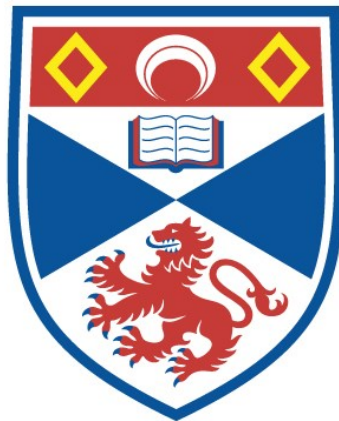# EVOLUTION OF SUPERMASSIVE BLACK HOLES: THE ROLE OF GALAXY MERGERS

Timothy Paul Hewlett

A Thesis Submitted for the Degree of PhD
at the
University of St Andrews

2019

Evolution of supermassive black holes:

the role of galaxy mergers

by

Timothy Paul Hewlett

Submitted for the degree of Doctor of Philosophy in Astrophysics

15$^{\text{th}}$ August 2018

University
of
St Andrews

Declaration

Candidate's declaration

I, Tim Hewlett, do hereby certify that this thesis, submitted for the degree of PhD, which is approximately 40,695 words in length, has been written by me, and that it is the record of work carried out by me, or principally by myself in collaboration with others as acknowledged, and that it has not been submitted in any previous application for any degree.

I was admitted as a research student at the University of St Andrews in September 2014.

I received funding from an organisation or institution and have acknowledged the funder(s) in the full text of my thesis.

Date   Signature of candidate

Supervisor's declaration

I hereby certify that the candidate has fulfilled the conditions of the Resolution and Regulations appropriate for the degree of PhD in the University of St Andrews and that the candidate is qualified to submit this thesis in application for that degree.

Date   Signature of supervisor

ii

Permission for publication

In submitting this thesis to the University of St Andrews we understand that we are giving permission for it to be made available for use in accordance with the regulations of the University Library for the time being in force, subject to any copyright vested in the work not being affected thereby. We also understand, unless exempt by an award of an embargo as requested below, that the title and the abstract will be published, and that a copy of the work may be made and supplied to any bona fide library or research worker, that this thesis will be electronically accessible for personal or research use and that the library has the right to migrate this thesis into new electronic forms as required to ensure continued access to the thesis.

I, Tim Hewlett, confirm that my thesis does not contain any third-party material that requires copyright clearance.

The following is an agreed request by candidate and supervisor regarding the publication of this thesis:

Printed copy

No embargo on print copy.

Electronic copy

No embargo on electronic copy.

Date   Signature of candidate

Date   Signature of supervisor

iii

iv

Underpinning Research Data or Digital Outputs

Candidate's declaration

I, Tim Hewlett, hereby certify that no requirements to deposit original research data or digital outputs apply to this thesis and that, where appropriate, secondary data used have been referenced in the full text of my thesis.

Date   Signature of candidate

Abstract

In this thesis the connections between galaxy mergers and the growth of supermassive black holes (SMBHs) are investigated. In chapter 2 we investigate the links between galaxy mergers and active galactic nuclei (AGN) as a function of luminosity, using visual classification of galaxy images and a new measure of morphological disturbance to identify mergers, testing the hypothesis that cataclysmic events such as mergers are required to drive the most rapid periods of accretion. We find no correlation between merger fraction and AGN luminosity, nor any difference in the total merger fractions of AGN and matched control galaxies. We also explore the possibility that the role of galaxy mergers evolves through cosmic time over the redshift range 0.5-2.2, finding a modest evolution, with higher redshift AGN more likely to be found in mergers than controls with 99% confidence. In chapter 3 we compare samples of simulated (Illustris) galaxies to real observations, exploring the efficacy of different techniques of merger-identification as a function of merger parameters (mass ratio, time elapsed since the merger, stellar mass, star formation rate etc.). Detailed analysis of the ability of structural parameters and visual classification to identify mergers allows for reinterpretation of observational studies in which AGN merger fractions are measured. We find the intrinsic merger fractions in relevant studies are likely several times higher than reported, with ∼50±20% of AGN in the local Universe triggered by galaxy mergers. In chapter 4 AGN are selected in the optical, infrared and radio to investigate the role of mergers in triggering AGN in different physical environments. We use structural parameters to estimate merger fractions. It is found that infrared-selected AGN are more likely to be associated with mergers than their controls, optically selected AGN are equivalent to controls and radio-selected AGN show fewer signs of interactions than their controls. Collectively, this is interpreted as evidence that a substantial fraction of SMBH growth is driven by galaxy mergers, but care needs to be taken to control for various observational biases, which can hide or dilute real underlying causal relations.

