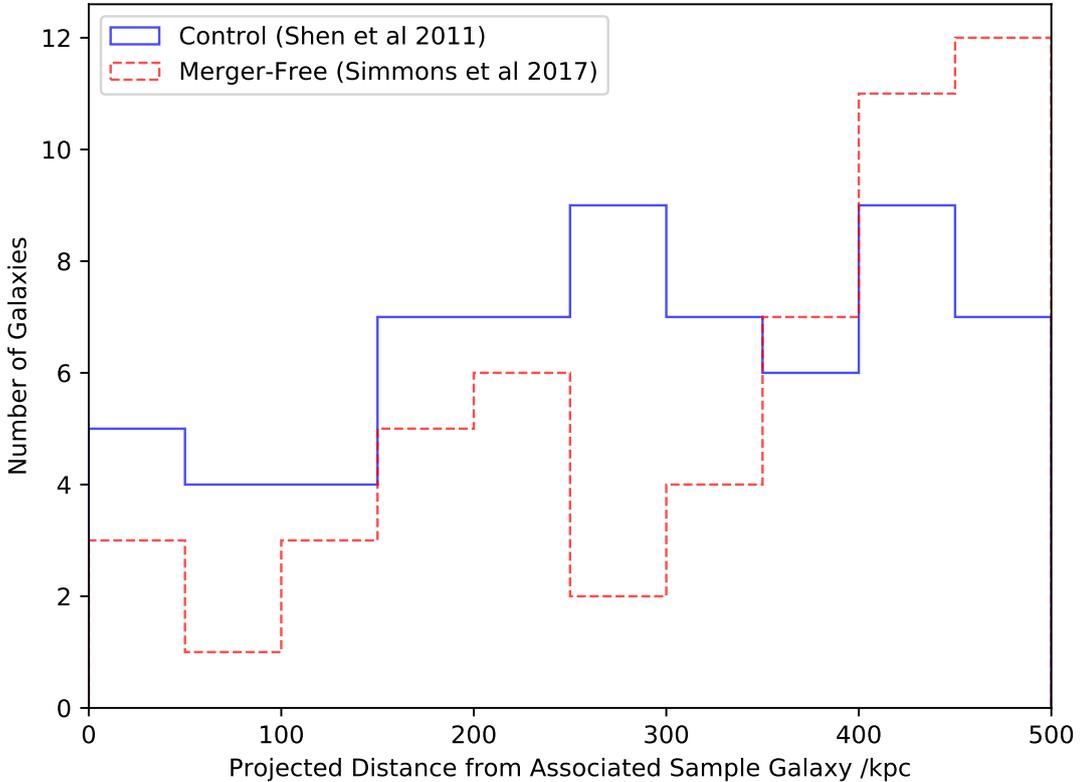
Figure 4.6: A non-cumulative histogram showing the distribution of distances of potential companion galaxies from their associated sample galaxies for galaxies within a 5 Mpc 'redshift distance' of their associated sample galaxy and within a 500 kpc projected distance around their sample galaxy. We use distance bins of size 50 kpc.

cylinder search at smaller projected distances from the sample galaxies, as is seen for both samples in Figure 4.6. This would be expected to have a greater effect on the control sample than the merger-free sample because the control sample has a greater average redshift than the merger-free sample (see Section 2.1.3). At the highest redshift of any galaxy in the control sample ($z = 0.253$), 1 arcmin (the angular separation at which fibre collisions become relevant [151]), corresponds to 238 kpc (3 sf). Therefore, the results of the cylinder search for potential companion galaxies to the control sample may have been affected by fibre collisions at any projected distance less than 238 kpc. Similarly, at the highest redshift of any galaxy in the merger-free sample ($z = 0.244$), 1 arcmin corresponds to 232

kpc. Therefore, the results of the cylinder search for potential companion galaxies to the merger-free sample at any projected distance less than 232 kpc may have been affected by fibre collisions. However, this would not fully explain the differences between distributions seen in Figure 4.6 as the control sample shows a greater number of companion galaxies than the merger-free sample at these low projected distances, which is the opposite of what would be expected if the effect occurred due to fibre collisions, and there are differences between the distributions at distances greater than 238 kpc, which are not affected by fibre collisions due to small separations between the sample galaxies and potential companions.

4.1.3   Potential Error in Distributions

The distributions shown throughout Section 4.1 are discrete distributions created by counting galaxies in the areas around the merger-free and control samples. If it is assumed that galaxies occur randomly and independently from one another (not in clusters, for example) in the areas around the merger-free and control samples, then the distributions shown throughout Section 4.1 may be modelled as Poisson distributions.

If it is assumed that the error in these distributions occurs independently and randomly throughout the distributions, then the error in these distributions may be modelled as Poisson variation [159]. Therefore, the error in these distributions may be calculated as:

$$\sqrt{N} \tag{4.1}$$

where N represents the number of counts (in this case galaxies) in a particular bin [159]. $\sqrt{N}$ may be added or subtracted from the measured count to give the theoretical maximum or minimum true value of the count.

In the case of the distributions that result from the cylinder searches that we conducted with 'redshift distance tolerances' of 500 kpc and 1 Mpc, these errors are very large compared to the counts themselves. For example, it the maximum number of galaxies in any bin shown in Figures 4.4 and 4.5 is 3. The error in this measurement would therefore be $\sqrt{3}$ (1.73 (3 sf)). This means that the theoretical true value of the count has a minimum value of 1.27 (3sf) and a maximum value of 4.73 (3sf). Since the true value of counts must be discrete, the true value may be any value in the range 1–5. In the case of these distributions, such a difference could potentially mean that the overall shape of the distributions may theoretically be very different from that which we have observed.

Although these error bars are not explicitly shown on the plots in Section 4.1 because their large relative size would mean that they dominated the plots, the fact that the error in these distributions is large compared to the numbers of counts is considered when performing K-S tests (see Section 3.5 for more information about these) to compare the shapes of the distributions, since it is less likely that distributions will vary significantly according to a K-S test if the overall sample size is small. This occurs because when the cumulative distributions

are normalised, as occurs during the K-S testing process, a single count (of one galaxy) would have a larger impact on the shape of a normalised distribution if the total sample size were smaller.

4.2   Analysis of Data from a Group Catalogue

We analyse the environments of galaxies in the merger-free and control samples which also appear in the Yang et al (2007) [115] group catalogue— 75 out of 101 galaxies for the merger-free sample and 53 out of 101 galaxies for the control sample (see Section 3.3).

4.2.1   Projected Distances of Companions from their Associated Sample Galaxies

The projected distance (using the RA and dec distances only, not considering redshift) was calculated for each galaxy in the group catalogue (described in Section 2.2.1) that was in the same group as a sample galaxy. This provides a measure of the environment in which each sample galaxy resides according to the group catalogue.

Figure 4.7 shows the projected distances of companion galaxies in each group from their associated sample galaxies[1]. The largest group, both in terms of projected distance and number of galaxies is one which contains a galaxy in the merger-free sample and the four groups containing the most galaxies contain a merger-free galaxy. This may indicate a tendency towards larger groups for merger-free galaxies, meaning they may reside in denser environments than the control galaxies. However, groups containing merger-free galaxies that are not among the largest four groups tend to be a similar size to the groups containing control galaxies.

Figure 4.8 shows the total number of galaxies associated with each sample by each projected distance. This is the result of summing the number of galaxies across all of the groups for each sample that are shown in Figure 4.7. This shows that at all projected distances, the merger-free sample has more companion galaxies overall than the control sample and the merger-free sample has many more companion galaxies overall than the control sample. This indicates a difference

---

[1]Groups containing a total of two galaxies (the sample galaxy and one companion) are not included in Figure 4.7 for clarity. For individual plots for each group, including those containing a total of two galaxies see Appendix B.
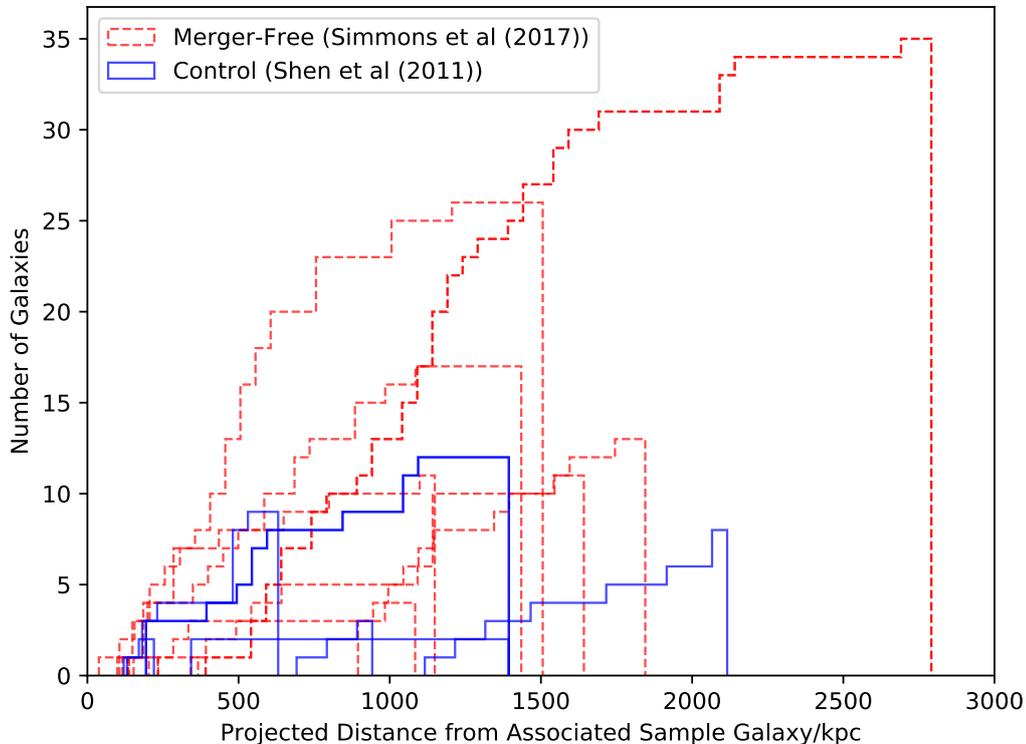
Figure 4.7: Cumulative histograms for each group showing the distances of companion galaxies from their associated sample galaxies in kpc. Each histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

in overall environment between the two samples, since the initial samples were of the same size (101 galaxies).

When a K-S test is performed to compare the two distributions shown in Figure 4.8, the distributions differ with a p value of 0.289, indicating a 1.1σ significance level. Even though the merger-free sample is associated with the largest groups and has the greatest number of overall companions, the difference in the distribution of galaxies in their environments is not statistically significant; we are unable to reject the null hypothesis that the projected distances of galaxies in groups associated with each sample are drawn from the same distribution.

Additionally, these results may be skewed by the largest groups, which are only associated with merger-free galaxies and could be anomalous. In order to analyse the differences between the smaller groups associated with merger-free sample galaxies and the groups associated with control sample galaxies, we repeat this
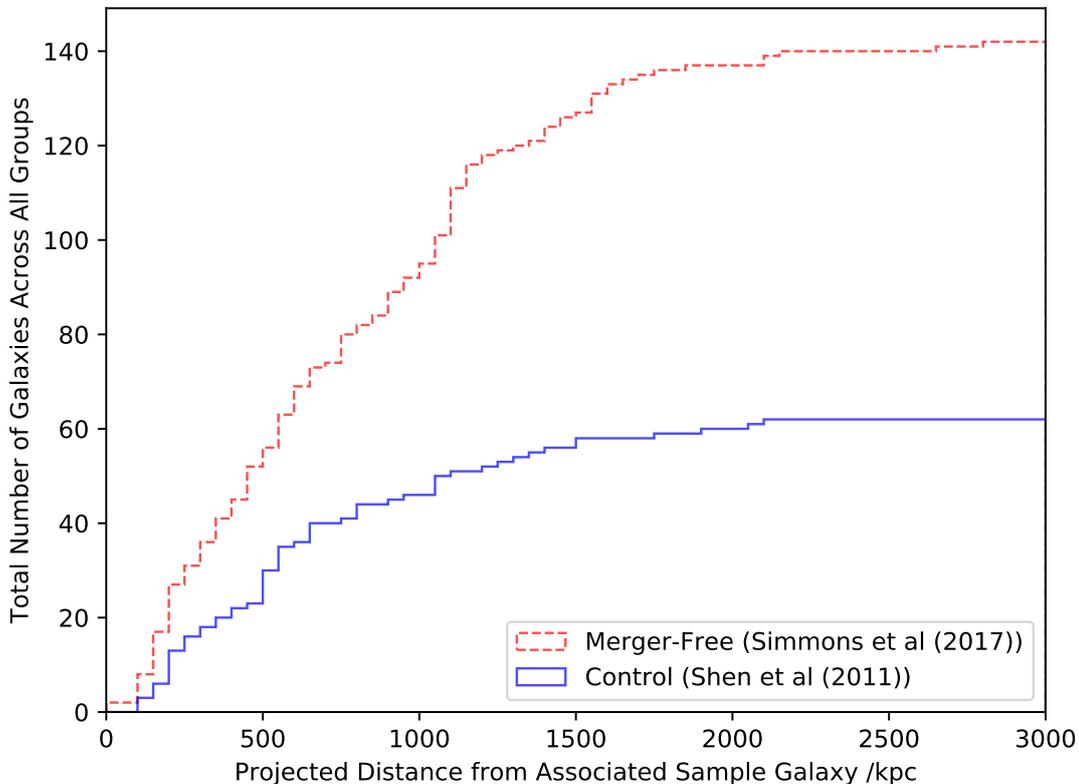
Figure 4.8: A cumulative frequency histogram showing the total number of companion galaxies associated with each sample by projected distance from their associated sample galaxy, according to the group catalogue we use. The bin size is set at 50 kpc.

analysis ignoring the largest two groups (in terms of numbers of galaxies). This results in Figure 4.9.

When a K-S test is performed to compare the new merger-free distribution to the control distribution, a p-value of 0.289 is obtained, indicating that the distributions differ with a significance of 1.1σ. This is the same result as when all groups are included, suggesting that the shape of the merger-free distribution is not being skewed by the largest two groups. As before, the distributions do not vary significantly in shape. However, there is still a greater overall number of galaxies in the merger-free sample than in the control sample. This may be interpreted to indicate that a difference in overall environment between the merger-free and control samples remains even when the two largest groups are
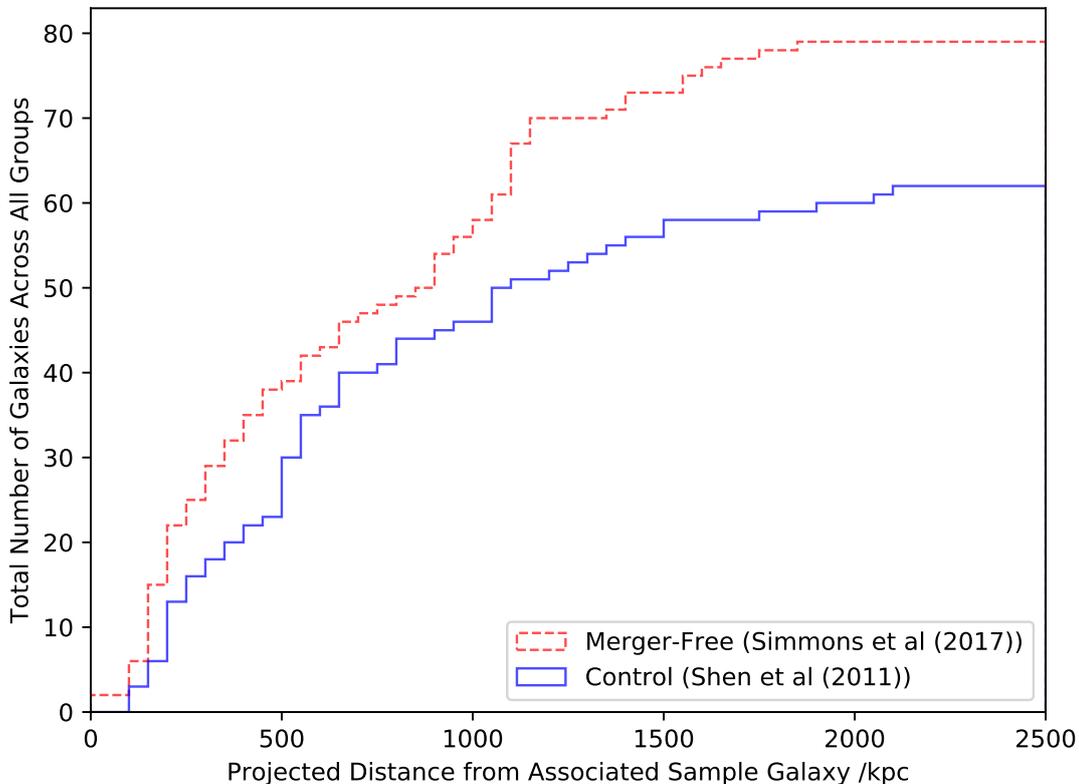
Figure 4.9: A cumulative frequency histogram showing the sum of frequency of galaxies for all groups in each sample by projected distance of companion galaxies from their associated sample galaxy, excluding the two largest groups in terms of number of galaxies. The bin size is set to 50 kpc.

not considered. This may have only occurred because more sample galaxies from merger-free sample than the control sample initially matched with galaxies in the Yang et al (2007) [115] group catalogue.

4.2.2 Potential Error in Distributions

As discussed in Section 4.1.3, the distances of potential companion galaxies from their associated sample galaxy may be modelled as a Poisson distribution. Similarly to the errors in Section 4.1, the error in the distributions shown throughout Section 4.2 may be modelled as Poisson errors. Therefore, the error in each 50 kpc bin, if the distributions were drawn as non-cumulative histograms, would

be a symmetrical error of $\sqrt{N}$, where N represents the number of galaxies in that bin. As in Section 4.1, this means that many of the errors would be very large compared to the number of galaxies being counted, given that some bins contain only one galaxy. This would especially affect the smallest groups. Such errors would accumulate when drawing a cumulative distribution showing the results of counting galaxies, which may give some uncertainty in the shapes of distributions, especially at small projected distances ($\lesssim 500$ kpc). As discussed in Section 4.1.3, the K-S test does consider the overall numbers of companion galaxies when comparing distributions by normalising the distributions before making a comparison.

4.3 Comparison Between Cylinder Search and Group Catalogue Results
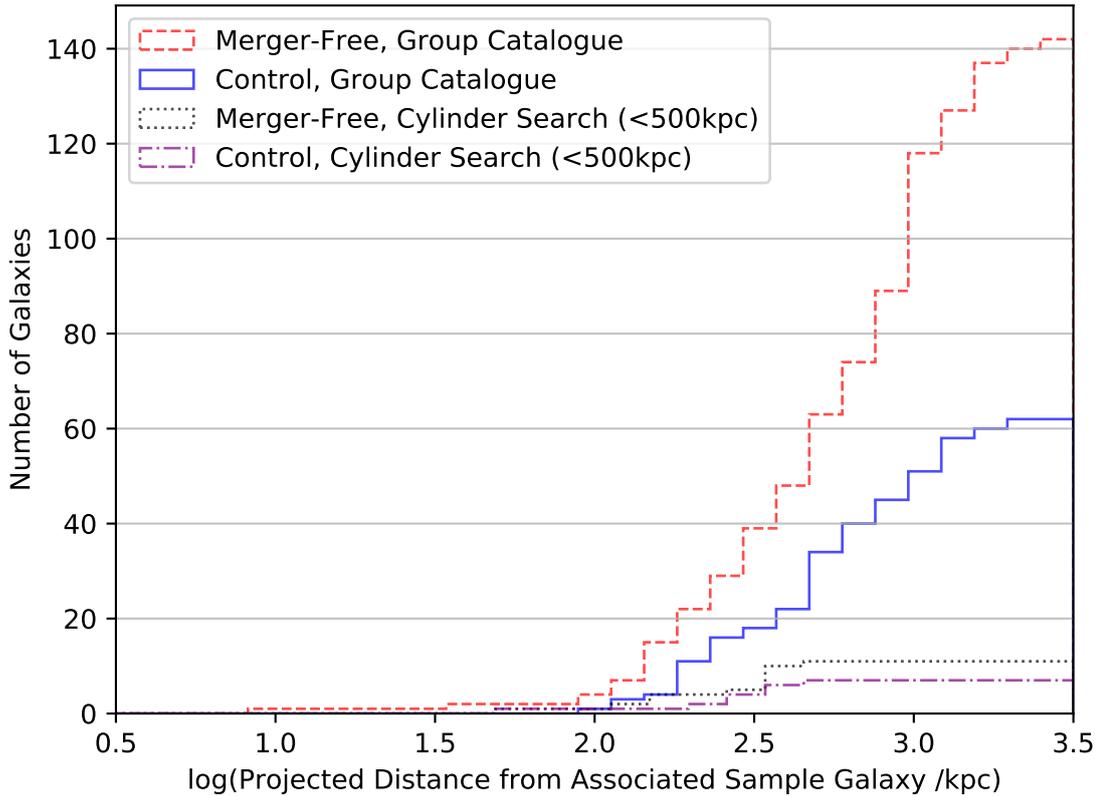
Figure 4.10: Cumulative frequency histograms showing the projected distances of companion galaxies from their associated sample galaxies for both samples for companions found using the group catalogue analysis and the cylinder search with a 'redshift distance tolerance' of 500 kpc. The x axis uses a $\log_{10}$ scale and the bin size is set to 0.1 log(kpc).