Contents

1

Introduction

1.1 The Observed Universe

Astronomy shapes how humans understand our place in the Universe and has long influenced the way we humans think about ourselves. When Galileo observed four moons of Jupiter in orbit around the giant planet in 1610 he contradicted the religious orthodoxy of the time: that the Universe was static, eternal and Geocentric. The works of Galileo, Newton and others served to profoundly alter how people thought about the Universe - as mechanistic and changing - with deep implications for technology, politics and society. Similarly, when Einstein wrote down his theories of special and general relativity (Einstein, 1905, 1915) our collective notions of existence were again overturned. Not only are objects in space dynamical, but spacetime itself - the newly integrated space and time - takes on a dynamical form; there are no special positions or times, only the relations between things. Humanity found itself once again metaphysically belittled; "nebulae" were established by Hubble in 1926 (Hubble, 1926) to be galaxies apart: "island universes", with recessional velocities correlated with distance (the Hubble relation). We now know that not only is the Solar System not unique, but is one of around 100 billion such systems in the Milky Way and the Milky Way is just one galaxy of hundreds of billions in the observable Universe. The Copernican principle, in which no position or time is special, is so firmly established that people are even left wondering (with some justification from inflationary and quantum theories) if our Universe may be simply one

universe within a perhaps infinite set of universes (the multiverse). After all, what makes this particular Universe, with its particular configuration of laws and particles, so special? Below lies a description of how astronomical data relate to observed astrophysical systems, and a brief description of the observations on which the standard model of cosmology is built. It is within the framework of this standard model that we understand the formation and evolution of galaxies and, in this thesis, probe the processes which lead to the growth of the Supermassive Black Holes (SMBHs) embedded within galaxies.

Until very recently (with the advent of gravitational wave astronomy) almost all astronomy has relied on the detection of light from astrophysical sources to interpret their physical conditions. Consequently, astronomers have become remarkably adept at teasing out information from any detected photons. The first and most simple information-containing quantity to be measured is the number of photons in some range of wavelengths. Such photometric data reveals the brightness of an object which, combined with information about the distance to the object, reveals the intrinsic luminosity; conversely, if the luminosity is known the distance can be deduced. In extragalactic astronomy photometry alone tends to have limited uses; spectroscopy, however, can reveal exquisite details about sources across the universe.

Spectroscopy relies on an understanding of atomic physics. When electrons transition between different quantised levels they emit or absorb photons of particular energies. A radiant gas will emit light with particular wavelengths corresponding to the atomic structure of the gas, which can be measured in a laboratory. The emission and absorption lines found in the spectra of astrophysical sources can be related to the known elements to deduce the chemical composition of the sources. Light can be doppler shifted and so information about the physical environment is encoded in the shape of emission/absorption lines: hot objects will produce broad lines, orbiting or outflowing gas can produce skewed or asymmetric lines, globally shifted lines indicate a bulk motion (or cosmological redshift). This is an extremely useful and important diagnostic, allowing measurement of the distance to galaxies (in some cosmological model); the dynamics of stars and gas in galaxies; black hole masses and spins; the mass, temperature and metallicity of stars etc.

Study of the extragalactic Universe has proceeded from the tentative observations of 10s of galaxies in the early 20th century to many millions mapped out across the Universe today (e.g. Hubble, 1926; Blanton et al., 2017). Using certain "standard candles", whose luminosity are known, Hubble measured the distance to other galaxies using the relation between apparent brightness, intrinsic luminosity and distance. Hubble found the universe to be expanding according to the Hubble law (Hubble, 1929), which can be written in the non-relativistic limit:

$$v \sim cz = H_0 D \tag{1.1}$$

where v is the recessional velocity, c is the speed of light, z is the redshift and H_0 is the Hubble constant. Hubble's law is understood to arise due to the expansion of the universe. Every point of space expands uniformly; over small distances the expansion is tiny, but integrating over large regions of space, such as the distances between galaxies, yields a significant expansion and causes galaxies to appear to be rushing apart. If galaxies are rushing apart today, in the past they must have been much closer together: this observation forms the bedrock of the big bang theory. Light travelling between two distant galaxies expands with the space, elongating the wavelength and lowering photon energies to redder colours ($E_\gamma = hc/\lambda$). The resultant cosmological redshift allows the radial distance to galaxies to be mapped.

Galaxy surveys tend to observe either large sections of sky to shallow depths, or focus on small areas to great depths. Wide-field surveys, such as the Sloan Digital Sky Survey (SDSS), can observe millions of galaxies in the local universe, building large statistical samples (Blanton et al., 2017). The smallest, dimmest structures will be missed, but rare luminous and distant sources in the field of view can be discovered. Deep imaging of small regions, where a telescope may be directed at a point of sky for weeks of integrated time, can discover galaxies too faint for the wide-field surveys to detect and small numbers of extreme objects in the early universe. With the latter technique lower bounds on the number of galaxies in the universe can be easily estimated. Around 10,000 galaxies are observed in the Hubble Ultra-Deep Field (Beckwith et al., 2006); this field covers only ~1/20th of a degree on each side and implies the existence of more than 100 billion currently detectable galaxies in the observable universe (note that the real number of galaxies is likely much greater than this, due to observational constraints).

Using (predominantly) wide-field surveys populations of galaxies in the local universe have been analysed and categorised; deep surveys have revealed elements of the early universe; and radio astronomy has elucidated the structure of the universe just ~400,000 years after the big bang (Planck-Collaboration, 2016). Some key observations are summarised here, before being placed in the context of the standard model of cosmology in sections 1.2 and 1.2.2.

The lives of galaxies are complex and interwoven. The details of how stars form is both broadly instructive and important to future discussion, and is described more quantitatively in section 1.2.2. For a broad discussion of extragalactic observations it is sufficient to say that for stars to form requires the existence of dense clouds of cool gas. Massive stars burn far hotter, therefore bluer and more quickly than lower mass stars, emitting thick, fast-flowing and metal-enriched winds (De Loore et al., 1977). The consequence of this short lifetime is that where there are massive stars, there is ongoing (or very recently ceased) star formation. The luminosity of a star, $L \propto \sim M^\alpha$ (where α ~2-4 depending on the mass of the star; Kuiper, 1938) so the light from a galaxy can be dominated by young massive stars for even modest star formation rates (SFRs). Massive stars formed in giant molecular clouds will heat and blow

away the surrounding gas, exposing the new star cluster. After just a few million years, the most massive stars will exhaust their fuel and explode in supernovae, injecting huge amounts of energy into the surrounding interstellar medium (ISM). Depending primarily on their mass, massive stars will collapse to form either neutron stars or black holes (BHs). In each galaxy sits a supermassive BH (SMBH; e.g. Kormendy & Richstone, 1995), the formation of which is poorly understood. Astonishing amounts of energy are released by gas accreting onto SMBHs and that energy may drive winds fast enough to entirely escape the potential wells of galaxies. The result is a complex ecosystem where gas falls into galaxies, is cycled through star formation, enriched with metals, and re-injected into the ISM and wider galactic environment in winds and supernovae ejecta (e.g. Vogelsberger et al., 2014a). The detailed physics underpinning this balance between inflows and outflows is extremely important to how galaxies and SMBHs form and evolve, with strong dependencies on mass, environment and redshift (Vogelsberger et al., 2014a). This will be discussed in more detail in section 1.3.3.

Galaxies are found to have quite a bimodal distribution in mass-colour-morphology space in the local universe: low-intermediate mass galaxies tend to be blue, star forming and disky; high mass galaxies tend to be elliptical or bulge-dominated structures with little star formation and a correspondingly red colour, though there is significant diversity among populations (Baldry et al., 2004, 2006). These ellipticals are often called "quenched", referring to the cessation of star formation within them. The small fraction of galaxies that do not fall into these classes are named "irregulars", with asymmetric shapes often due to rapid and patchy star formation or the merger of galaxies.