Figures 4.10, 4.11 and 4.12 compare the results of the cylinder searches (Section 4.1) to the results of our analysis of a well-studied group catalogue (Section 4.2). The figures show the results for the cylinder searches performed using 'redshift distance tolerances' of 500 kpc, 1 Mpc and 5 Mpc respectively.

The frequency for any of the cylinder search results does not increase at $\log_{10}$ (Projected Distance from Associated Sample Galaxy) $> \log_{10}(500$ kpc) ($\sim$
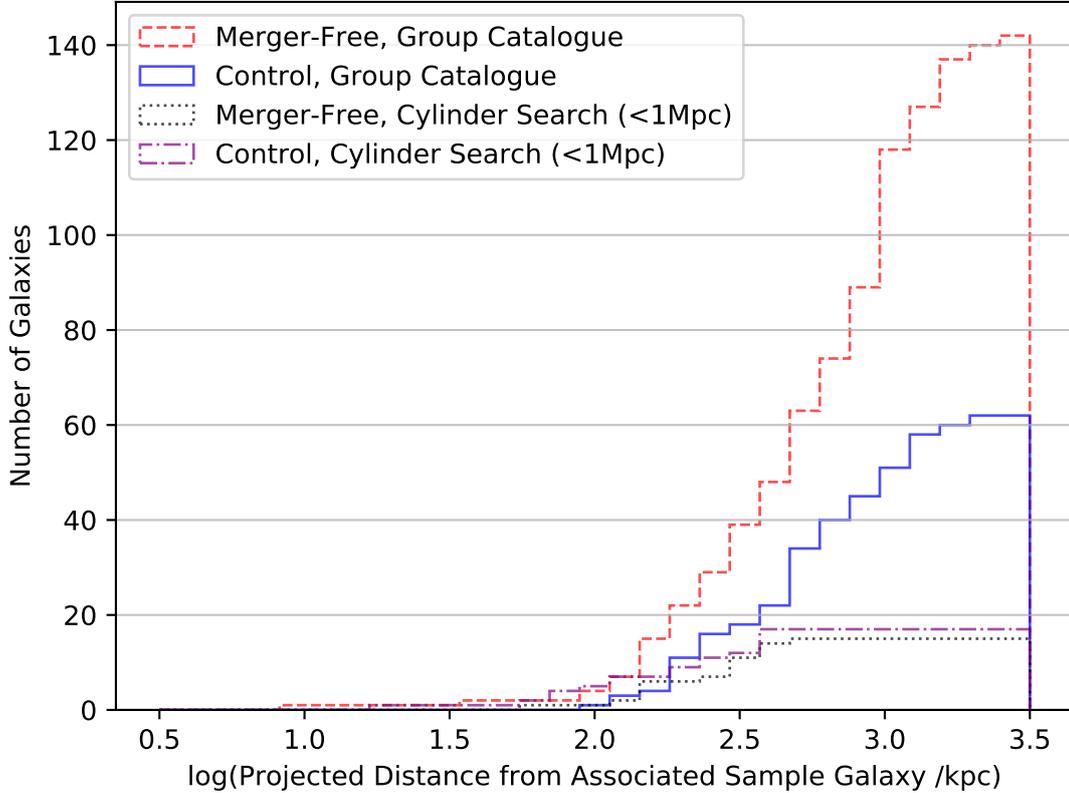
Figure 4.11: Cumulative frequency histograms showing the projected distances of companion galaxies from their associated sample galaxies for both samples for companions found using the group catalogue analysis and the cylinder search with a 'redshift distance tolerance' of 1 Mpc. The x axis uses a $\log_{10}$ scale and the bin size is set to 0.1 log(kpc).

2.7 $\log_{10}$(kpc)) because the cylinder searches were only conducted up to a projected distance of 500 kpc from each sample galaxy, whereas no distance limits were placed on the group catalogue. This makes the results difficult to compare past this point. The presence of companions in the group catalogue at further projected distances than those considered for the cylinder search indicates that, had the cylinder search been extended to further projected distances, more interacting companion galaxies would have been found. However, at further projected distances we cannot be certain that the companion galaxies are gravitationally interacting with the sample galaxies and therefore it is likely that galaxies that are not interacting with the sample galaxies would also have been found if we
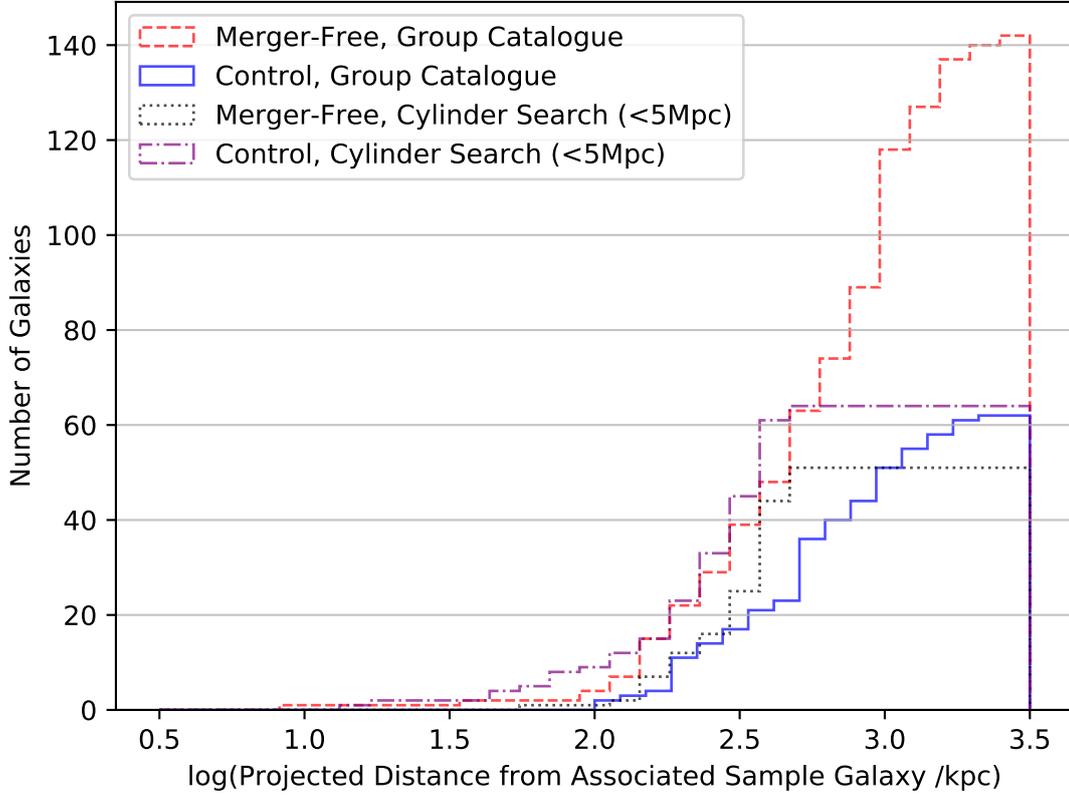
Figure 4.12: Cumulative frequency histograms showing the projected distances of companion galaxies from their associated sample galaxies for both samples for companions found using the group catalogue analysis and the cylinder search with a 'redshift distance tolerance' of 5 Mpc. The x axis uses a $\log_{10}$ scale and the bin size is set to 0.1 log(kpc).

were to expand the cylinder search radius.

When a KS test is performed to compare the distributions for each of the cylinder searches to the group catalogue results for projected distances of less than 500 kpc from an associated sample galaxy (the range in which the results are comparable) for each sample, the p-values and statistical significances obtained are as shown in Table 4.1.

The results shown in Table 4.1 indicate that there is a significant difference between the distributions of projected distances found using each method for the same sample for all of the 'redshift distance tolerances' used in the cylinder searches and for both samples. These differences are more prominent for greater

Sample	'RDT'	P-Value and Significance
Merger-Free (Simmons et al (2017))	500 kpc	1.19×10^{-4} (3.8σ)
Merger-Free (Simmons et al (2017))	1 Mpc	3.92×10^{-6} (4.6σ)
Merger-Free (Simmons et al (2017))	5 Mpc	1.11×10^{-14} (7.7σ)
Control (Shen et al (2011))	500 kpc	5.25×10^{-3} (2.8σ)
Control (Shen et al (2011))	1 Mpc	1.18×10^{-5} (4.4σ)
Control (Shen et al (2011))	5 Mpc	7.05×10^{-12} (6.9σ)

Table 4.1: The results of KS statistical tests comparing the distributions of projected distances of companion galaxies from their associated sample galaxies for the cylinder searches (using each 'redshift distance tolerance' (RDT)) and the group catalogue for each sample.

'redshift distance tolerances'. This shows that the distributions of potential companion galaxies found when using each method differ significantly in shape. This may be a result of differences in which galaxies each method consider to be potential companions. For example, a cylinder search will consider any galaxy within a certain distance of a sample galaxy to be a potential companion, while the Yang et al (2007) [115] algorithm may not find them to be interacting. This may lead to cylinder searches finding more potential companions than the group finder. Similarly, the Yang et al (2007) [115] group finder algorithm may find galaxies that are outside of the cylinder search area to be linked to a sample galaxy, which may lead to the group catalogue containing more companion galaxies than are found by a cylinder search at a given projected distance from the sample galaxy.

4.4 Environment Coefficients

4.4.1 Environment Coefficients of Sample Galaxies
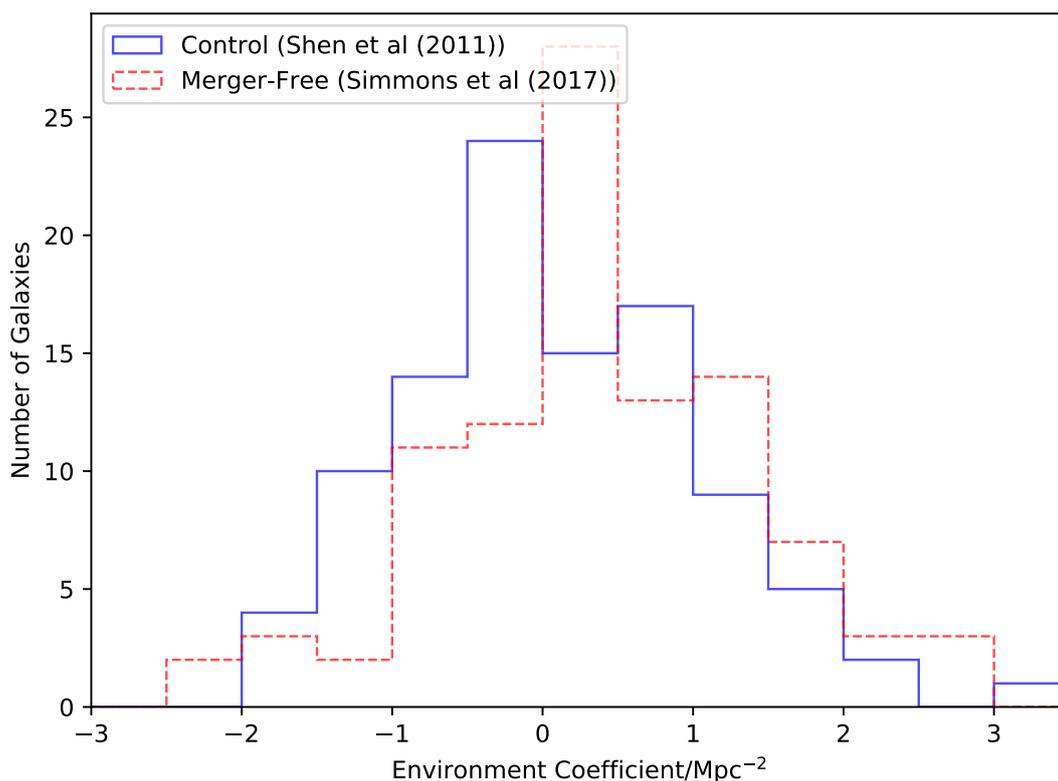
Figure 4.13: Histograms comparing the environment coefficients obtained for galaxies in the merger-free and control samples. The bin size is set to 0.5 Mpc^{-2}.

Figure 4.13 shows the frequency distribution of environment coefficients for galaxies in both of the samples that we consider.

There is a slight difference in peak environment coefficient between the merger-free and control samples. In the merger-free sample, the 0–0.5Mpc^{-2} bin contains the greatest number of galaxies, whereas in the control sample, the -0.5–0 Mpc^{-2} bin contains the greatest number of galaxies. The merger-free distribution also appears to tend towards greater environment coefficients than the control sample. This may be interpreted to indicate that galaxies in the merger-free sample may

reside in denser environments to galaxies in the control sample. When a K-S statistical test is performed comparing these two distributions, a p-value of 0.00919 is obtained, indicating that the distributions differ with a statistical significance of 2.6σ. Therefore, the differences between these distributions are of marginal significance.

In the results detailed in this section, we do not include three galaxies in the merger-free sample which— upon inspection of Hubble Space Telescope images processed whilst the research presented in this dissertation was ongoing— appeared to be either elliptical galaxies or currently undergoing a merger, although a full morphological analysis of these images, which would confirm or disconfirm this, has yet to be conducted and is outside the scope of this dissertation. These galaxies had been included in previous sections of this dissertation. Figure 4.14 shows the environment coefficients of both samples when these galaxies remain included. When a K-S statistical test is performed comparing the merger-free distribution shown in Figure 4.13 to that shown in Figure 4.14, a p-value of 1.00 is obtained, indicating there is almost no difference between the two distributions and therefore we can conclude that removing these three galaxies from the merger-free sample has not significantly affected our results.
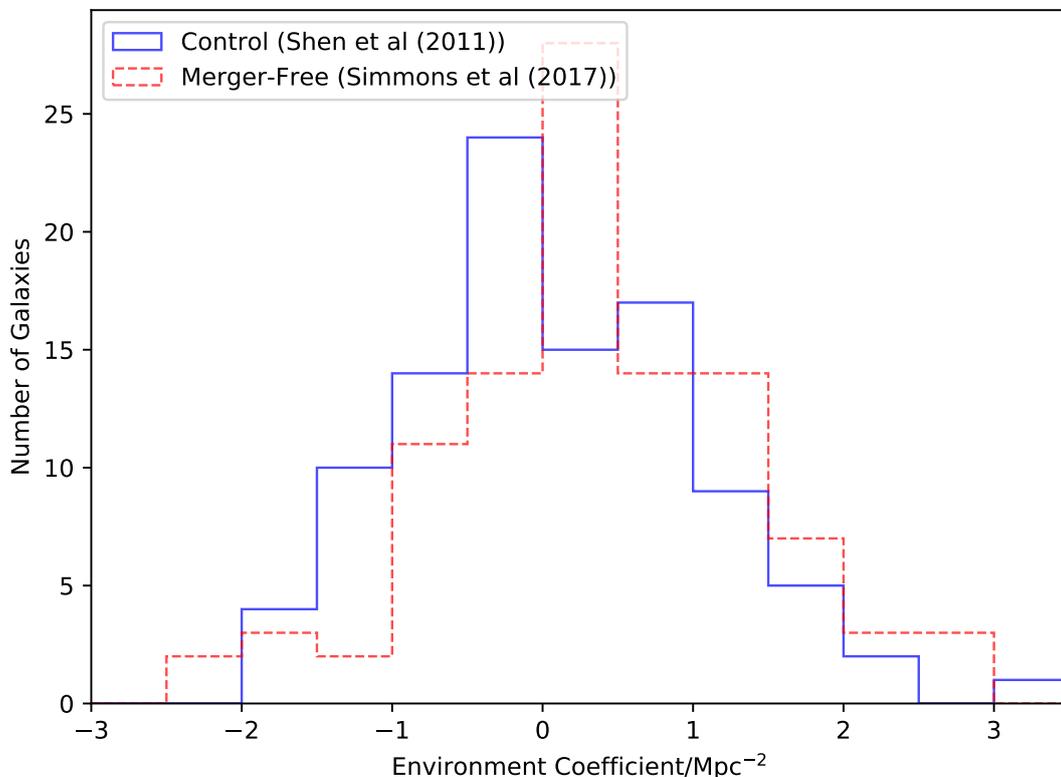
Figure 4.14: Histograms comparing the environment coefficients obtained for galaxies in the merger-free and control samples with all galaxies included. The bin size is set to 0.5 Mpc^{-2}.

4.4.2 Potential Error in Distributions

As in previous sections, the distribution of the environment coefficients of galaxies in the merger-free and control samples may be modelled as a Poisson distribution, assuming that potential companion galaxies are randomly distributed. Therefore, the error in a discrete binned distribution of galaxies with various environment coefficients, as shown in Figure 4.14 would be $\sqrt{N}$ for each bin of 0.5 Mpc^{-2}, where N represents the number of galaxies in that bin (see Section 4.1.3 for a further discussion of Poisson errors). Therefore, in the case of bins containing few galaxies, the error in the count may be large compared to the number of galaxies in the bin. In the case of the -2.5– -2 bin for the merger-free sample, which contains 2 galaxies, for example, the Poisson error would be $\sqrt{2}$ (1.41 (3sf)).

Since the values of a count must be discrete, this would indicate that the true value of the number of galaxies in that bin may be as few as 0 or as many as 3. Such differences would affect the overall shape of the distribution.

However, the bin in which the merger-free distribution peaks remains certain as the most populated bin for this distribution contains 28 galaxies, giving an error of $\sqrt{28}$ (5.29 (3 sf)), meaning the lowest possible true value of the peak is 22. The greatest possible value of the second most populated bin (which contains 15 galaxies and therefore has an error of $\sqrt{15}$ (3.87 (3 sf))), is 19. Therefore, the most populated and the second most populated bins cannot overlap for the merger-free distributions. The same is not true of the control sample distribution, in which the most populated and second most populated bins contain 24 and 17 galaxies respectively. This would mean that the lowest possible value of the most populated bin would be 19 and the highest possible value of the second most populated bin would be 22. Also, the two bins which contain 16 galaxies have a maximum true value of 20 galaxies, meaning that these bins may also be the location of the true peak for the control distribution. This casts some uncertainty over the difference in shape between the merger-free and control distributions of environment coefficients, since within error they may peak in the same bin or the control sample may peak at a greater environment coefficient than the merger-free sample.

4.4.3 Comparison of Sample Environment Coefficients to those of the Baldry et al (2006) Sample

In this section, we compare the environment coefficients that we calculated for the two samples we consider to the environment coefficients calculated by Baldry et al (2006) [152] for the version of their sample that is based on the seventh data release of the Sloan Digital Sky Survey [107],[110]. We do this in order to ensure that our environment coefficients are reasonable and— since most of the Baldry et al (2006) [152] sample is at a lower redshift than the merger-free and control samples (all galaxies in the Baldry et al (2006) [152] sample are at $z < 0.14$ whereas galaxies in the merger-free and control samples are at $0.025 \leq z < 0.275$ (as shown in Section 2.1))— to assess any effect that redshift may have when applying the method of Baldry et al (2006) [152] to samples containing slightly higher-redshift galaxies, such as the merger-free and control samples.