Looking to greater distances, earlier in the history of the universe, the state of galaxies shifts. The fraction of irregular galaxies dominates by z∼3, as does the SFR density (Buitrago et al., 2013; Bouwens et al., 2014). Remarkably, there are already giant quenched elliptical galaxies and SMBHs in place, with masses comparable to those observed today, 10 billion years later (e.g. Alexander et al., 2008). In general, however, galaxies have lower average masses (though there is a bias to detect the more massive, brighter galaxies in surveys) and are more compact (Trujillo et al., 2007). The highest redshift observations of today probe galaxies out to z∼10, finding the rarest, highest mass objects just ∼400 million years after the big bang (Oesch et al., 2016). These observations inform the picture of an evolving SFR density through time, climbing to z∼2-3 and then falling again to the present time (Bouwens et al., 2014).

Underpinning the formation and growth of galaxies is dark matter, comprising ∼80% of the matter in the Universe and apparently only interacting gravitationally. Through comparison of the mass inferred by the motions of galaxies to the mass inferred from luminous components (i.e. stars) it was discovered by Fritz Zwicky in the 1930s (Zwicky, 1933) that galaxy clusters appear to contain far more mass than can be accounted for in stars. Perhaps the simplest expla-

nation for this disparity is to suppose there exists some form of "dark matter" which interacts gravitationally but not electromagnetically, hence is non-luminous. Neutrinos are one example of such a particle, though neutrinos do not constitute the majority of dark matter. Later evidence for the existence of dark matter came from Vera Rubin, who studied galaxy rotation curves and found that the orbital angular momenta of stars was far too high for those stars to remain gravitationally bound by the luminous matter alone (Rubin & Ford, 1970). Comprising ∼80% of the matter in the Universe, dark matter underpins the formation and growth of galaxies. There has never been a direct detection of dark matter (e.g. from energy deposits when dark matter particles scatter off atomic nuclei in a detector), but its existence is inferred from several lines of reasoning, including the dynamics of visible matter (Zwicky, 1933; Rubin & Ford, 1970), structure formation theory, gravitational lensing experiments (Clowe et al., 2004) and measurements of the Cosmic Microwave Background Radiation (CMB, see below Planck-Collaboration, 2016). These indirect measurements place constraints on the possible nature of dark matter: for instance, if dark matter were dynamically hot, massive structures could not form in the early universe since dark matter particles would have high enough kinetic energies to escape protogalactic potential wells (ruling out neutrinos). Galaxy formation and evolution is now understood within the framework where dark matter is dynamically cold, accounts for ∼80% of the matter density of the universe, and sows the seeds for galaxy formation by gravitational collapse.

The other dark component of the universe is perhaps even more mysterious: dark energy accounts for ∼70% of the energy density of the universe and drives the accelerated expansion of the Universe at late times ($z \lesssim 0.4$). That the expansion of the Universe was accelerating was discovered in 1998 (Riess et al., 1998) from observations of type 1a supernovae, extremely luminous standard candles. Since their luminosity can be directly inferred the distance to them is (relatively) simply calculated and the rate of cosmic expansion through time can be accurately determined. This result has since been confirmed from analysis of the CMB and large scale structure.

Galaxies, the dark matter halos they are embedded within and the gas strewn between them (intergalactic medium, IGM) form vast networks. Strands of galaxies called filaments connect immense galaxy clusters and groups (e.g. Bharadwaj et al., 2004), where several filaments have knotted together to form structures with masses of $\sim 10^{13} - 10^{15} M_{\odot}$, and circumnavigate great empty voids with radii (in the local universe) ∼150Mpc. These structures together make up the Cosmic Web. Though both elliptical- and disk-galaxies are found in all environments there exists an environmental bias: elliptical, quenched galaxies tend to live in high-density regions of the universe, in rich clusters of galaxies; spiral galaxies tend to be found in smaller groups, filaments and the field. Although galaxies, stars and planets form pockets of complexity, on large scales the universe looks remarkably homogenous, with a sta-

tistically equivalent distribution of galaxies at all positions; on large scales a self-similar fractal structure emerges.