Figure 4.15 displays this comparison as a set of histograms showing the environment coefficient (x axis) against normalised frequency (number of galaxies). It was necessary to normalise the number of galaxies due to samples being of very different sizes. The Baldry et al (2006) [152] sample contains 265466 galaxies, whereas the samples that we consider contain only 101 galaxies each. In Figure 4.15, the frequency is normalised such that the total area beneath each histogram is equal to 1 Mpc^{-2}.

It is visible from Figure 4.15 that the distribution of environment coefficients for the Baldry et al (2006) [152] sample has a narrower range than those of the other samples and the peak of the Baldry et al (2006) [152] distribution occurs within the $-0.5 \leq \sigma < 0$ Mpc^{-2} bin, which is the same bin in which the peak occurs for the control sample distribution, but not for the merger-free distribution, whose peak occurs in the $0 \leq \sigma < 0.5$ bin. When a K-S statistical test is performed comparing the environment coefficients Baldry et al (2006) [152] sample to those of the merger-free sample, a p-value of 1.43×10^{-13} is obtained, indicating that the distributions differ with a statistical significance of 7.4σ. When a K-S statistical test is performed comparing the environment coefficients Baldry et al (2006) [152] sample to those of the control sample, a p-value of 2.64×10^{-4} is obtained, indicating that the distributions differ with a statistical significance
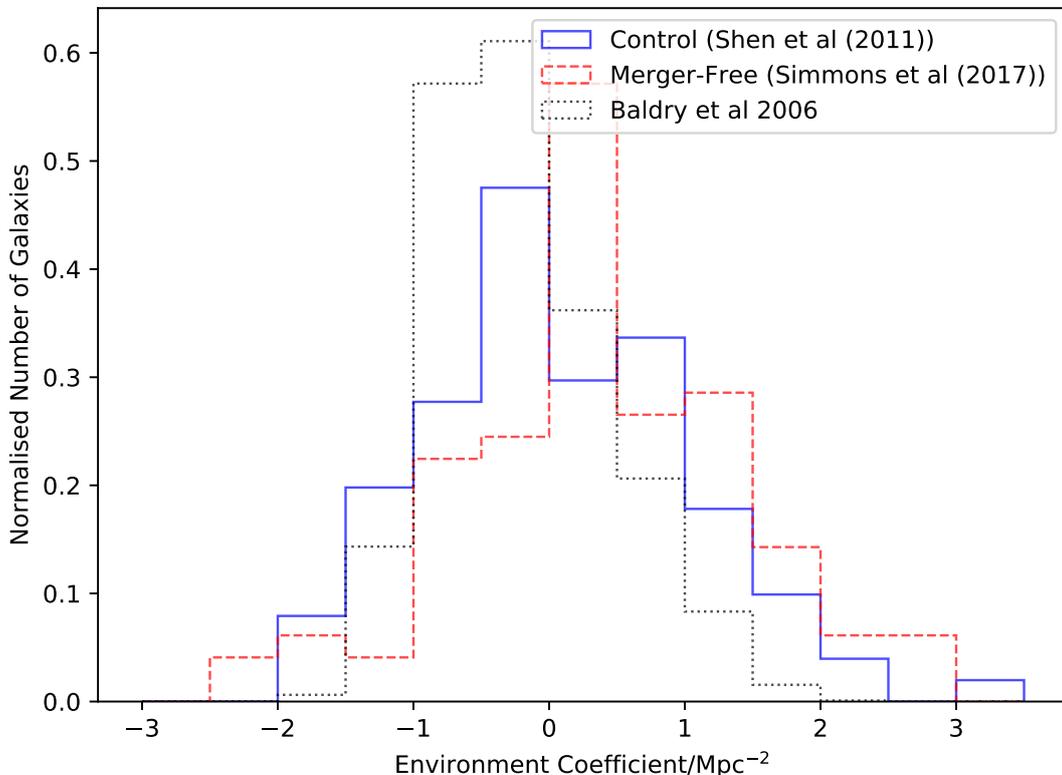
Figure 4.15: Normalised histograms comparing the environment coefficients obtained for galaxies in the two samples that we consider to those obtained in the Baldry et al (2006) study for a different sample of galaxies. The bin size is set to 0.5 Mpc^{-2}.

of 3.6σ. The large size of the Baldry et al (2006) [152] sample is likely to cause any differences between samples to be of high statistical significance.

These differences may occur due to greater uncertainty being present in the redshift, and hence the distance, of the merger-free and control sample galaxies and their neighbours compared to the Baldry et al (2006) [152] sample, or due to the lack of redshifts available for neighbouring galaxies to the merger-free and control samples causing some sample galaxies to have lower environment coefficients than they would have if redshifts were available for all galaxies— however, the latter explanation does not account for an wider range appearing at higher environment coefficients as well as lower ones. Alternatively, this could be a result of the difference in redshift between the Baldry et al (2006) [152] and the

merger-free and control samples. It should also be noted that we have not applied any additional selection criteria to the Baldry et al (2006) [152] sample, therefore factors such as differences in the stellar mass of galaxies between samples may result in differences between distributions.

4.4.4 Redshift-Environment Coefficient Relation
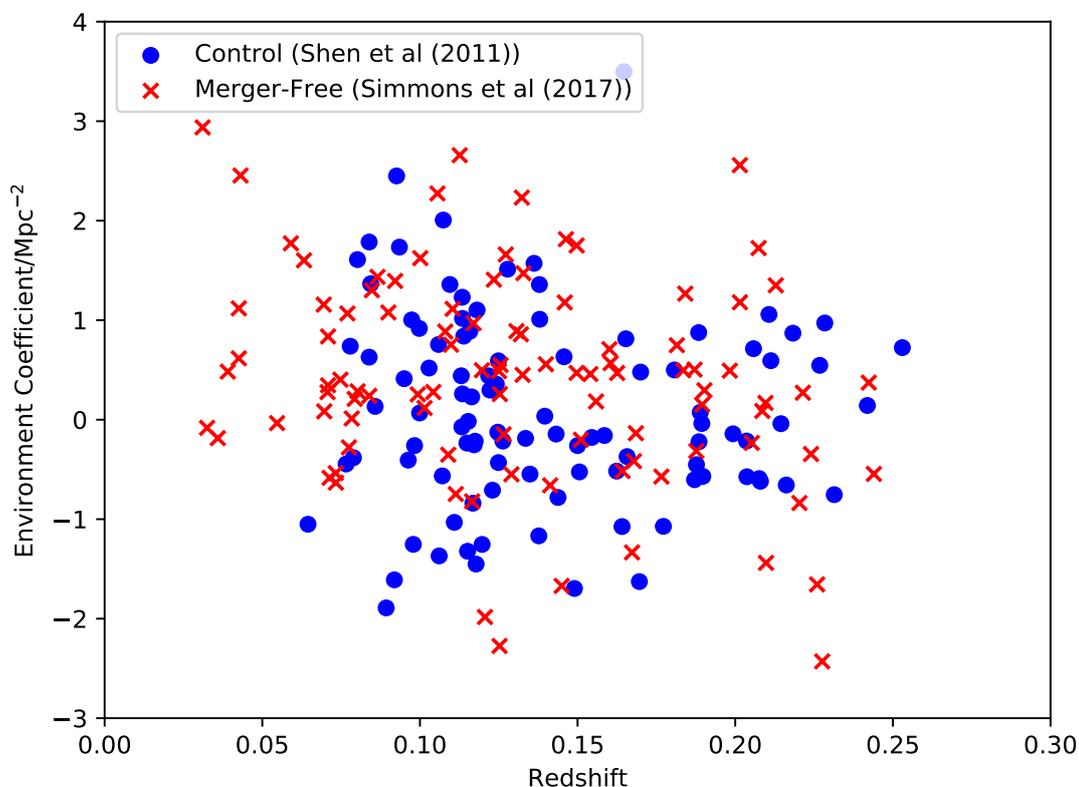
Figure 4.16: A scatter plot showing the relationship between environment coefficient and redshift for each of the sample galaxies.

Figure 4.16 shows the distribution of galaxies in terms of their environment coefficient (y axis) and redshift (x axis). This shows a fairly random distribution of points for both samples, with no clear correlation between redshift and environment coefficient visible for either sample. To give a quantitative value of the correlation, we calculate the Pearson correlation coefficients between redshift and environment coefficient for each sample. This gives a value of -0.248 (3sf) for

the merger-free sample and -0.0600 (3sf) for the control sample. This indicates a slight negative correlation between redshift and environmental coefficient for the merger-free sample and almost no correlation between redshift and environmental coefficient for the control sample. The negative correlation found for the merger-free sample may be due to a detection bias, since neighbours to galaxies at higher redshifts may be fainter than those at lower redshifts, and hence less likely to be detected. This may affect our calculation of environment coefficients since we only consider galaxies with redshifts available in SDSS DR9 [107],[109]. The same correlation is not found in the control sample, despite the same method and survey being used when searching for neighbouring galaxies and calculating the environment coefficient for both samples. This may occur because, as discussed in Section 2.1, the merger-free sample contains some galaxies at lower redshifts than any of those in the control sample.

It is also possible that a detection bias against galaxies within 55–60 arcsecs of other galaxies due to fibre collisions is present. This would have a greater effect on galaxies at higher redshifts than on those at lower redshifts because all angular scales are smaller at greater redshifts than at lower redshifts for equal linear distances. Hence, fibre collisions could cause a negative correlation between redshift and environment coefficient as is shown in Figure 4.16.

However, the negative correlation found between redshift and environment coefficient for the merger-free sample is small, and therefore the effect of selection bias and the difference in the effect of fibre collisions differences in redshift within a sample may not be considered to have significantly impacted our results for either sample when considered individually.

The difference redshift range between the merger-free and control samples may have impacted our results. As is shown in Figure 4.16, the samples are not matched in redshift and are also not matched in environment. The slight negative correlation between redshift and environment coefficient in the merger-free sample occurs due to some galaxies with redshifts lower than any of those in the control sample having very high environment coefficients. This could lead to the merger-free sample appearing to have a greater average environment coefficient than the control sample due to the merger-free sample containing galaxies at lower redshifts than the control sample. In order to compare these samples

more fairly, either control galaxies at lower redshifts would need to be added to the control sample or galaxies at redshifts lower than any of those in the control sample would need to be removed from the merger-free sample.

Chapter 5

Conclusion

5.1 Summary

We investigate the possible relationship between supermassive black hole growth and minor gravitational interactions short of mergers. To this end, we compare two samples of galaxies: the Simmons et al (2017) [55] sample of bulgeless AGN host galaxies, which we refer to as the 'merger-free' sample (fully described in Section 2.1.1), and the 'control' sample (fully described in Section 2.1.2), a subset drawn from the Shen et al (2011) [53] sample, whose redshift is approximately matched to that of the 'merger-free' sample, but is slightly higher on average—the merger-free sample has an average redshift of $\langle z \rangle = 0.132 \pm 0.053$, whilst the control sample has an average redshift of $\langle z \rangle = 0.141 \pm 0.045$[1]. These average redshifts are not central for either sample, since the redshift distribution of both samples is skewed towards lower redshifts. It is also important to note the caveat that AGN in the control sample generally have greater black hole masses and greater bolometric luminosities than galaxies in the merger-free sample. This is possibly a result of the samples not being entirely matched in redshift.

We employ three methods to analyse the environments of the galaxies in both samples: three cylinder searches for potential companion galaxies, with search

---

[1]The errors in average redshifts shown are the standard deviations of each redshift distribution around the mean value for each sample.

radii of 500 kpc and depths of 500 kpc, 1 Mpc and 5 Mpc respectively either side of the sample galaxies (Section 3.2); a search for galaxies that are in the same groups as sample galaxies in the well-studied Yang et al (2007) [115] group catalogue (Section 3.3) and calculation of environment coefficients for each of the sample galaxies using the method of Baldry et al (2006) [152] (Section 3.4).

From the cylinder searches, we find that the distributions of projected distances of potential companion galaxies from their associated sample galaxies do not vary significantly when search depths of 500 kpc and 1 Mpc were applied (the samples differed with statistical significance levels of 0.6σ (500 kpc) and 0.1σ (1 Mpc)), but did vary with marginal statistical significance (2.1σ) when a search depth of 5 Mpc was applied. At projected distances >350 kpc, we found more potential companion galaxies to the merger-free sample than to the control sample, but at projected distances ≤350 kpc, we found less potential companion galaxies to the merger-free sample than to the control sample. Whilst we showed in Section 3.5 that galaxies at 5 Mpc from a sample galaxy would not be directly gravitationally interacting with the sample galaxy, the large-scale environment may have an effect on the evolution of galaxies through conformity [149],[150]. Also, it is possible, although speculative, that galaxies found to have a high difference in redshift from the sample galaxies may not be at high distances from the sample galaxies, but travelling at high peculiar velocities relative to the sample galaxies, since our calculation of distances based on redshifts assumes that all redshift differences between galaxies originate from the expansion of the Universe. Such galaxies passing others at high peculiar velocities might transfer enough kinetic energy to funnel gas into a supermassive black hole, thereby contributing to increased AGN activity. If these 'fly-by' galaxies were passing the sample galaxies at high peculiar velocities and causing a disturbance which contributed to black hole growth, this would explain the increased number of galaxies at greater projected distances for the merger-free sample compared to the control sample. Alternatively, the observed difference in the number of potential companion galaxies to the merger-free sample across projected distances from the sample galaxies may occur due to the increased effect of fibre collisions at lower projected distances from the sample galaxies, but this would not explain the difference in distributions between the merger-free sample and the control

67

sample, as in this case the control sample would be expected to be more severely affected than the merger-free sample at low projected distances from the sample galaxies due to the control sample's higher average redshift, and this is not apparent in the results of our cylinder search with a 5 Mpc search depth. It is also important to note that the results of all of our cylinder searches are subject to large Poisson errors relative to the numbers of galaxies counted due to the small numbers of potential companion galaxies that were counted. Therefore, these results are subject to large uncertainties, although this is accounted for somewhat in the normalisation process when calculating the statistical significance of our results.

Our analysis of the environments of the merger-free and control samples of galaxies based on the Yang et al (2007) [115] group catalogue found no statistically significant difference (a difference at the 1.1σ statistical significance level) between the merger-free and control samples in terms of the projected distances of companion galaxies from their associated sample galaxies. However, a much greater total number of companion galaxies was found for the merger-free sample compared to the control sample. This could be a result of the merger-free sample having a greater overlap with the Yang et al (2007) [115] group catalogue than the control sample or could be interpreted to indicate that galaxies in the merger-free sample reside in denser environments than galaxies in the control sample—further studies and/or analysis would be required to determine whether this indicates a physical difference between the samples. Also, the two largest groups of galaxies were those associated with the merger-free sample. This result is less affected by the variation in overlap with the group catalogue between the samples (although the probability of a galaxy being in a large group may be greater for a larger sample). The two largest groups being associated with the merger-free sample could be interpreted as an indicator that being a member of a large group, and hence having a large number of gravitationally interacting companion galaxies, contributes to supermassive black hole growth in bulgeless galaxies.

From our calculation of environment coefficients for the merger-free and control samples, we find that there is a marginally significant (2.6σ) difference between the environment coefficients of the merger-free and control samples. Galaxies in the merger-free sample are found to typically have greater environment co-

efficients than those in the control sample. This could indicate that the galaxies in the merger-free sample reside in denser environments than those in the control sample and if this is also found to be the case in future studies, that may indicate that residing in a higher density environment contributes to supermassive black hole growth in bulgeless galaxies.

Across all the methods of environmental analysis that we have employed, we find that there is either no statistically significant difference in environment between the merger-free and control samples or that there is a difference in environment of marginal significance. The results of our calculation of environment coefficients indicates that the galaxies in the merger-free sample may potentially reside in overall denser environments than galaxies in the control sample. The results of our cylinder search with a 5 Mpc search depth suggest that, whilst the nearby environment of galaxies in the merger-free sample may be less dense than that of galaxies in the control sample, their environment at projected distances > 350 kpc may be denser than that of galaxies in the control sample. However, since the statistical significance of both of these results is marginal, then further studies would be required in order to state this with more certainty. Our results indicate that it is possible for high-density environments to be contributing factors in, or responsible for, supermassive black hole growth in galaxies with no recent history of mergers (bulgeless galaxies). Our results may also support the possibility of gravitational interactions with high-velocity 'fly-by' galaxies contributing to or being responsible for supermassive black hole growth in bulgeless galaxies. However, the low statistical significance of our results means that we are currently unable to reject the null hypothesis that minor gravitational interactions do not affect supermassive black hole growth in bulgeless galaxies.

In addition, all of our results could have been affected by fibre collisions. These occur when two sources in the Sloan Digital Sky Survey (SDSS) have an angular separation of less than 55 arcsecs, since this is the minimum separation of the centres of two SDSS fibres [107]. Fibre collisions could have a significant effect at other separations also, especially those of less than 1 arcmin [151]. In our case, this would mean any two potential companion galaxies with separations $\lesssim 1$ arcmin between them may have been affected by fibre collisions, which would have caused one of the galaxies to not have an available spectroscopic redshift

measurement, and hence either not have a calculable three-dimensional distance or have a relatively large error in this distance due to the use of a photometric redshift.

In the case of the results of the cylinder search, since at the highest redshift of any galaxy in the merger-free sample ($z = 0.244$), an angular separation of 1 arcmin corresponds to a linear distance of 232 kpc, there may be a lack of recorded potential companion galaxies with smaller separations than this from their associated sample galaxy due to fibre collisions. This is similar for the control sample, whose maximum redshift is $z = 0.253$, at which an angular separation of 1 arcmin corresponds to a linear distance of 238 kpc. Therefore, fibre collisions may partially explain the lack of potential companion galaxies at projected distances of ≤ 350 kpc seen in the cylinder search with a 5 Mpc depth for the merger-free sample. This is in addition to the more general effect of fibre collisions potentially 'removing' any galaxies with similarly small separations from any other galaxies from our results. As shown in Table 3.2, $> 99\%$ of galaxies within a 500 kpc projected distance of galaxies in the merger-free and control samples did not have available spectroscopic redshifts, and hence would have either used a photometric redshift (with an error of $\sim 21\%$) to calculate their three-dimensional distance or would not have been included in our results due to having no available redshift measurement. This is likely to be due to a mixture of fibre collisions and galaxies being too faint to measure their redshift. 60% of galaxies within a 500 kpc projected distance of galaxies in the merger-free sample and 70% of galaxies within a 500 kpc projected distance of the control sample did not have any available redshift measurement, meaning that the results of the cylinder searches performed around the control sample may have been disproportionately affected by 'missing' potential companion galaxies compared to those performed around the merger-free sample.

Fibre collisions affect $\sim 7\%$ of potential companion galaxies that might be covered by the Yang et al (2007) group study for the sample that we analysed [115]. In this case, since only spectroscopic redshifts were used, this caused these galaxies to not be included in the results due to not having a spectroscopic redshift measurement available from SDSS DR7 [110]. Therefore, the results of our

analysis of the results of the Yang et al (2007) group finder [115] may be incomplete due to fibre collisions.

In the case of our calculation of environment coefficients for galaxies in the merger-free and control samples, fibre collisions are likely to have affected the overall results in a similar way to the cylinder search results— by causing some galaxies to be 'missing' from the final results and causing others to have their three-dimensional distance calculated using a photometric redshift rather than a spectroscopic redshift. Since all angular scales correspond to greater linear distances at higher redshifts and the control sample has a higher average redshift than the merger-free sample, fibre collisions may have had a greater effect on the environment coefficients of galaxies in the control sample than on the environment coefficients of galaxies in the merger-free sample. This is one possible explanation for the difference in environment coefficients seen between the merger-free and control samples. However, further work involving matching the samples completely in redshift and correcting for fibre collisions would be required to determine whether or not this is the cause of the difference.

5.2   Further Work

Given the marginal statistical significance of our results, we suggest that further work should be undertaken in order to ascertain whether or not statistically significant relationship is present between the environment of bulgeless galaxies and supermassive black hole growth. The Large Synoptic Survey Telescope (LSST), whose first light is expected in 2020, will survey the entire southern sky every night, producing extremely large amounts of data [160],[161]. This would allow us to gather much larger samples of bulgeless AGN host galaxies as well as control AGN host galaxies at similar redshifts to the bulgeless samples. Unsupervised machine learning, alongside other methods currently in use such as visual classification by humans via Galaxy Zoo [112], can be used as a tool to quickly analyse the morphology of galaxies [162], allowing us to identify bulgeless galaxies in the LSST data. Analysis of such large samples would clarify the environments in which bulgeless galaxies reside and may provide a more definitive answer as to whether or not this differs between bulgeless and control AGN host galaxies.

As for the currently available Simmons et al (2017) [55] sample, whilst images from the Hubble Space Telescope (HST) are currently available for some of the sample galaxies, providing higher resolution images than those available from the Sloan Digital Sky Survey (SDSS) [107] and allowing detailed morphological analysis, such images are not yet available for all galaxies in the sample. As mentioned in Sections 2.1.1.4 and 4.4.1, some sample galaxies which have HST images available appear to not be bulgeless galaxies as was previously thought. This could also be the case for other galaxies in the sample for which HST images are not yet available, hence it is important to take high resolution images of all the sample galaxies and conduct a detailed morphological analysis on these images and the HST images already available for the sample.

In Section 4.3, we show that the Yang et al (2007) [115] group catalogue finds companion galaxies to galaxies in both the merger-free and control samples at greater projected distances from sample galaxies than our cylinder searches covered. In order to make a full comparison between the group catalogue and cylinder searches, and to find potential companion galaxies at these greater projected distances— especially for sample galaxies not included in the Yang et al

(2007) [115] group catalogue— it would be necessary to extend the radii of our cylinder searches to at least the maximum projected distance of a companion galaxy in the group catalogue. However, this would increase the risk of galaxies that were not gravitationally interacting with the sample galaxies being considered as potential companions.

Further work may also be necessary in order to account for the effect of fibre collisions on our results. Without such corrections, galaxies that are within 55–60 arcsecs of another galaxy in the Sloan Digital Sky Survey [107] may not have an available spectroscopic redshift measurement [107],[151], causing them to either not be included in our results or to have a large error in their redshift. One possible method of correcting for fibre collisions is that employed by Yang et al (2007) [115], in which galaxies affected by fibre collisions are assigned the same redshift as the galaxies with which they collide, except in cases where this would cause the galaxy's implied absolute magnitude to be excessively large. Whilst we do not use the corrected sample when analysing the results of the Yang et al (2007) group finder in this work, performing the same analysis on the corrected sample provided by Yang et al (2007) and comparing this to the results in Section 4.2 would allow us to assess and account for the effect of fibre collisions on these results. The same correction method could also be applied to the other results in this dissertation, however it is important to note that whilst the assumption that galaxies that are affected by fibre collisions have the same redshifts as the galaxies with which they collide is reasonable in many cases, this would not produce entirely accurate results as such galaxies are unlikely to share the exact same redshift as their neighbouring galaxies.

In addition, the difference in redshift between the merger-free and control samples may cause each sample to be affected differently by fibre collisions. Further work— involving matching the samples fully in redshift by either removing galaxies in the merger-free sample with redshifts lower than any of those in the control sample or including control galaxies from a different sample that are at these lower redshifts— would be necessary to ensure that the differing effects of fibre collisions between the samples did not translate to a difference in environment in our results.

Appendix A

Derivation of Free Fall Time Equation

The following is a derivation of Equation 3.5 from Newton's Law of Gravitation and Newton's Laws of Motion [148], used to calculate the free fall time between two galaxies.

Given the definition of gravitational force from Newton's Law of Gravitation:

$$F_{grav} = \frac{GMm}{R^2}$$

where F_{grav} represents the total gravitational force, G represents Newton's gravitational constant, M represents the mass of the system we are calculating the free fall time for, m represents a second mass both being acted upon by and exerting gravitational force on the system we are calculating the free fall time for and R represents the radius of the system.

And the general force equation from Newton's Second Law of Motion:

$$F_{acceleration} = ma$$

where $F_{acceleration}$ represents the total force due to acceleration, m represents the mass of the accelerating object— in this case the same small test mass as in the previous equation and a represents the acceleration.

Assuming all motion is due to the action of gravitational force acting upon the small test mass:

$$F_{grav} = F_{acceleration}$$
$$\therefore ma = \frac{GMm}{R^2}$$
$$\implies a = \frac{GM}{R^2}$$

Given the definition of acceleration:

$$a = \frac{dv}{dt} = \frac{d^2r}{dt^2}$$
$$\therefore \frac{d^2r}{dt^2} = \frac{GM}{R^2}$$

where v represents the velocity of the accelerating object, t represents the time elapsed and r represents the distance travelled by the accelerating object.

We can integrate this twice to obtain the free fall time:

$$\frac{dr}{dt} = \int_0^t \frac{GM}{R^2} dt = \frac{GM}{R^2} t$$

$$\implies r = \int_0^t \frac{GM}{R^2} t \, dt = \frac{GM}{2R^2} t^2$$

$$t = t_{ff} \, , \; r = R$$

$$\therefore t_{ff} = \sqrt{\frac{2R^3}{GM}}$$

where t_{ff} represents the free fall time.

Appendix B

Individual plots for each group in Figure 4.7

This chapter displays individual plots for each group of companion galaxies shown in Figure 4.7. Also included are plots showing groups for which there was only one companion galaxy (therefore the entire group contained two galaxies including the sample galaxy), which were not included in Figure 4.7.