As a clue to the origin of the structural homogeneity of the Universe, Penzias and Wilson serendipitously discovered a uniform radio-frequency radiation field across the whole sky in the 1960s (Wilson & Penzias, 1967). This observation was a key prediction of the Big Bang model of the beginning of the Universe, in which it was conjectured that a hot Big Bang would produce huge quantities of radiation which would have been redshifted by cosmological expansion since. Observations of the CMB have increased dramatically in precision and sensitivity since its initial discovery. It is now known that the CMB has an almost perfectly isotropic temperature of T~2.7K across the whole sky, with temperature variations of just 1 part in 100,000 (Planck-Collaboration, 2016). Analysis of the CMB - its spectrum, polarisation and anisotropies - has yielded some of the most stringent constraints on theories of the formation and evolution of the Universe.

Now that some key observations of the local and distant Universe have been described, in section 1.2 the standard model of cosmology is established, in which structure forms as detailed in section 1.2.2. The physics of black holes and AGN are established in section 1.3, before relations between galaxies and AGN are outlined in section 1.4.

1.2 The Standard Model of Cosmology

In this section the standard model of cosmology is described, from the hot big bang through to speculation about the future of the Universe. The conclusion that the Universe began in a hot dense state follows simply and naturally from the observation that galaxies are receding from each other. Reverse the trajectories of galaxies and their paths converge at some point in the past. A simple estimate of how long ago the Universe began comes from the rate of recession observed today, measured by the Hubble constant: $t_0 = 1/H_0 \sim 70 \text{kms}^{-1}\text{Mpc}^{-1} \sim 4\times10^{17}\text{s}$ ~14 billion years ago. If the Universe began in a singularity how could it then expand, what drove it to do so, and how did the structures present today emerge? There are many physical theories which seek to explain the origins of the Universe: Inflation (Guth, 1981), Conformal Cyclic Cosmology (Penrose, 2012), String Cosmology (e.g. Horowitz & Marolf, 1998), Loop Quantum Cosmology (e.g. Bojowald, 2002), bounce and black hole cosmologies (e.g. Smolin, 1991; Garriga et al., 2016). The most developed and successful of these is Inflation, forming a bedrock for modern cosmological models. Although the exact details of the first instants after the big bang are not of crucial importance to this thesis, the basic ideas of inflation are presented below. First, the essentials of 'standard' cosmology are covered in section 1.2.1, motivating the development of Inflation.

1.2.1 Classical Cosmology

Relativity forms the foundation of modern cosmology and so the fundamentals are briefly established here before describing the cosmological model. The Newtonian view is of a Universe which is static and infinite; the gravitational force between two objects is dependent only on their masses and separation, the force itself operates instantly. Einstein showed how this was false with the development of special and general relativity. Fundamental to the development of special relativity is the fact that the speed of light is the same in all inertial frames. It follows that moving clocks must tick slower than stationary clocks, as demonstrated by the "Einstein clock": a beam of light reflecting between two vertically displaced mirrors constitutes the time-keeping mechanism. By moving the whole clock horizontally it is clear the path-length of the light has elongated and, since the speed of light is constant, the duration between ticks increases correspondingly. Similarly, the length of a moving object appears contracted to some "stationary" observer. The essential insight extending special relativity to include gravity is the equivalence principle: gravitation is locally indistinguishable from acceleration in flat space. The principle of equivalence serves to simply demonstrate how spacetime is curved: any body in a gravitational field will experience the same acceleration, regardless of mass. Light must therefore bend in a gravitational field; however, since light always takes the shortest distance between two points this implies the spacetime must be curved. Gravity, in this view, is as a fictitious force (such as the centrifugal force) simply caused by particles following the geodesics of curved spacetime. The core ideas of general relativity can be summarised poetically in the words of John Wheeler: "Matter tells spacetime how to curve, spacetime tells matter how to move".

The Friedmann-Robertson-Walker metric (in spherical polar coordinates) defines the causal structure of spacetime:

$$ds^2 = c^2 dt^2 - a(t)^2 \left(\frac{dx^2}{1 - kx^2} + x^2 [d\theta^2 + sin^2\theta d\phi^2] \right) \tag{1.2}$$

where x is the the comoving angular distance, k is the curvature term normalised to take on values of 0, 1 and -1 for flat, closed and open Universes respectively, and a(t) is the scale factor (size) of the Universe normalised to a_0=1 at present day. Our own Universe is measured to be flat to within <1% (De Bernardis et al., 2000). The fate of a Universe in classical cosmology is determined by the balance of curvature, pressure and gravity; pressure exerts a force that drives expansion while gravity slows expansion down. The Friedmann equation expresses the cosmic evolution of each component:

$$\frac{\dot{a}(t)^2}{a(t)^2} = H(t)^2 = \frac{8\pi G}{3}\left(\frac{\rho_r}{a(t)^4} + \frac{\rho_m}{a(t)^3}\right) - \frac{kc^2}{a(t)^2} + \frac{\Lambda c^2}{3} \qquad (1.3)$$

where ρ_m is the density of matter today, ρ_r is the radiation density today (including neutrinos), Λ is the cosmological constant, H(t) is the Hubble parameter and a dot denotes a derivative with respect to time. Einstein originally introduced Λ as a means of allowing the Universe to be static (since the influence of matter, radiation and curvature could be balanced by some extra factor), without an a priori theoretical justification for its existence but with the Gauge freedom to do so (the divergence of the energy-momentum tensor is invariant under a uniform translation). It is now thought by some that the cosmological constant does exist and gives rise to the observed dark energy (Riess et al., 1998), but the basic theoretical prediction of its value (as due to pressure from virtual particles) overestimates the vacuum pressure by 120 orders of magnitude (Carroll, 2001). Presumably, the vacuum pressure terms entirely or almost entirely cancel, and/or dark energy has some other theoretical origin (e.g. entropic acceleration; Easson et al., 2011). The Friedmann equation can facilitate many kinds of Universe (collapsing, expanding, static) and has been known for nearly a century; by measuring the values of the different components (ρ, k, Λ, H) predictions for our own Universe can be made from the model.

Radiation, matter and the cosmological constant each respond differently under cosmic expansion. Both the number density of photons and the energy of each photon decreases as redshifted with the expansion of the Universe, so Ω_r has the strongest dependence on a(t). In the early Universe, as a(t) is small, expansion is driven by radiation pressure; $\dot{a}$ is large and the Universe expands rapidly, decelerating as radiation pressure declines with expansion. Since the mass of particles remains constant (unlike photon energies) while the number density of matter particles decreases, the pressure due to matter changes less rapidly as the Universe expands than the pressure from radiation. Gravity eventually begins to dominate, slowing expansion. An open Universe (k=-1) drives accelerated expansion of the Universe while a closed Universe (k=+1) does the opposite. The $c^2/a(t)^2$ in the curvature term ensures that if the Universe deviates much from flatness at early times it will either expand too rapidly for structures to form, or rapidly collapse back in on itself if open or closed respectively. If the Universe is approximately flat today, standard cosmology suggests that there must have been remarkably fine-tuned values of the density and Hubble parameter in the early Universe so as to avoid near-instant collapse or such rapid expansion that structures could not form. Dark energy takes the form of a cosmological constant in the Friedmann equation, thus its contribution is constant in space. Since the Universe is expanding the amount of dark energy grows with $a(t)^2$ and so dominates the expansion of the Universe at late times. In a Universe dominated by a cosmological constant, with negligible matter, radiation and curvature, integrating the Friedmann equation yields an exponentially expanding Universe. Such a state would dilute

the energy of the Universe to the point where structures disintegrate, entropy is maximised and novel processes become impossible, termed the heat death of the Universe.

The big bang model makes a unique prediction: such a Universe would be exceedingly hot in the distant past; as it cools quarks can condense to form baryons and some of those baryons later condense into Deuterium, Helium, and Lithium nuclei (no larger, due to the rapidity of Beryllium decay) while most remains as Hydrogen. The Universe remains too hot for electrons and protons to combine to form atoms until $z\sim1100$. At this point the Universe becomes rapidly full of a neutral gas, becoming suddenly transparent to photons (which are no longer scattering off free electrons). These photons have energy corresponding to the temperature of the Universe at that time and are then free to travel relatively unperturbed to detectors on Earth today, providing a snapshot of the Universe at just $\sim$400,000 years old. As they travel, the photons from this surface of last scattering are redshifted by cosmic expansion, arriving as microwaves (hence the cosmic microwave background). The CMB has been measured to have a near-perfect blackbody spectrum with a peak wavelength at $T\sim2.7K$ (Planck-Collaboration, 2016): this is exceptional evidence that the whole observable Universe was in thermal equilibrium at one particular time. The CMB has almost uniform emission in all directions, but with tiny temperature deviations corresponding to 1 part in 100,000 (Planck-Collaboration, 2016). Over-dense regions are warmer than sparser regions; these slight perturbations to uniformity can grow under gravitational forces to form massive structures in the Universe at later times.