Figure B.1: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Figure B.2: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Figure B.3: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Figure B.4: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Figure B.5: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Figure B.6: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
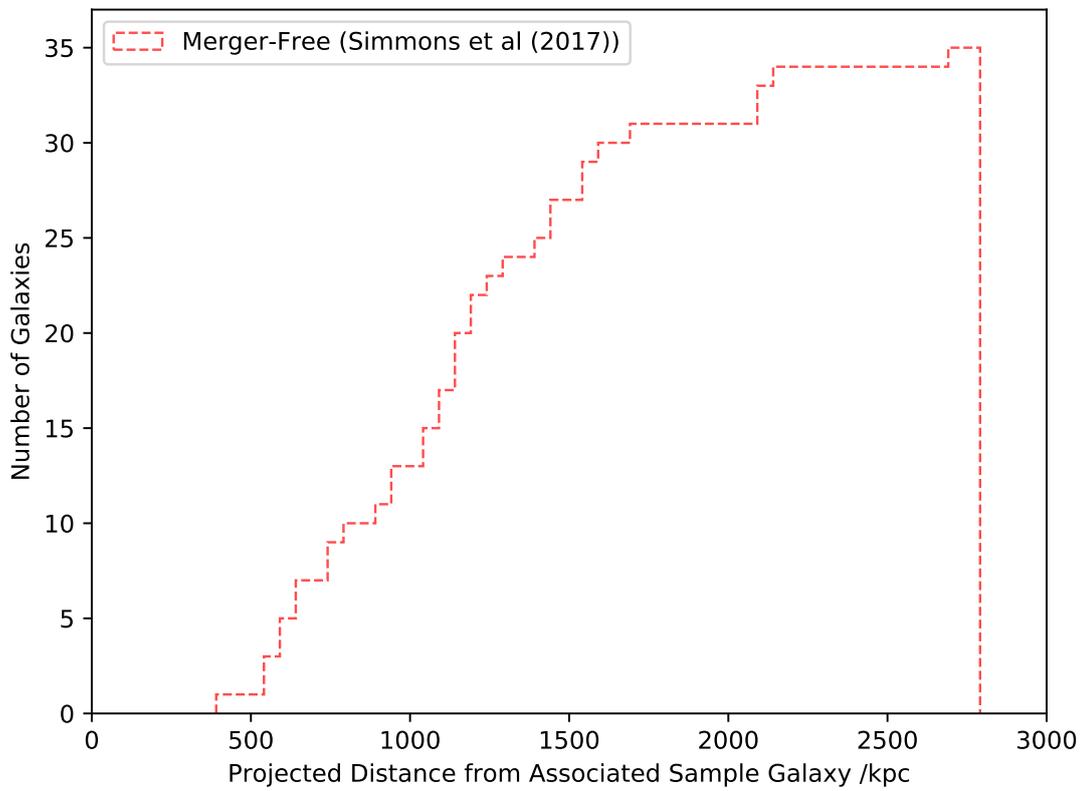
Figure B.7: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Figure B.8: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
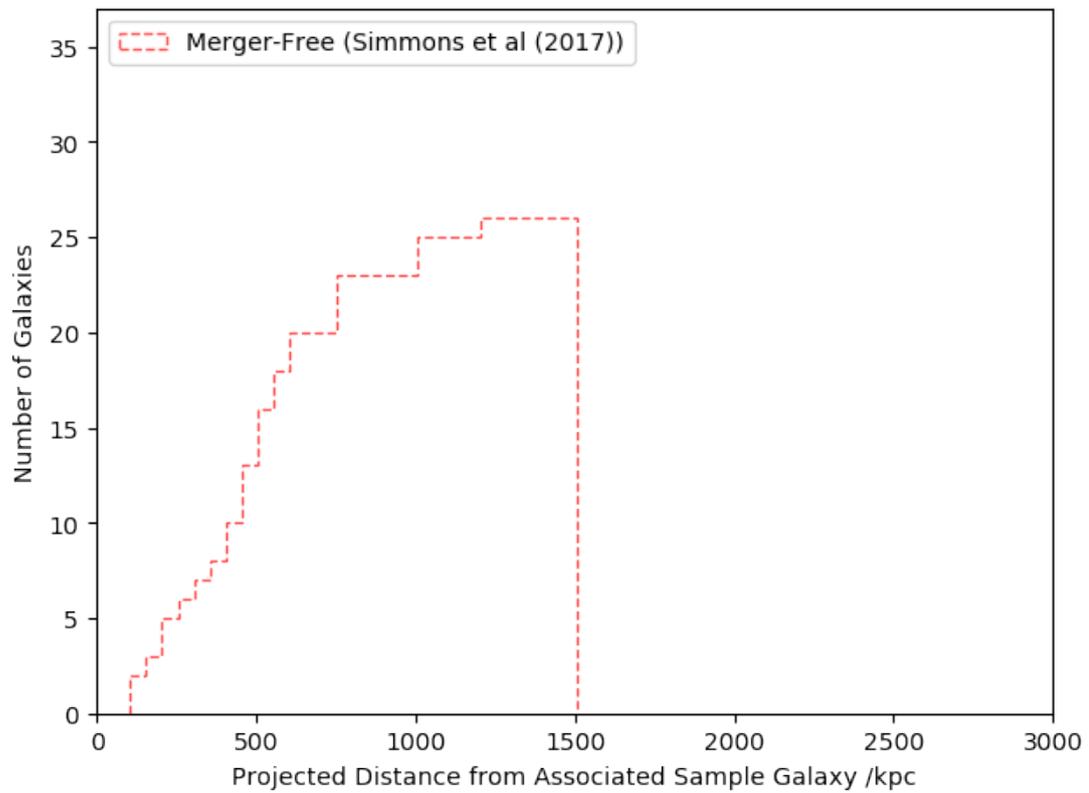
Figure B.9: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
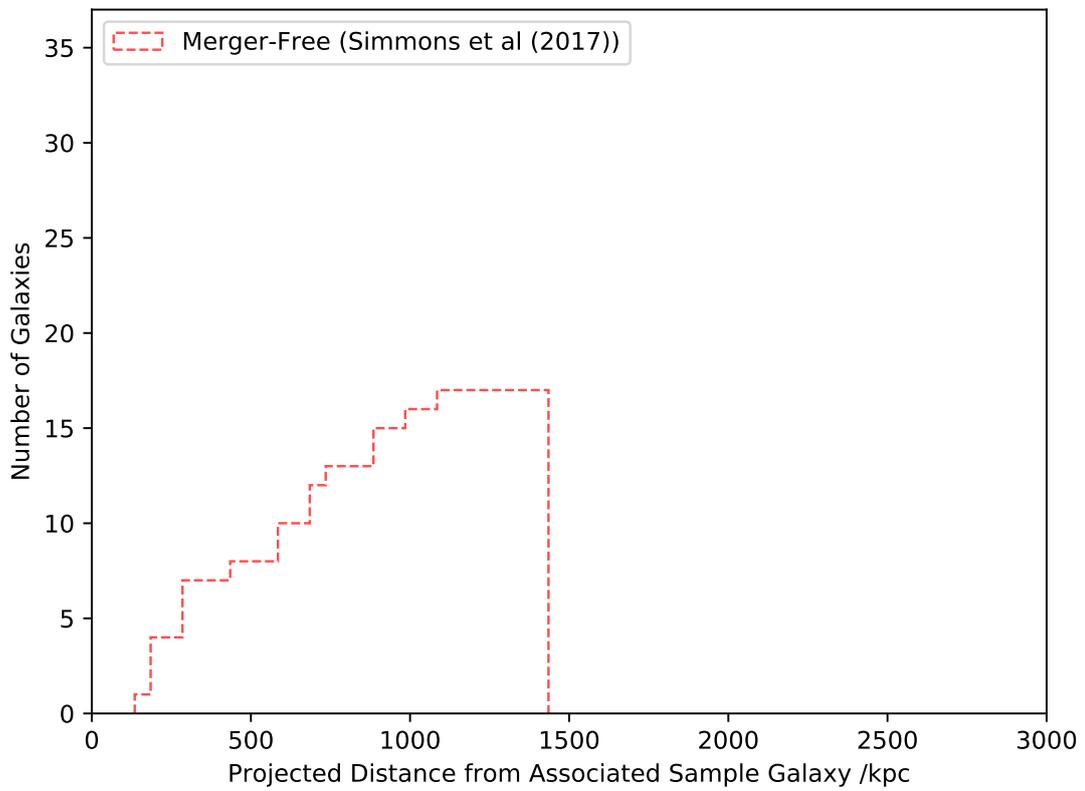
Figure B.10: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Figure B.11: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Simmons et al (2017) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
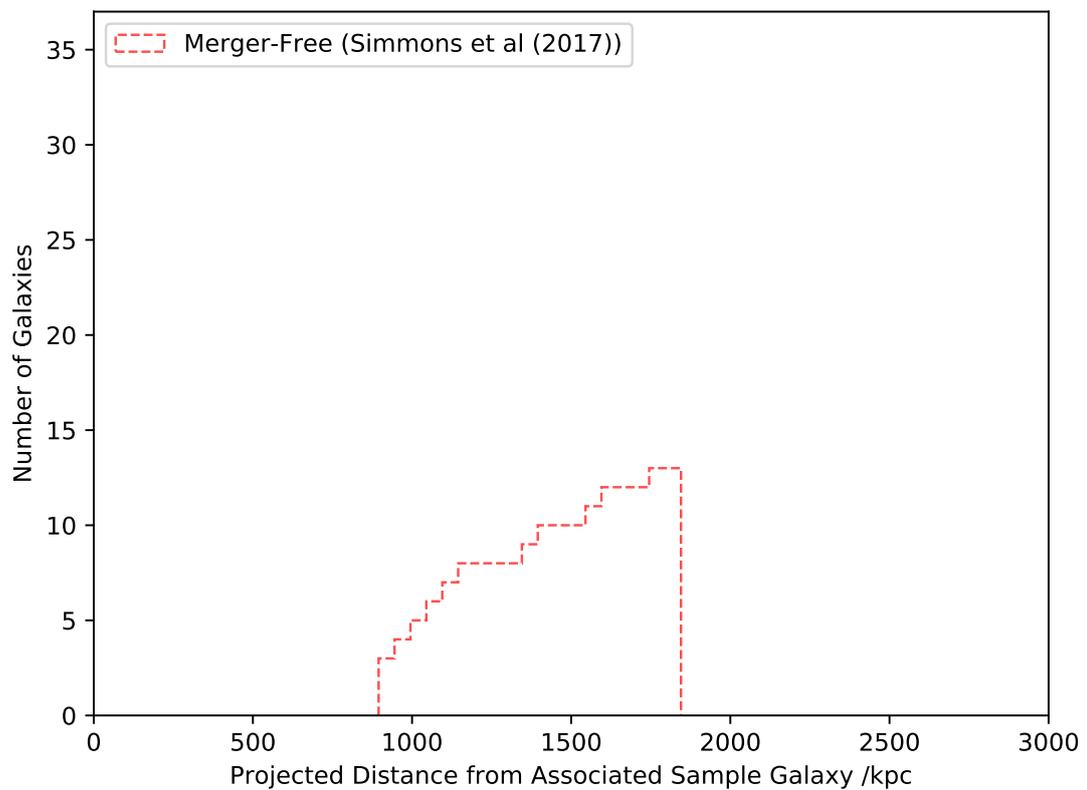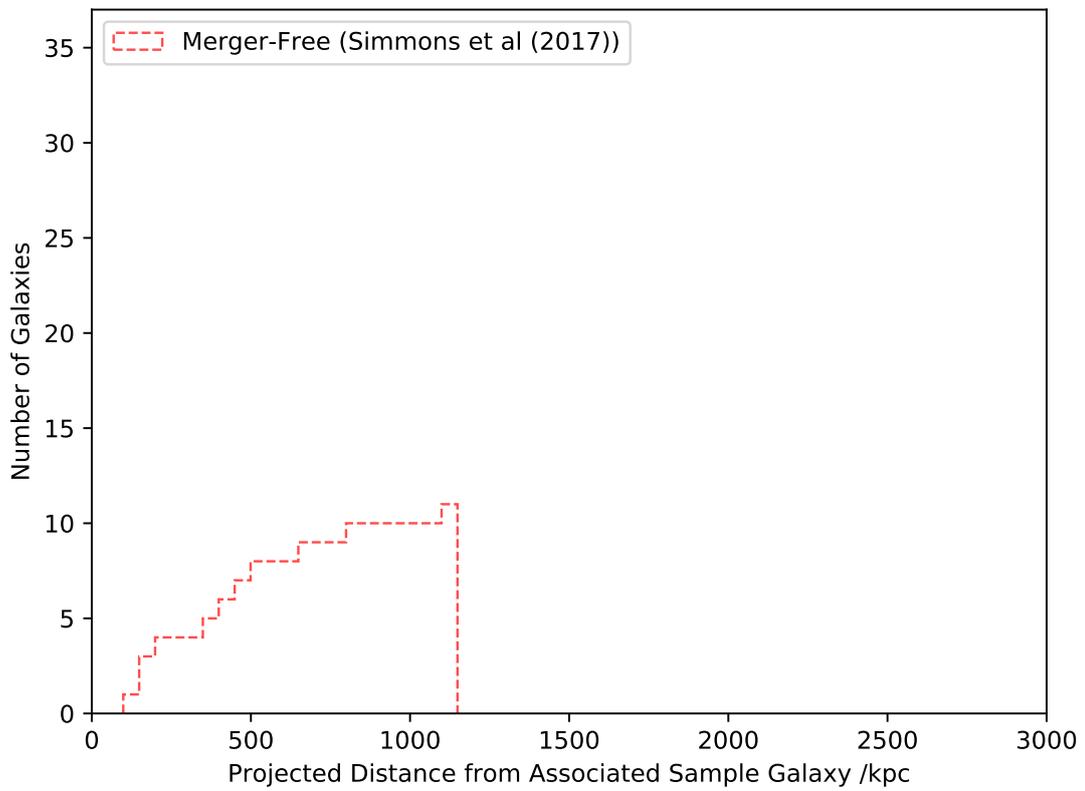
Figure B.12: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Shen et al (2011)) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
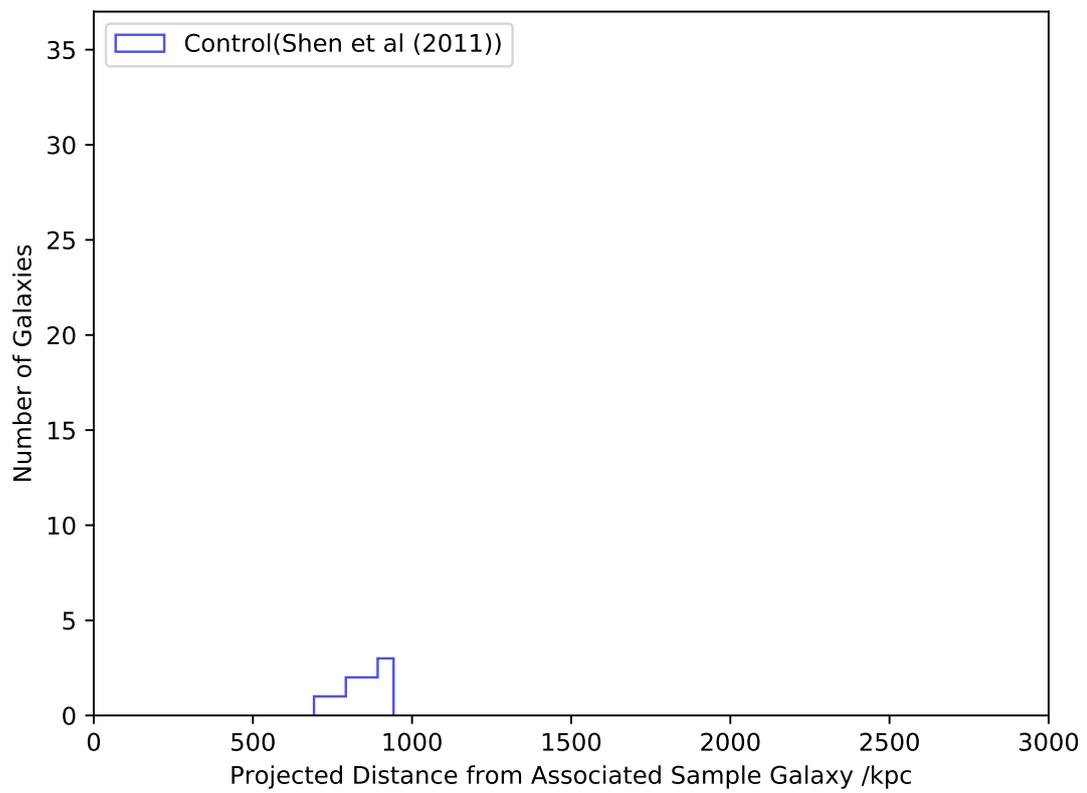
Figure B.13: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Shen et al (2011)) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
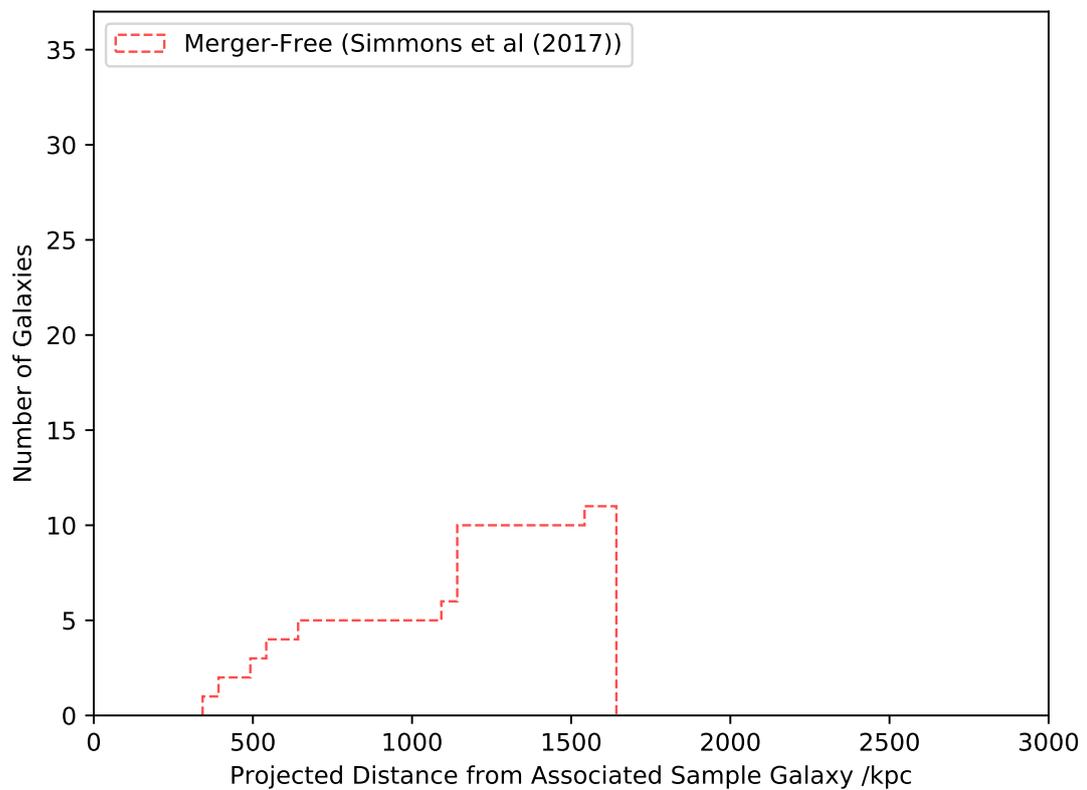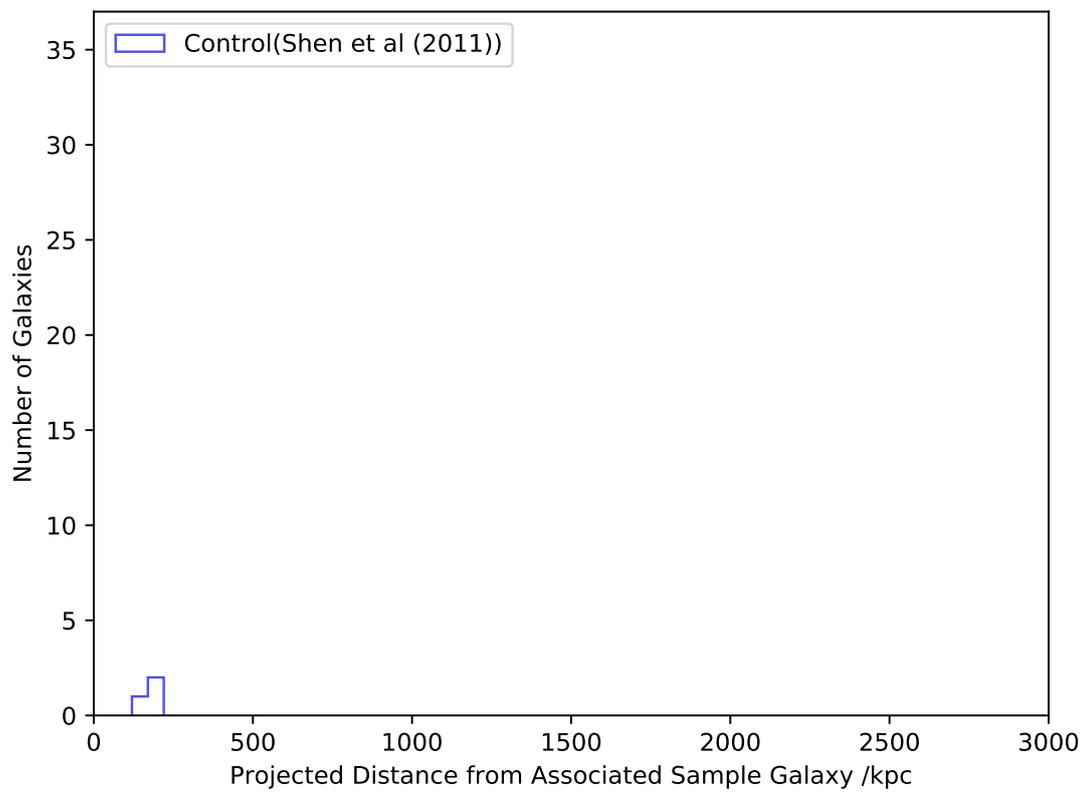
Figure B.14: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Shen et al (2011)) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
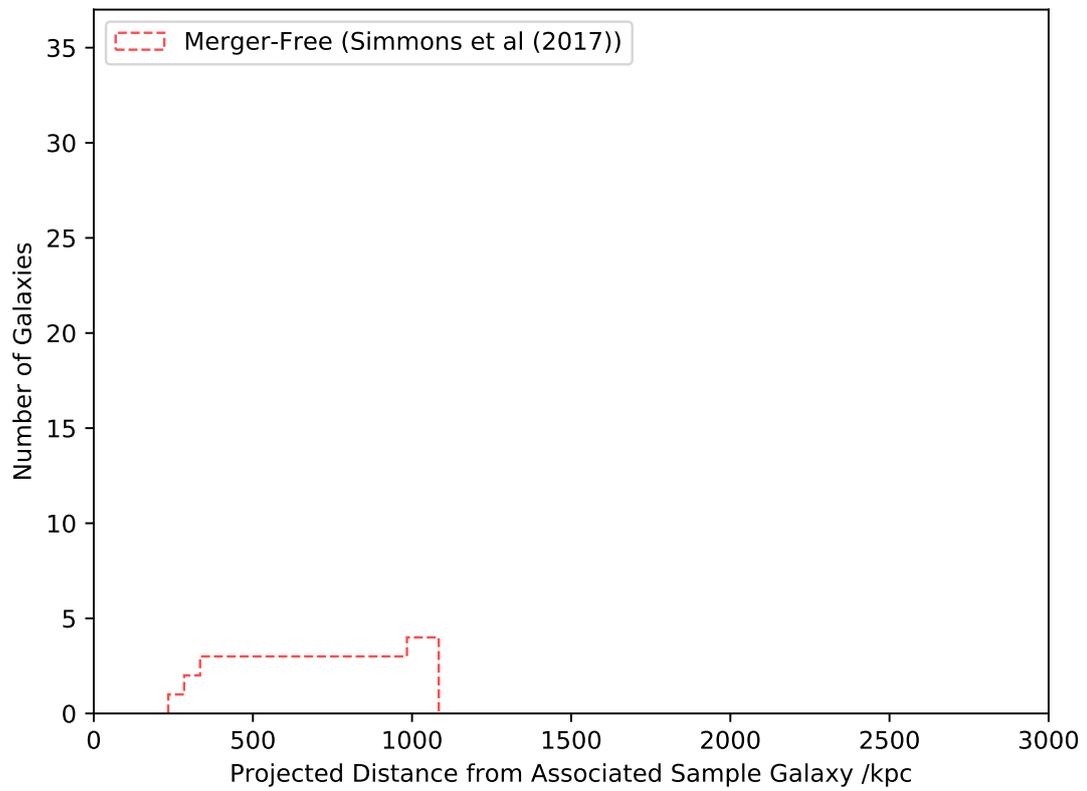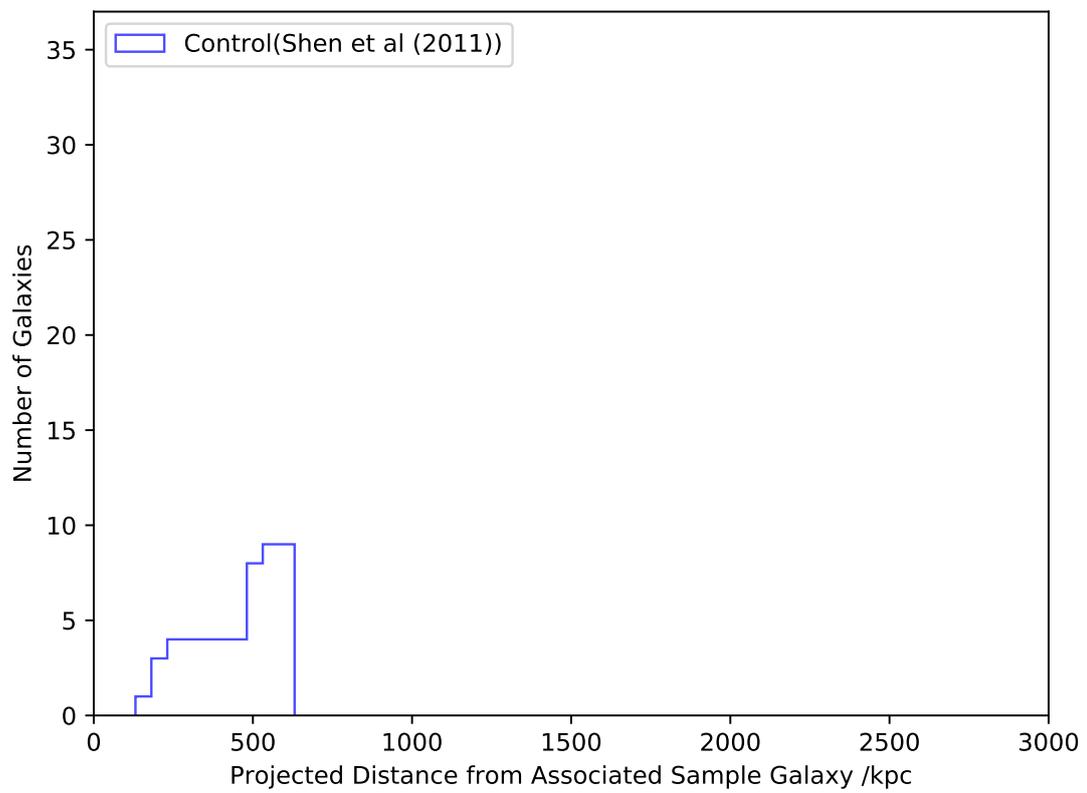
Figure B.15: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Shen et al (2011)) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
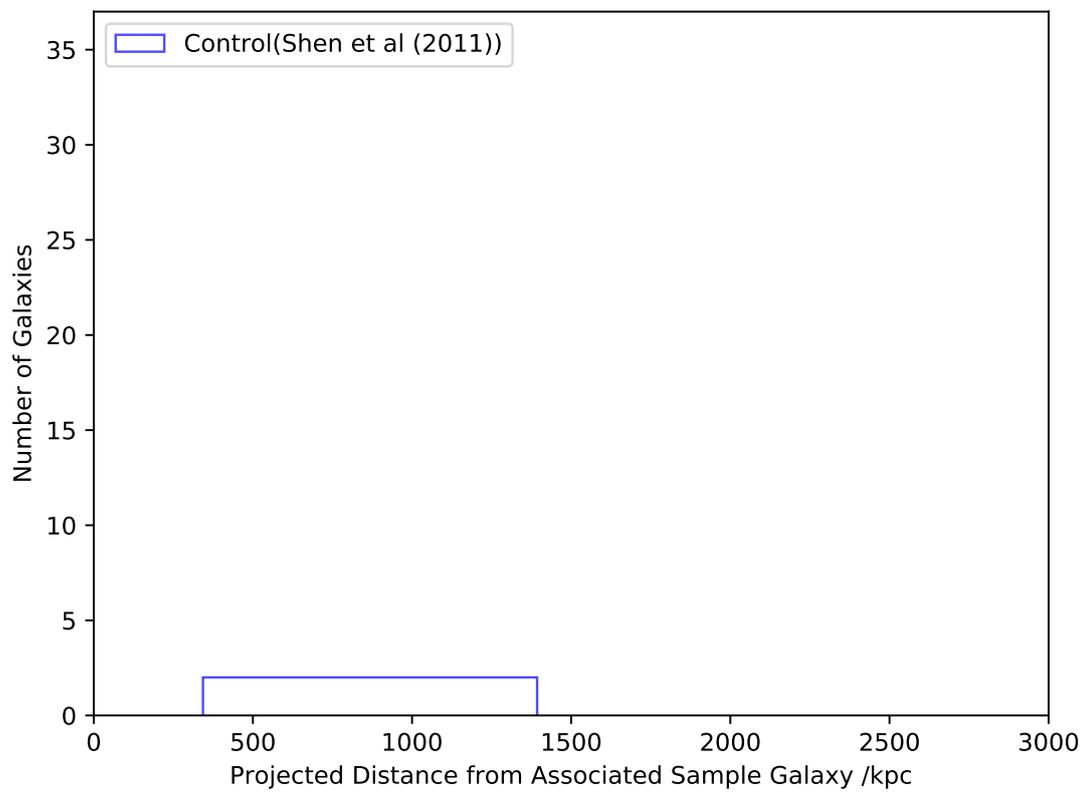
Figure B.16: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Shen et al (2011)) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.
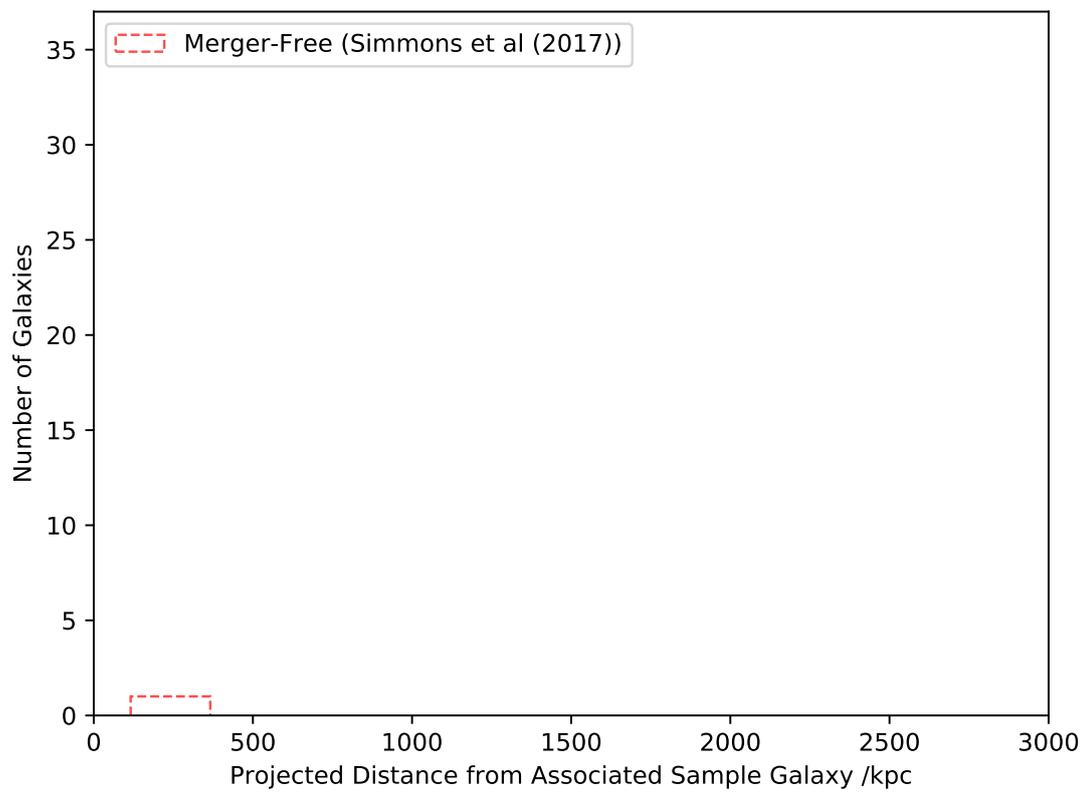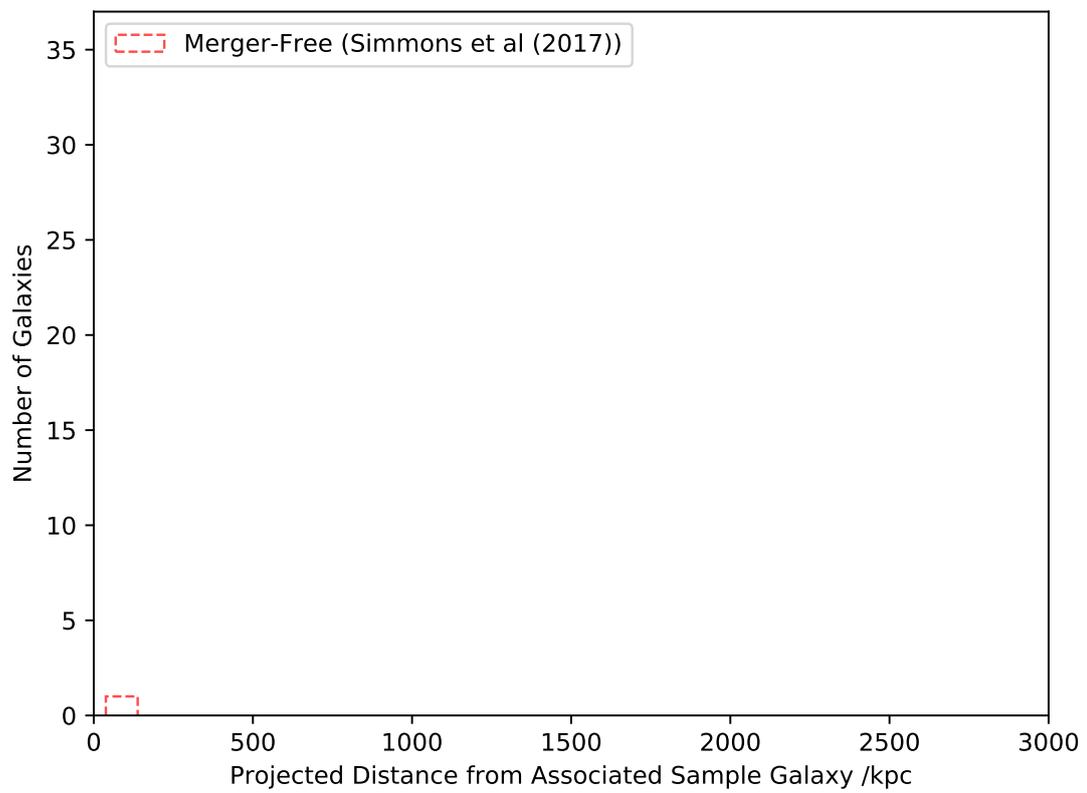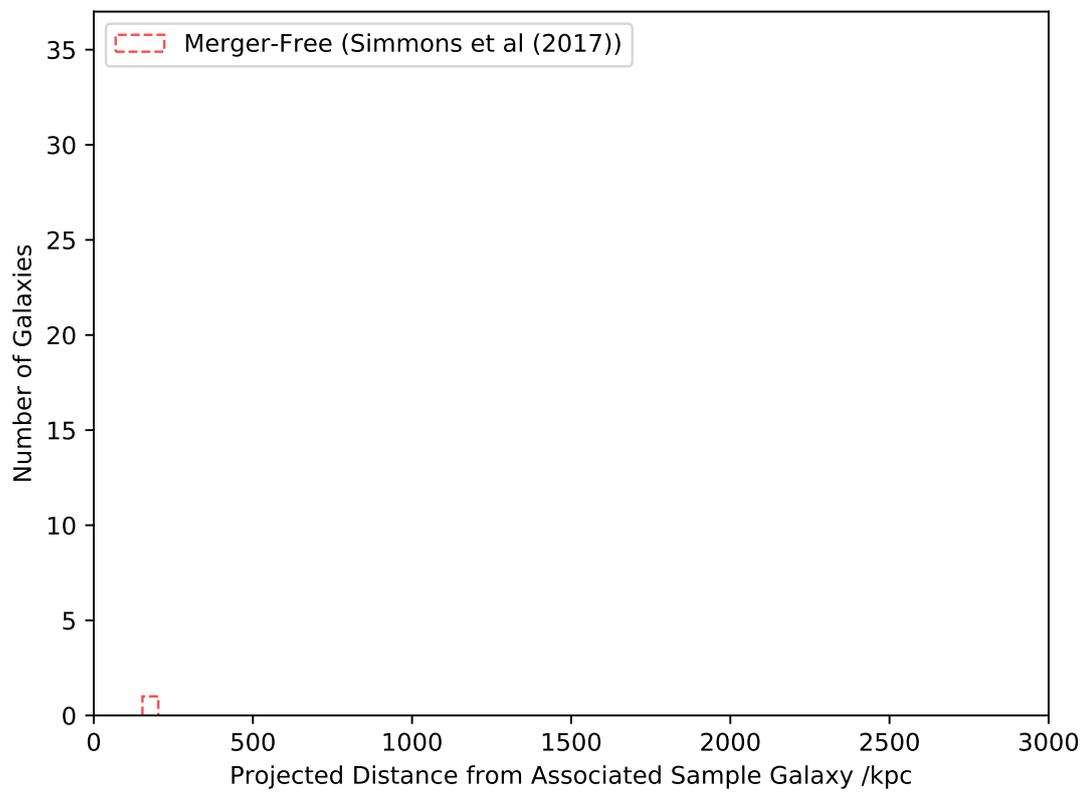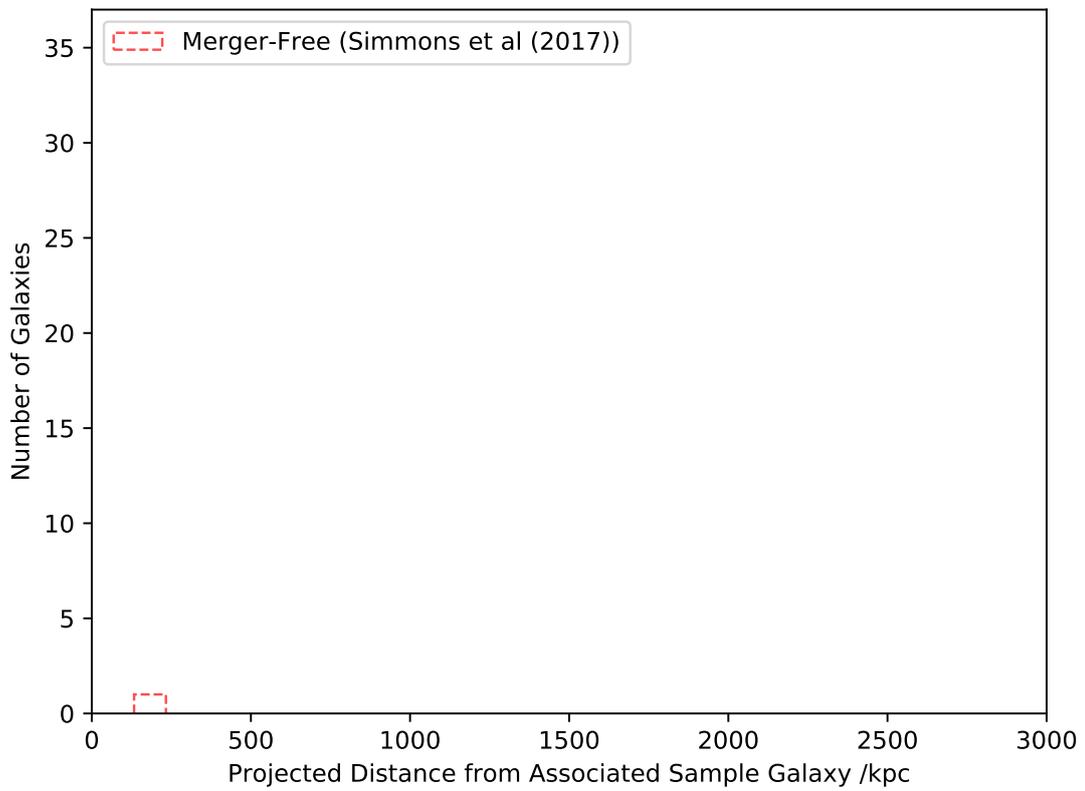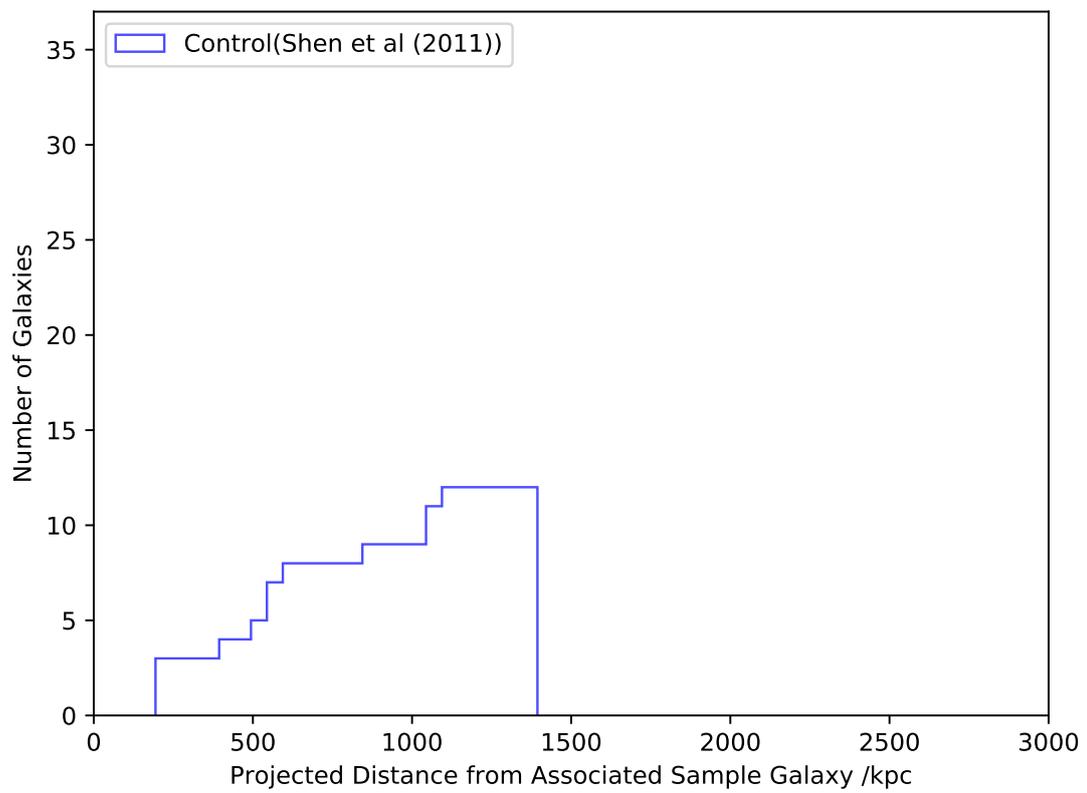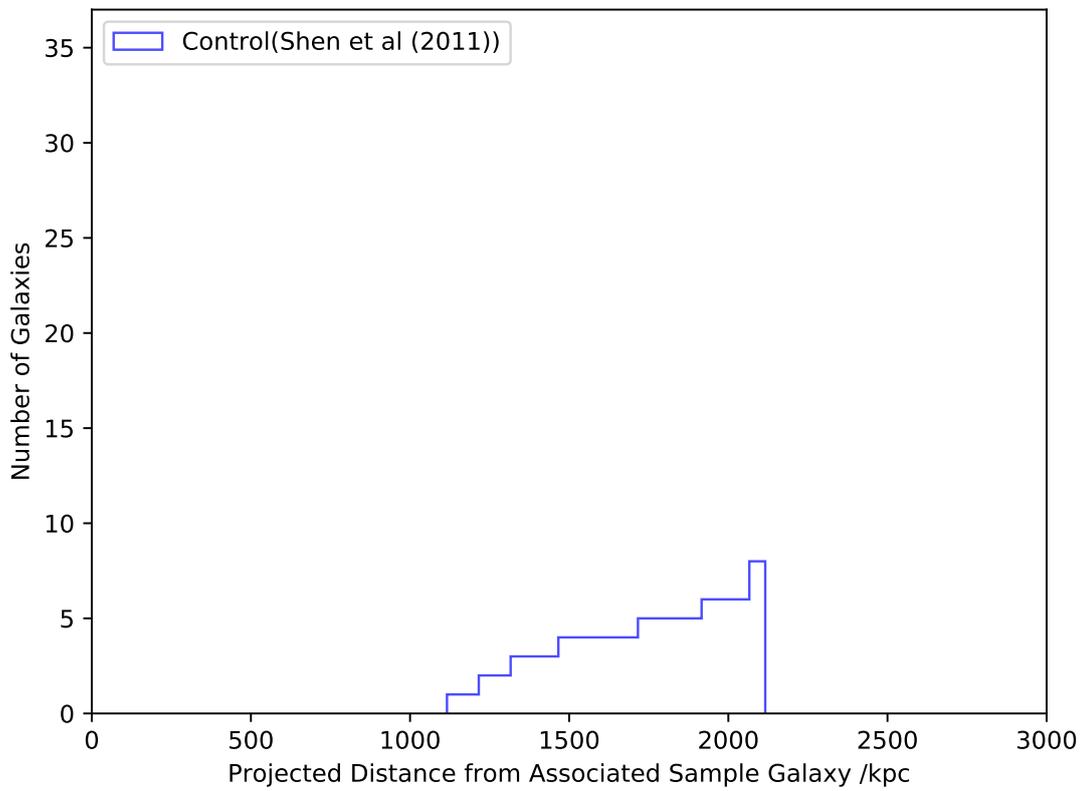
Figure B.17: Cumulative histogram for one group from Yang et al (2007) showing the distances of companion galaxies from their associated sample galaxy from Shen et al (2011)) in kpc. The histogram ends at the distance of the furthest galaxy in the group. The bin size is set at 50 kpc.

Appendix C

An adjustment to the plots and significance levels of some cylinder search results

It was found whilst making corrections to this dissertation that a small number of potential companion galaxies were repeated in the cylinder search results with 1 Mpc and 5 Mpc depths. Whilst this does not affect the general shape of the plots, it does somewhat affect the significance levels of these results, therefore it is necessary to mention this here. The new plots are provided below in Figures C.1, C.2, C.3 and C.4.

In the case of the cylinder search with a 1 Mpc depth, this affected only one galaxy, which was a potential companion to a galaxy in the control sample. This caused the significance level gained from the K-S test to become 0.2σ, with a p-value of 0.847. This result is therefore still not statistically significant.

In the case of the cylinder search with a 5 Mpc depth, this affected 4 galaxies that were potential companions to the merger-free sample and 3 galaxies that were potential companions to the control sample. This caused the significance level gained from the K-S test to become 1.6σ, which would mean that, whilst the differences between samples in this search are of much higher significance than those of the searches with 500 kpc and 1 Mpc depths, and therefore should not be

ignored, this value is not of marginal statistical significance as stated elsewhere in the dissertation. Also, the number of potential companion galaxies in the merger-free sample now peaks in the 400–450 kpc bin rather than in the 450–500 kpc bin as stated elsewhere in this dissertation.
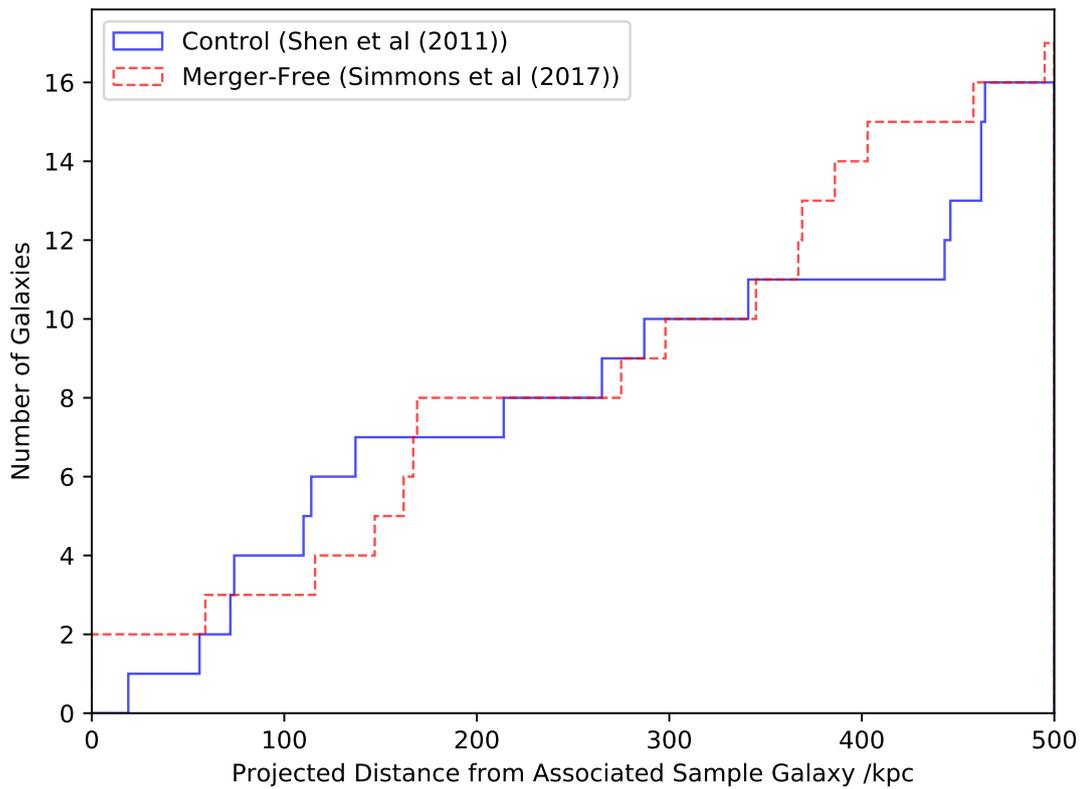
Figure C.1: A cumulative histogram showing the results of a cylinder search for galaxies that are within 500 kpc projected distance of any sample galaxy and within 1 Mpc 'redshift distance' of any sample galaxy. The projected distance (not considering redshift) between each galaxy found in the search and its associated sample galaxy (x) is plotted against cumulative frequency (y). A bin size of 1 kpc is used.
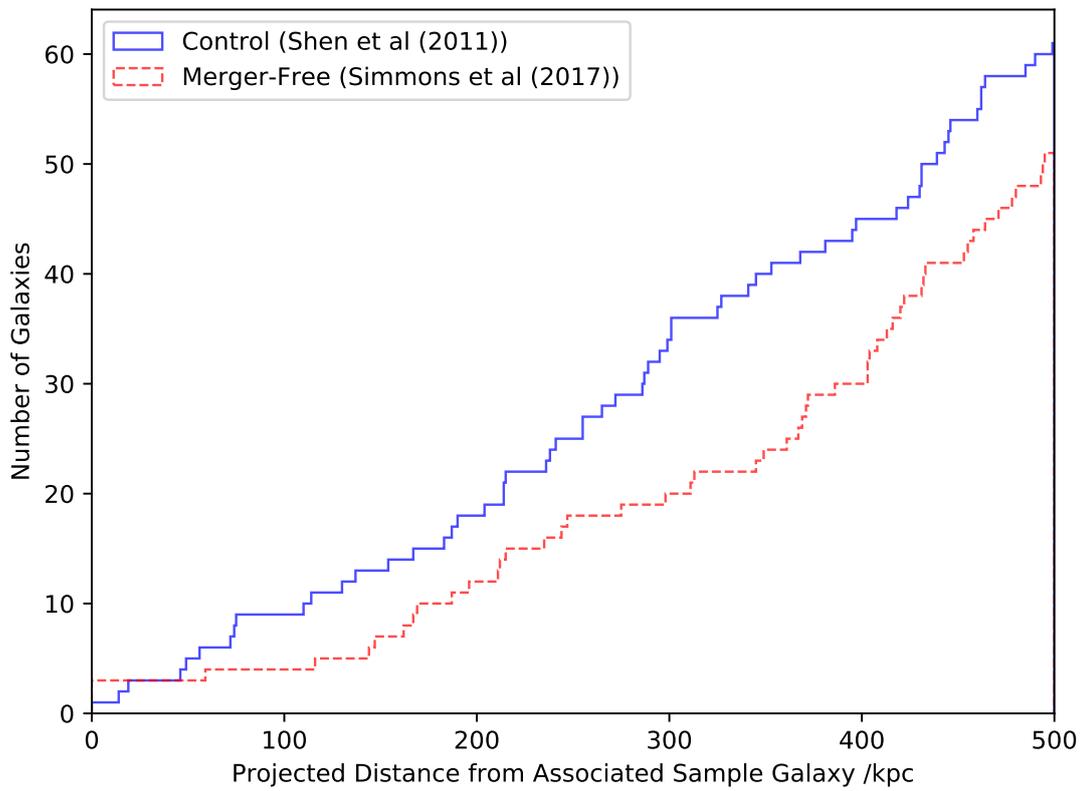
Figure C.2: A cumulative histogram showing the results of a cylinder search for galaxies that are within 500 kpc projected distance of any sample galaxy and within 5 Mpc 'redshift distance' of any sample galaxy. The projected distance (not considering redshift) between each galaxy found in the search and its associated sample galaxy (x) is plotted against cumulative frequency (y). A bin size of 1 kpc is used.
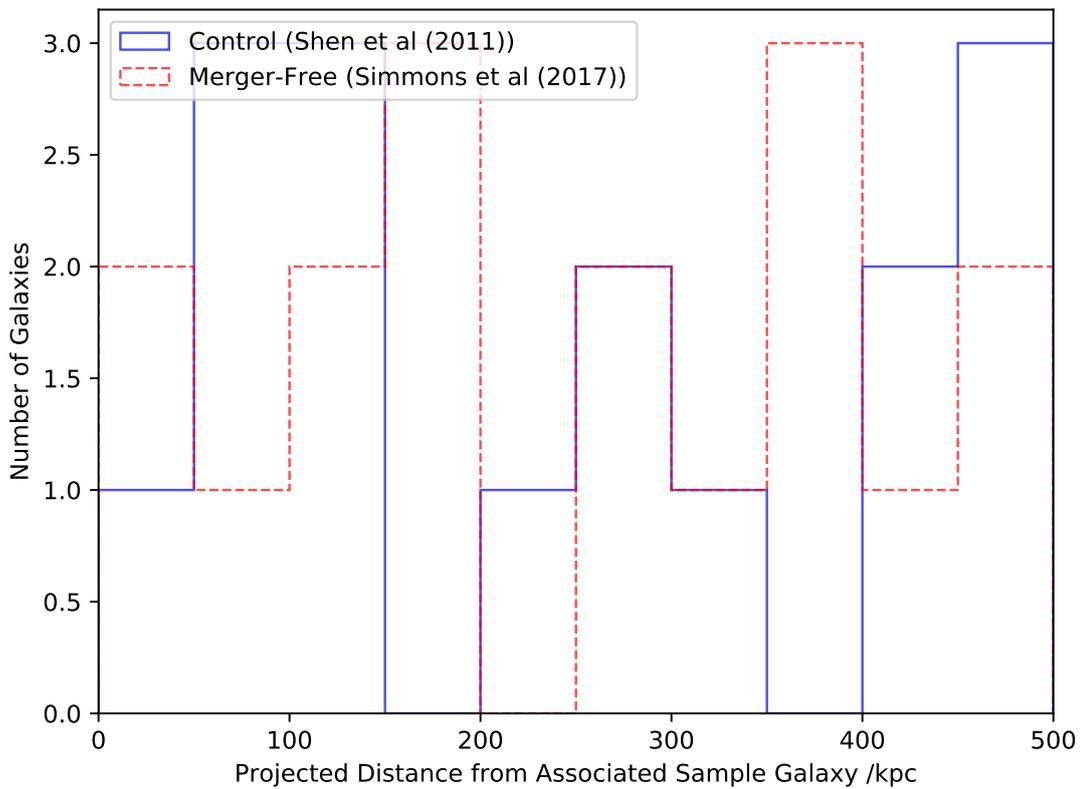
Figure C.3: A non-cumulative histogram showing the distribution of distances of potential companion galaxies from their associated sample galaxies for galaxies within a 1 Mpc 'redshift distance' of their associated sample galaxy and within a 500 kpc projected distance around their sample galaxy. We use distance bins of size 50 kpc.
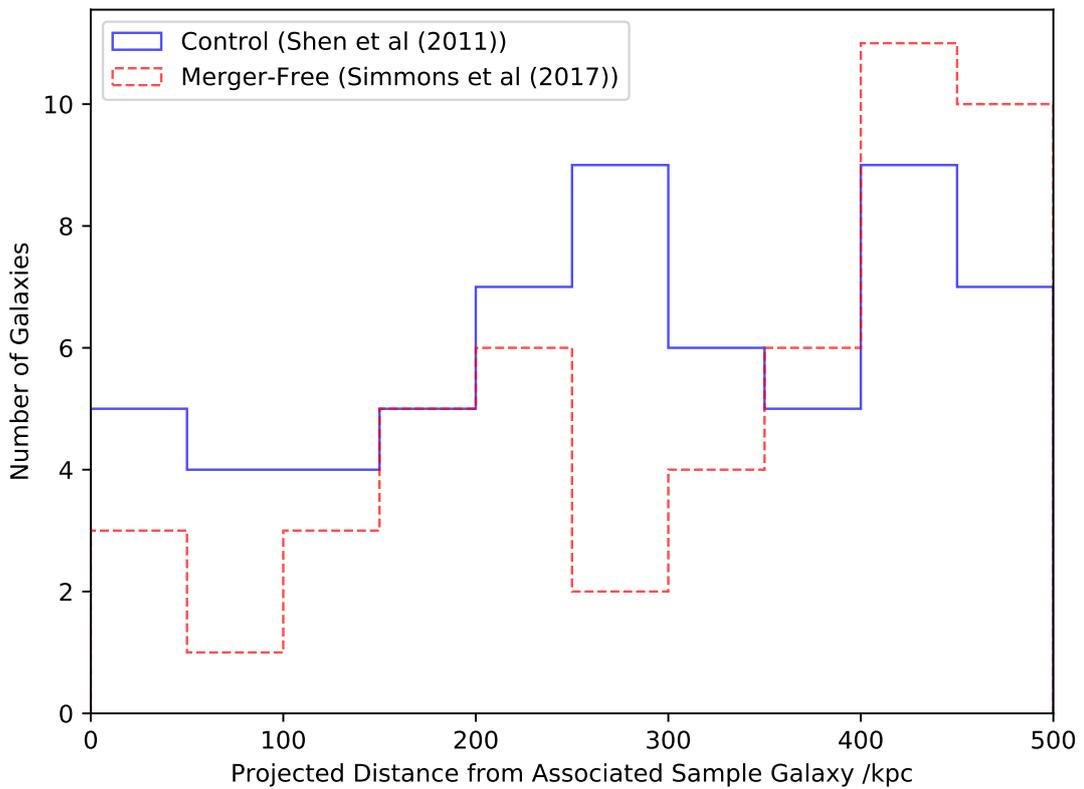
Figure C.4: A non-cumulative histogram showing the distribution of distances of potential companion galaxies from their associated sample galaxies for galaxies within a 5 Mpc 'redshift distance' of their associated sample galaxy and within a 500 kpc projected distance around their sample galaxy. We use distance bins of size 50 kpc.

References

[1] Astropy Collaboration. Astropy: A community Python package for astronomy. A&A, 558:A33, Oct 2013.

[2] Stéfan van der Walt, S. Chris Colbert, and Gaël Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science and Engineering*, 13(2):22–30, Mar 2011.

[3] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python, 2001. [Online; accessed September 9, 2019].

[4] John D. Hunter. Matplotlib: A 2D Graphics Environment. *Computing in Science and Engineering*, 9(3):90–95, May 2007.

[5] Wes McKinney. Data structures for statistical computing in python. In Stéfan van der Walt and Jarrod Millman, editors, *Proceedings of the 9th Python in Science Conference*, pages 51 – 56, 2010.

[6] M. B. Taylor. TOPCAT & STIL: Starlink Table/VOTable Processing Software. In P. Shopbell, M. Britton, and R. Ebert, editors, *Astronomical Data Analysis Software and Systems XIV*, volume 347 of *Astronomical Society of the Pacific Conference Series*, page 29, Dec 2005.

[7] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, et al. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. ApJS, 208(2):20, Oct 2013.

[8] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, et al. Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, page arXiv:1807.06209, Jul 2018.

[9] R. J. Bouwens, G. D. Illingworth, I. Labbe, P. A. Oesch, M. Trenti, C. M. Carollo, et al. A candidate redshift z~10 galaxy and rapid changes in that population at an age of 500Myr. Nature, 469(7331):504–507, Jan 2011.

[10] Richard S. Ellis, Ross J. McLure, James S. Dunlop, Brant E. Robertson, Yoshiaki Ono, Matthew A. Schenker, et al. The Abundance of Star-forming Galaxies in the Redshift Range 8.5-12: New Results from the 2012 Hubble Ultra Deep Field Campaign. ApJ, 763(1):L7, Jan 2013.

[11] David Sobral, Jorryt Matthee, Behnam Darvish, Daniel Schaerer, Bahram Mobasher, Huub J. A. Röttgering, et al. Evidence for PopIII-like Stellar Populations in the Most Luminous Lyman-α Emitters at the Epoch of Reionization: Spectroscopic Confirmation. ApJ, 808(2):139, Aug 2015.

[12] John Magorrian, Scott Tremaine, Douglas Richstone, Ralf Bender, Gary Bower, Alan Dressler, et al. The Demography of Massive Dark Objects in Galaxy Centers. AJ, 115(6):2285–2305, Jun 1998.

[13] Laura Ferrarese and Holland Ford. Supermassive Black Holes in Galactic Nuclei: Past, Present and Future Research. Space Sci. Rev., 116(3-4):523–624, Feb 2005.

[14] John Kormendy and Luis C. Ho. Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. ARA&A, 51(1):511–653, Aug 2013.

[15] E. P. Hubble. A general study of diffuse galactic nebulae. ApJ, 56:162–199, Oct 1922.

[16] E. P. Hubble. Extragalactic nebulae. ApJ, 64:321–369, Dec 1926.

[17] E. P. Hubble. *Realm of the Nebulae.* 1936.

[18] J. Kormendy. Observations of Galaxy Structure and Dynamics. *Saas-Fee Advanced Course*, 12:115, Jan 1982.

[19] R. J. Smethurst, C. J. Lintott, B. D. Simmons, K. Schawinski, P. J. Marshall, S. Bamford, et al. Galaxy Zoo: evidence for diverse star formation histories through the green valley. MNRAS, 450(1):435–453, Jun 2015.

[20] Alar Toomre and Juri Toomre. Galactic Bridges and Tails. ApJ, 178:623–666, Dec 1972.

[21] David Sobral, Philip N. Best, Ian Smail, James E. Geach, Michele Cirasuolo, Timothy Garn, et al. The dependence of star formation activity on environment and stellar mass at z∼1 from the HiZELS-Hα survey. MNRAS, 411(1):675–692, Feb 2011.

[22] A. Dressler. Galaxy morphology in rich clusters: implications for the formation and evolution of galaxies. ApJ, 236:351–365, Mar 1980.

[23] Luigi Guzzo, Michael A. Strauss, Karl B. Fisher, Riccardo Giovanelli, and Martha P. Haynes. Redshift-Space Distortions and the Real-Space Clustering of Different Galaxy Types. ApJ, 489(1):37–48, Nov 1997.

[24] Tomotsugu Goto, Chisato Yamauchi, Yutaka Fujita, Sadanori Okamura, Maki Sekiguchi, Ian Smail, et al. The morphology-density relation in the Sloan Digital Sky Survey. MNRAS, 346(2):601–614, Dec 2003.

[25] Arjen van der Wel. The Dependence of Galaxy Morphology and Structure on Environment and Stellar Mass. ApJ, 675(1):L13, Mar 2008.

[26] Steven P. Bamford, Robert C. Nichol, Ivan K. Baldry, Kate Land, Chris J. Lintott, Kevin Schawinski, et al. Galaxy Zoo: the dependence of morphology and colour on environment*. MNRAS, 393(4):1324–1352, Mar 2009.

[27] Ramin A. Skibba, Steven P. Bamford, Robert C. Nichol, Chris J. Lintott, Dan Andreescu, Edward M. Edmondson, et al. Galaxy Zoo: disentangling the environmental dependence of morphology and colour. MNRAS, 399(2):966–982, Oct 2009.

[28] Ramin A. Skibba, Karen L. Masters, Robert C. Nichol, Idit Zehavi, Ben Hoyle, Edward M. Edmondson, et al. Galaxy Zoo: the environmental dependence of bars and bulges in disc galaxies. MNRAS, 423(2):1485–1502, Jun 2012.

[29] Hong Bae Ann. Environment Dependence of Disk Morphology of Spiral Galaxies. *Journal of Korean Astronomical Society*, 47(1):1–13, Feb 2014.

[30] G. Fasano, B. M. Poggianti, D. Bettoni, M. D'Onofrio, A. Dressler, B. Vulcani, et al. Morphological fractions of galaxies in WINGS clusters: revisiting the morphology-density paradigm. MNRAS, 449(4):3927–3944, Jun 2015.

[31] Ana Paulino-Afonso, David Sobral, Behnam Darvish, Bruno Ribeiro, Arjen van der Wel, John Stott, et al. VIS^{3}COS. II. Nature and nurture in galaxy structure and morphology. A&A, 630:A57, Oct 2019.

[32] J. A. Sellwood and A. Wilkinson. Dynamics of barred galaxies. *Reports on Progress in Physics*, 56(2):173–256, Feb 1993.

[33] Allan Sandage. *The Hubble Atlas of Galaxies*. 1961.

[34] E. Athanassoula. On the nature of bulges in general and of box/peanut bulges in particular: input from N-body simulations. MNRAS, 358(4):1477–1488, Apr 2005.

[35] John Kormendy and Jr. Kennicutt, Robert C. Secular Evolution and the Formation of Pseudobulges in Disk Galaxies. ARA&A, 42(1):603–683, Sep 2004.

[36] R. Genzel, A. Burkert, N. Bouché, G. Cresci, N. M. Förster Schreiber, A. Shapley, et al. From Rings to Bulges: Evidence for Rapid Secular Galaxy Evolution at z ~2 from Integral Field Spectroscopy in the SINS Survey. ApJ, 687(1):59–77, Nov 2008.

[37] David Izquierdo-Villalba, Silvia Bonoli, Daniele Spinoso, Yetli Rosas-Guevara, Bruno M. B. Henriques, and Carlos Hernández-Monteagudo. The build-up of pseudo-bulges in a hierarchical universe. MNRAS, 488(1):609–632, Sep 2019.

[38] J. Kormendy. Rotation of the bulge components of barred galaxies. ApJ, 257:75–88, Jun 1982.

[39] Marta Volonteri, Francesco Haardt, and Piero Madau. The Assembly and Merging History of Supermassive Black Holes in Hierarchical Models of Galaxy Formation. ApJ, 582(2):559–573, Jan 2003.

[40] Marta Volonteri and Martin J. Rees. Rapid Growth of High-Redshift Black Holes. ApJ, 633(2):624–629, Nov 2005.

[41] Marta Volonteri. Formation of supermassive black holes. A&A Rev., 18(3):279–315, Jul 2010.

[42] Martin G. Haehnelt and Martin J. Rees. The formation of nuclei in newly formed galaxies and the evolution of the quasar population. MNRAS, 263(1):168–178, Jul 1993.

[43] Abraham Loeb and Frederic A. Rasio. Collapse of Primordial Gas Clouds and the Formation of Quasar Black Holes. ApJ, 432:52, Sep 1994.

[44] Daniel J. Eisenstein and A. Loeb. Origin of Quasar Progenitors from the Collapse of Low-Spin Cosmological Perturbations. ApJ, 443:11, Apr 1995.

[45] Volker Bromm and Abraham Loeb. Formation of the First Supermassive Black Holes. ApJ, 596(1):34–46, Oct 2003.

[46] Savvas M. Koushiappas, James S. Bullock, and Avishai Dekel. Massive black hole seeds from low angular momentum material. MNRAS, 354(1):292–304, Oct 2004.

[47] Giuseppe Lodato and Priyamvada Natarajan. Supermassive black hole formation during the assembly of pre-galactic discs. MNRAS, 371(4):1813–1823, Oct 2006.

[48] Isaac Shlosman, Juhan Frank, and Mitchell C. Begelman. Bars within bars: a mechanism for fuelling active galactic nuclei. Nature, 338(6210):45–47, Mar 1989.

[49] Mitchell C. Begelman, Marta Volonteri, and Martin J. Rees. Formation of supermassive black holes by direct collapse in pre-galactic haloes. MNRAS, 370(1):289–298, Jul 2006.

[50] Mitchell C. Begelman and Martin J. Rees. The fate of dense stellar systems. MNRAS, 185:847–860, Dec 1978.

[51] Toshikazu Ebisuzaki, Junichiro Makino, Takeshi Go Tsuru, Yoko Funato, Simon Portegies Zwart, Piet Hut, et al. Missing Link Found? The "Runaway" Path to Supermassive Black Holes. ApJ, 562(1):L19–L22, Nov 2001.

[52] Gregory A. Shields. A Brief History of Active Galactic Nuclei. PASP, 111(760):661–678, Jun 1999.

[53] Yue Shen, Gordon T. Richards, Michael A. Strauss, Patrick B. Hall, Donald P. Schneider, Stephanie Snedden, et al. A Catalog of Quasar Properties from Sloan Digital Sky Survey Data Release 7. *The Astrophysical Journal Supplement Series*, 194(2):45, Jun 2011.

[54] Brooke D. Simmons, Chris Lintott, Kevin Schawinski, Edward C. Moran, Anna Han, Sugata Kaviraj, et al. Galaxy Zoo: bulgeless galaxies with growing black holes. MNRAS, 429(3):2199–2211, Mar 2013.

[55] B. D. Simmons, R. J. Smethurst, and C. Lintott. Supermassive black holes in disc-dominated galaxies outgrow their bulges and co-evolve with their host galaxies. MNRAS, 470(2):1559–1569, Sep 2017.

[56] J. G. Bolton, G. J. Stanley, and O. B. Slee. Positions of Three Discrete Sources of Galactic Radio-Frequency Radiation. Nature, 164(4159):101–102, Jul 1949.

[57] Clive Tadhunter. Radio AGN in the local universe: unification, triggering and evolution. A&A Rev., 24(1):10, Jun 2016.

[58] E. E. Salpeter. Accretion of Interstellar Matter by Massive Objects. ApJ, 140:796–800, Aug 1964.

[59] D. Lynden-Bell and M. J. Rees. On quasars, dust and the galactic centre. MNRAS, 152:461, Jan 1971.

[60] Martin J. Rees. Black Hole Models for Active Galactic Nuclei. ARA&A, 22:471–506, Jan 1984.

[61] Mitchell C. Begelman, Roger D. Blandford, and Martin J. Rees. Theory of extragalactic radio sources. *Reviews of Modern Physics*, 56(2):255–351, Apr 1984.

[62] John Kormendy and Douglas Richstone. Inward Bound—The Search For Supermassive Black Holes In Galactic Nuclei. ARA&A, 33:581, Jan 1995.

[63] M. P. Véron-Cetty and P. Véron. The emission line spectrum of active galactic nuclei and the unifying scheme. A&A Rev., 10:81–133, Jan 2000.

[64] Event Horizon Telescope Collaboration, Kazunori Akiyama, Antxon Alberdi, Walter Alef, Keiichi Asada, Rebecca Azulay, et al. First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. ApJ, 875(1):L1, Apr 2019.

[65] R. D. Blandford and M. J. Rees. A "twin-exhaust" model for double radio sources. MNRAS, 169:395–415, Dec 1974.

[66] Jose-Luis Gomez. *Observations and Simulations of Relativistic Jets*, volume 589, page 169. 2002.

[67] C. Urry. AGN Unification: An Update. In Gordon T. Richards and Patrick B. Hall, editors, *AGN Physics with the Sloan Digital Sky Survey*, volume 311 of *Astronomical Society of the Pacific Conference Series*, page 49, Jun 2004.

[68] Ezequiel Treister and C. Megan Urry. Active Galactic Nuclei Unification and the X-Ray Background. ApJ, 630(1):115–121, Sep 2005.

[69] B. M. Peterson. *The Broad-Line Region in Active Galactic Nuclei*, volume 693, page 77. 2006.

[70] Thaisa Storchi-Bergmann and Allan Schnorr-Müller. Observational constraints on the feeding of supermassive black holes. *Nature Astronomy*, 3:48–61, Jan 2019.

[71] Vivienne Wild, Guinevere Kauffmann, Tim Heckman, Stéphane Charlot, Gerard Lemson, Jarle Brinchmann, et al. Bursty stellar populations and obscured active galactic nuclei in galaxy bulges. MNRAS, 381(2):543–572, Oct 2007.

[72] João Calhau, David Sobral, Andra Stroe, Philip Best, Ian Smail, Bret Lehmer, et al. The growth of typical star-forming galaxies and their supermassive black holes across cosmic time since z 2. MNRAS, 464(1):303–311, Jan 2017.

[73] James Aird, Alison L. Coil, John Moustakas, Michael R. Blanton, Scott M. Burles, Richard J. Cool, et al. PRIMUS: The Dependence of AGN Accretion on Host Stellar Mass and Color. ApJ, 746(1):90, Feb 2012.

[74] J. Aird, A. L. Coil, and A. Georgakakis. X-rays across the galaxy population - III. The incidence of AGN as a function of star formation rate. MNRAS, 484(3):4360–4378, Apr 2019.

[75] Mauricio Cisternas, Knud Jahnke, Angela Bongiorno, Katherine J. Inskip, Chris D. Impey, Anton M. Koekemoer, et al. Secular Evolution and a Non-evolving Black-hole-to-galaxy Mass Ratio in the Last 7 Gyr. ApJ, 741(1):L11, Nov 2011.

[76] Francine R. Marleau, Dominic Clancy, and Matteo Bianconi. The ubiquity of supermassive black holes in the Hubble sequence. MNRAS, 435(4):3085–3095, Nov 2013.

[77] Alessandro Marconi and Leslie K. Hunt. The Relation between Black Hole Mass, Bulge Mass, and Near-Infrared Luminosity. ApJ, 589(1):L21–L24, May 2003.

[78] Nadine Häring and Hans-Walter Rix. On the Black Hole Mass-Bulge Mass Relation. ApJ, 604(2):L89–L92, Apr 2004.

[79] Darren J. Croton. Evolution in the black hole mass-bulge mass relation: a theoretical perspective. MNRAS, 369(4):1808–1812, Jul 2006.

[80] Vivienne Wild, Timothy Heckman, and Stéphane Charlot. Timing the starburst-AGN connection. MNRAS, 405(2):933–947, Jun 2010.

[81] Zachary Schutte, Amy Reines, and Jenny Greene. The Black Hole – Bulge Mass Relation including Dwarf Galaxies Hosting Active Galactic Nuclei. *arXiv e-prints*, page arXiv:1908.00020, Jul 2019.

[82] Laura Ferrarese and David Merritt. A Fundamental Relation between Supermassive Black Holes and Their Host Galaxies. ApJ, 539(1):L9–L12, Aug 2000.

[83] Karl Gebhardt, Ralf Bender, Gary Bower, Alan Dressler, S. M. Faber, Alexei V. Filippenko, et al. A Relationship between Nuclear Black Hole Mass and Galaxy Velocity Dispersion. ApJ, 539(1):L13–L16, Aug 2000.

[84] John Kormendy and Karl Gebhardt. Supermassive black holes in galactic nuclei. In J. Craig Wheeler and Hugo Martel, editors, *20th Texas Symposium on relativistic astrophysics*, volume 586 of *American Institute of Physics Conference Series*, pages 363–381, Oct 2001.

[85] Nicholas J. McConnell and Chung-Pei Ma. Revisiting the Scaling Relations of Black Hole Masses and Host Galaxy Properties. ApJ, 764(2):184, Feb 2013.

[86] Darren J. Croton, Volker Springel, Simon D. M. White, G. De Lucia, C. S. Frenk, L. Gao, et al. The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. MNRAS, 365(1):11–28, Jan 2006.

[87] D. B. Sanders, B. T. Soifer, J. H. Elias, B. F. Madore, K. Matthews, G. Neugebauer, et al. Ultraluminous Infrared Galaxies and the Origin of Quasars. ApJ, 325:74, Feb 1988.

[88] Philip F. Hopkins and Lars Hernquist. Fueling Low-Level AGN Activity through Stochastic Accretion of Cold Gas. ApJS, 166(1):1–36, Sep 2006.

[89] Chien Y. Peng. How Mergers May Affect the Mass Scaling Relation between Gravitationally Bound Systems. ApJ, 671(2):1098–1107, Dec 2007.

[90] Philip F. Hopkins, Rachel S. Somerville, Lars Hernquist, Thomas J. Cox, Brant Robertson, and Yuexing Li. The Relation between Quasar and Merging Galaxy Luminosity Functions and the Merger-driven Star Formation History of the Universe. ApJ, 652(2):864–888, Dec 2006.

[91] Mauricio Cisternas, Knud Jahnke, Katherine J. Inskip, Jeyhan Kartaltepe, Anton M. Koekemoer, Thorsten Lisker, et al. The Bulk of the Black Hole Growth Since z ~1 Occurs in a Secular Universe: No Major Merger-AGN Connection. ApJ, 726(2):57, Jan 2011.

[92] Sara L. Ellison, Akshara Viswanathan, David R. Patton, Connor Bottrell, Alan W. McConnachie, Stephen Gwyn, et al. A definitive merger-AGN connection at z∼ 0 with CFIS: mergers have an excess of AGN and AGN hosts are more frequently disturbed. MNRAS, 487(2):2491–2504, Aug 2019.

[93] Thaisa Storchi-Bergmann. The Coevolution of Galaxies and Supermassive Black Holes in the Near Universe. *Brazilian Journal of Physics*, 43(5-6):383–388, Dec 2013.

[94] G. Martin, S. Kaviraj, M. Volonteri, B. D. Simmons, J. E. G. Devriendt, C. J. Lintott, et al. Normal black holes in bulge-less galaxies: the largely quiescent, merger-free growth of black holes over cosmic time. *Monthly Notices of the Royal Astronomical Society*, 476(2):2801–2812, May 2018.

[95] R. J. Smethurst, B. D. Simmons, C. J. Lintott, and J. Shanahan. Secularly powered outflows from AGNs: the dominance of non-merger driven supermassive black hole growth. MNRAS, 489(3):4016–4031, Nov 2019.

[96] Gwang-Ho Lee, Jong-Hak Woo, Myung Gyoon Lee, Ho Seong Hwang, Jong Chul Lee, Jubee Sohn, et al. Do Bars Trigger Activity in Galactic Nuclei? ApJ, 750(2):141, May 2012.

[97] Seulhee Oh, Kyuseok Oh, and Sukyoung K. Yi. Bar Effects on Central Star Formation and Active Galactic Nucleus Activity. ApJS, 198(1):4, Jan 2012.

[98] N. Menci, F. Fiore, A. Bongiorno, and A. Lamastra. Relative growth of black holes and the stellar components of galaxies. A&A, 594:A99, Oct 2016.

[99] Min Du, Victor P. Debattista, Juntai Shen, Luis C. Ho, and Peter Erwin. Black Hole Growth in Disk Galaxies Mediated by the Secular Evolution of Short Bars. ApJ, 844(2):L15, Aug 2017.

[100] Gordon T. Richards, Xiaohui Fan, Heidi Jo Newberg, Michael A. Strauss, Daniel E. Vanden Berk, Donald P. Schneider, et al. Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Quasar Sample. AJ, 123(6):2945–2975, Jun 2002.

[101] F. E. Bauer, D. M. Alexander, W. N. Brandt, D. P. Schneider, E. Treister, A. E. Hornschemeier, et al. The Fall of Active Galactic Nuclei and the Rise of Star-forming Galaxies: A Close Look at the Chandra Deep Field X-Ray Number Counts. AJ, 128(5):2048–2065, Nov 2004.

[102] Daniel Stern, Peter Eisenhardt, Varoujan Gorjian, Christopher S. Kochanek, Nelson Caldwell, Daniel Eisenstein, et al. Mid-Infrared Selection of Active Galaxies. ApJ, 631(1):163–168, Sep 2005.

[103] R. Edelson and M. Malkan. Reliable Identifications of Active Galactic Nuclei from the WISE, 2MASS, and ROSAT All-Sky Surveys. ApJ, 751(1):52, May 2012.

[104] Edward L. Wright, Peter R. M. Eisenhardt, Amy K. Mainzer, Michael E. Ressler, Roc M. Cutri, Thomas Jarrett, et al. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. AJ, 140(6):1868–1881, Dec 2010.

[105] M. F. Skrutskie, R. M. Cutri, R. Stiening, M. D. Weinberg, S. Schneider, J. M. Carpenter, et al. The Two Micron All Sky Survey (2MASS). AJ, 131(2):1163–1183, Feb 2006.

[106] W. Voges, B. Aschenbach, Th. Boller, H. Bräuninger, U. Briel, and others Burkert. The ROSAT all-sky survey bright source catalogue. A&A, 349:389–405, Sep 1999.

[107] Donald G. York, J. Adelman, Jr. Anderson, John E., Scott F. Anderson, James Annis, Neta A. Bahcall, et al. The Sloan Digital Sky Survey: Technical Summary. AJ, 120(3):1579–1587, Sep 2000.

[108] Marie Martig, Frédéric Bournaud, Darren J. Croton, Avishai Dekel, and Romain Teyssier. A Diversity of Progenitors and Histories for Isolated Spiral Galaxies. ApJ, 756(1):26, Sep 2012.

[109] Christopher P. Ahn, Rachael Alexandroff, Carlos Allende Prieto, Scott F. Anderson, Timothy Anderton, Brett H. Andrews, et al. The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. ApJS, 203(2):21, Dec 2012.

[110] Kevork N. Abazajian, Jennifer K. Adelman-McCarthy, Marcel A. Agüeros, Sahar S. Allam, Carlos Allende Prieto, Deokkeun An, et al. The Seventh Data Release of the Sloan Digital Sky Survey. ApJS, 182(2):543–558, Jun 2009.

[111] Donald P. Schneider, Gordon T. Richards, Patrick B. Hall, Michael A. Strauss, Scott F. Anderson, Todd A. Boroson, et al. The Sloan Digital Sky Survey Quasar Catalog. V. Seventh Data Release. AJ, 139(6):2360, Jun 2010.

[112] Chris J. Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, et al. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. MNRAS, 389(3):1179–1189, Sep 2008.

[113] Chris Lintott, Kevin Schawinski, Steven Bamford, Anåe Slosar, Kate Land, Daniel Thomas, et al. Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies. MNRAS, 410(1):166–178, Jan 2011.

[114] E. L. Wright. A Cosmology Calculator for the World Wide Web. PASP, 118(850):1711–1715, Dec 2006.

[115] Xiaohu Yang, H. J. Mo, Frank C. van den Bosch, Anna Pasquali, Cheng Li, and Marco Barden. Galaxy Groups in the SDSS DR4. I. The Catalog and Basic Properties. ApJ, 671(1):153–170, Dec 2007.

[116] Michael R. Blanton, David J. Schlegel, Michael A. Strauss, J. Brinkmann, Douglas Finkbeiner, Masataka Fukugita, et al. New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys. AJ, 129(6):2562–2578, Jun 2005.

[117] Nikhil Padmanabhan, David J. Schlegel, Douglas P. Finkbeiner, J. C. Barentine, Michael R. Blanton, Howard J. Brewington, et al. An Improved Photometric Calibration of the Sloan Digital Sky Survey Imaging Data. ApJ, 674(2):1217–1233, Feb 2008.

[118] Xiaohu Yang, H. J. Mo, Frank C. van den Bosch, Youcai Zhang, and Jiaxin Han. Evolution of the Galaxy-Dark Matter Connection and the Assembly of Galaxies in Dark Matter Halos. ApJ, 752(1):41, Jun 2012.

[119] Xiaohu Yang, H. J. Mo, Frank C. van den Bosch, and Y. P. Jing. A halo-based galaxy group finder: calibration and application to the 2dFGRS. MNRAS, 356(4):1293–1307, Feb 2005.

[120] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. ApJ, 292:371–394, May 1985.

[121] Surhud More, Andrey V. Kravtsov, Neal Dalal, and Stefan Gottlöber. The Overdensity and Masses of the Friends-of-friends Halos and Universality of Halo Mass Function. ApJS, 195(1):4, Jul 2011.

[122] Alexander Knebe, Steffen R. Knollmann, Stuart I. Muldrew, Frazer R. Pearce, Miguel Angel Aragon-Calvo, Yago Ascasibar, et al. Haloes gone MAD: The Halo-Finder Comparison Project. MNRAS, 415(3):2293–2318, Aug 2011.

[123] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, et al. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. ApJS, 170(2):377–408, Jun 2007.

[124] V. Petrosian. Surface Brightness and Evolution of Galaxies. ApJ, 210:L53, Dec 1976.

[125] Franco D. Albareti, Carlos Allende Prieto, Andres Almeida, Friedrich Anders, Scott Anderson, Brett H. Andrews, et al. The 13th Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-IV Survey Mapping Nearby Galaxies at Apache Point Observatory. ApJS, 233(2):25, Dec 2017.

[126] Shadab Alam, Franco D. Albareti, Carlos Allende Prieto, F. Anders, Scott F. Anderson, Timothy Anderton, et al. The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. ApJS, 219(1):12, Jul 2015.

[127] Róbert Beck, László Dobos, Tamás Budavári, Alexander S. Szalay, and István Csabai. Photometric redshifts for the SDSS Data Release 12. MNRAS, 460(2):1371–1381, Aug 2016.

[128] Daniel J. Eisenstein, David H. Weinberg, Eric Agol, Hiroaki Aihara, Carlos Allende Prieto, Scott F. Anderson, et al. SDSS-III: Massive Spectroscopic Surveys of the Distant Universe, the Milky Way, and Extra-Solar Planetary Systems. AJ, 142(3):72, Sep 2011.

[129] Stephen A. Smee, James E. Gunn, Alan Uomoto, Natalie Roe, David Schlegel, Constance M. Rockosi, et al. The Multi-object, Fiber-fed Spectrographs for the Sloan Digital Sky Survey and the Baryon Oscillation Spectroscopic Survey. AJ, 146(2):32, Aug 2013.

[130] Matthew Colless, Gavin Dalton, Steve Maddox, Will Sutherland, Peder Norberg, Shaun Cole, et al. The 2dF Galaxy Redshift Survey: spectra and redshifts. MNRAS, 328(4):1039–1063, Dec 2001.

[131] Matthew Colless, Bruce A. Peterson, Carole Jackson, John A. Peacock, Shaun Cole, Peder Norberg, et al. The 2dF Galaxy Redshift Survey: Final Data Release. *arXiv e-prints*, pages astro–ph/0306581, Jun 2003.

[132] D. Heath Jones, Will Saunders, Matthew Colless, Mike A. Read, Quentin A. Parker, Fred G. Watson, et al. The 6dF Galaxy Survey: samples, observational techniques and the first data release. MNRAS, 355(3):747–763, Dec 2004.

[133] D. Heath Jones, Mike A. Read, Will Saunders, Matthew Colless, Tom Jarrett, Quentin A. Parker, et al. The 6dF Galaxy Survey: final redshift release (DR3) and southern large-scale structures. MNRAS, 399(2):683–698, Oct 2009.

[134] Marc Davis, Sandra M. Faber, Jeffrey Newman, Andrew C. Phillips, Richard S. Ellis, Charles C. Steidel, et al. *Science Objectives and Early Results of the DEEP2 Redshift Survey*, volume 4834 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 161–172. 2003.

[135] Jeffrey A. Newman, Michael C. Cooper, Marc Davis, S. M. Faber, Alison L. Coil, Puragra Guhathakurta, et al. The DEEP2 Galaxy Redshift Survey: Design, Observations, Data Reduction, and Redshifts. ApJS, 208(1):5, Sep 2013.

[136] S. P. Driver, D. T. Hill, L. S. Kelvin, A. S. G. Robotham, J. Liske, P. Norberg, et al. Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. MNRAS, 413(2):971–995, May 2011.

[137] I. K. Baldry, M. Alpaslan, A. E. Bauer, J. Bland-Hawthorn, S. Brough, M. E. Cluver, et al. Galaxy And Mass Assembly (GAMA): AUTOZ spectral redshift measurements, confidence and errors. MNRAS, 441(3):2440–2451, Jul 2014.

[138] Alison L. Coil, Michael R. Blanton, Scott M. Burles, Richard J. Cool, Daniel J. Eisenstein, John Moustakas, et al. The PRIsm MUlti-object Survey (PRIMUS). I. Survey Overview and Characteristics. ApJ, 741(1):8, Nov 2011.

[139] Richard J. Cool, John Moustakas, Michael R. Blanton, Scott M. Burles, Alison L. Coil, Daniel J. Eisenstein, et al. The PRIsm MUlti-object Survey (PRIMUS). II. Data Reduction and Redshift Fitting. ApJ, 767(2):118, Apr 2013.

[140] B. Garilli, L. Guzzo, M. Scodeggio, M. Bolzonella, U. Abbas, C. Adami, et al. The VIMOS Public Extragalactic Survey (VIPERS). First Data Release of 57 204 spectroscopic measurements. A&A, 562:A23, Feb 2014.

[141] L. Guzzo, M. Scodeggio, B. Garilli, B. R. Granett, A. Fritz, U. Abbas, et al. The VIMOS Public Extragalactic Redshift Survey (VIPERS). An unprecedented view of galaxies and large-scale structure at 0.5 < z < 1.2. A&A, 566:A108, Jun 2014.

[142] O. Le Fèvre, Y. Mellier, H. J. McCracken, S. Foucaud, S. Gwyn, M. Radovich, et al. The VIRMOS deep imaging survey. I. Overview, survey strategy, and CFH12K observations. A&A, 417:839–846, Apr 2004.

[143] B. Garilli, O. Le Fèvre, L. Guzzo, D. Maccagni, V. Le Brun, S. de la Torre, et al. The Vimos VLT deep survey. Global properties of 20,000 galaxies in the I_{AB} < 22.5 WIDE survey. A&A, 486(3):683–695, Aug 2008.

[144] Michael J. Drinkwater, Russell J. Jurek, Chris Blake, David Woods, Kevin A. Pimbblet, Karl Glazebrook, et al. The WiggleZ Dark Energy Survey: survey design and first data release. MNRAS, 401(3):1429–1452, Jan 2010.

[145] David Parkinson, Signe Riemer-Sørensen, Chris Blake, Gregory B. Poole, Tamara M. Davis, Sarah Brough, et al. The WiggleZ Dark Energy Survey: Final data release and cosmological results. Phys. Rev. D, 86(10):103518, Nov 2012.

[146] S. J. Lilly, O. Le Fèvre, A. Renzini, G. Zamorani, M. Scodeggio, T. Contini, et al. zCOSMOS: A Large VLT/VIMOS Redshift Survey Covering 0<z<3 in the COSMOS Field. ApJS, 172(1):70–85, Sep 2007.

[147] Simon J. Lilly, Vincent Le Brun, Christian Maier, Vincenzo Mainieri, Marco Mignoli, Marco Scodeggio, et al. The zCOSMOS 10k-Bright Spectroscopic Sample. ApJS, 184(2):218–229, Oct 2009.

[148] Isaac Newton. *Philosophiae naturalis principia mathematica, vol. 1 - 4.* 1760.

[149] Angela M. Berti, Alison L. Coil, Peter S. Behroozi, Daniel J. Eisenstein, Aaron D. Bray, Richard J. Cool, et al. PRIMUS: One- and Two-halo Galactic Conformity at 0.2 < z < 1. ApJ, 834(1):87, Jan 2017.

[150] Victor F. Calderon, Andreas A. Berlind, and Manodeep Sinha. Small- and large-scale galactic conformity in SDSS DR7. MNRAS, 480(2):2031–2045, Oct 2018.

[151] Yjan A. Gordon, Kevin A. Pimbblet, and Matt S. Owers. The Impact of Spectroscopic Fiber Collisions on the Observed Angular Correlation Function. *Research Notes of the American Astronomical Society*, 2(3):132, Jul 2018.

[152] I. K. Baldry, M. L. Balogh, R. G. Bower, K. Glazebrook, R. C. Nichol, S. P. Bamford, et al. Galaxy bimodality versus stellar mass and environment. MNRAS, 373(2):469–483, Dec 2006.

[153] A. Kolmogoroff. Confidence limits for an unknown distribution function. *The Annals of Mathematical Statistics*, 12(4):461–463, Dec 1941.

[154] N. V. Smirnov. Approximate laws of distribution of random variables from empirical data. *Uspehi Matem. Nauk*, 10:179–206, 1944.

[155] D. A. Darling. The kolmogorov-smirnov, cramer-von mises tests. *The Annals of Mathematical Statistics*, 28(4):823–838, 1957.

[156] J. A. Peacock. Two-dimensional goodness-of-fit testing in astronomy. *Monthly Notices of the Royal Astronomical Society*, 202:615–627, Feb 1983.

[157] Auguste Bravais. *Analyse mathématique sur les probabilités des erreurs de situation d'un point.* Impr. Royale, Paris, 1844.

[158] Karl Pearson. Mathematical Contributions to the Theory of Evolution. III. Regression, Heredity, and Panmixia. *Philosophical Transactions of the Royal Society of London Series A*, 187:253–318, Jan 1896.

[159] W. G. Cochran. The analysis of variance when experimental errors follow the poisson or binomial laws. *The Annals of Mathematical Statistics*, 11(3):335–347, 1940.

[160] J. Anthony Tyson. Large Synoptic Survey Telescope: Overview. In J. Anthony Tyson and Sidney Wolff, editors, Proc. SPIE, volume 4836 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 10–20, Dec 2002.

[161] Ž. Ivezić, S. M. Kahn, J. A. Tyson, B. Abel, E. Acosta, R. Allsman, et al. LSST: From Science Drivers to Reference Design and Anticipated Data Products. ApJ, 873:111, March 2019.

[162] G. Martin, S. Kaviraj, A. Hocking, S. C. Read, and J. E. Geach. Galaxy morphological classification in deep-wide surveys via unsupervised machine learning. MNRAS, 491(1):1408–1426, Jan 2020.