It is remarkable that the fate of the Universe is entirely described by a few simple equations in classical cosmology, but a few outstanding questions remain. The Universe appears homogeneous on large scales, but regions separated by some angle ($\gtrsim2°$) were apparently never causally connected. How are these regions in thermal equilibrium without ever having been in contact? The curvature of the universe is measured to be so close to unity, with the energy density of the Universe extremely close to the critical value (above which the Universe would re-collapse). Some overall curvature, any small deviation from this balance, would cause the Universe to almost instantly re-collapse (for positive curvature) or expand too rapidly for structures to form (negative curvature); why should the Universe choose such precise initial conditions so as to give rise to a Universe capable of hosting conscious observers? Inflation addresses these issues by positing that the Universe went through a period of exponential growth right at the outset of its formation. Causally disconnected regions today were once in thermal equilibrium, and the density of the Universe is rapidly varied during inflation, falling off to close to the critical value as the Universe expands. The final advantage of inflation comes in its ability to explain the origins of structure formation and thus the power spectrum of the CMB. During inflation, quantum fluctuations occur throughout the spacetime. Long-wave fluctuations of scalar fields and spacetime curvature are blown up to large scales during inflation and provide density perturbations which can grow into galaxies and larger structures at late

times.

1.2.2 Growth of Structure

To first order, cosmological parameters set the evolution of the Universe: the strength of gravity (set by Newton's constant, G) relative to the other fundamental forces, the curvature of the Universe, the vacuum pressure of space etc. Small changes to any of these could render the formation of structure impossible. Those structures make up a beautiful, interconnected, multi-dependent ecosystem which alters the future of the cosmos in turn. Gas collapses to form stars; stars enrich gas with heavy elements; gas is distributed through the Universe by feedback from stars and black holes; new stars with different properties form from the chemically-enriched gas; planets and solar systems form. As the Universe evolves, new phenomena become possible which were previously unimaginable. In this section the processes which lead to formation of the first stars, black holes and galaxies are outlined, along with the evolution of those components. The interdependence between different components is elaborated upon in section 1.4.

Post-inflation (assuming inflation is the correct theory of the early Universe), the quantum fluctuations stretched into large-scale density fluctuations can begin to grow under gravitational collapse into larger structures. During the radiation dominated era ($z \lesssim 3000$) particles and photons behave as a single fluid (plasma); pressure from photons and neutrinos inhibit the formation of structure and density perturbations can only grow logarithmically. Collapse into some over-density increases the opposing force of radiation pressure, causing matter in that same region to re-expand (e.g. Eisenstein et al., 2005). Oscillations of matter on large scales are thus established in this period of the history of the Universe with marked impact on the formation of galaxies and galaxy clusters. Outflowing gas propagates as a sound wave, with an over-dense spherical shell of size corresponding to the sound speed of the plasma multiplied by the time spent in the radiation-dominated era. These over-dense shells propagate from different positions (at the peaks of primordial fluctuations) and overlap like waves from raindrops on a pond. This is measured today in the large-scale structure of the Universe in baryon acoustic oscillations (Eisenstein et al., 2005). Once the Universe has expanded sufficiently electrons cool and are captured by protons, the Universe becomes neutral, photons decouple from matter and escape as the CMB (Planck-Collaboration, 2016). Without this source of radiation pressure, matter collapses more rapidly into denser regions. As denser regions grow, they acquire a stronger gravitational pull and accrete gas more rapidly, thus becoming more massive and growing more rapidly etc. Resultantly, over-densities grow non-linearly following the release of the CMB.

This much can be deduced from thermal physics and the cosmological model, and is sup-

ported by observations of the CMB and large-scale structure today. The following era, between redshifts of z~1000 - 10, makes up the "dark ages" of astronomy, before the first observations of galaxies. Astronomers have to interpolate from the slight inhomogeneities recorded in the CMB to the relatively massive galaxies and black holes observed at z$\lesssim$10 to deduce how structure may have developed in this time. The standard model for structure growth is a hierarchical assembly, where small objects merge to form larger structures while gas is accreted from the surrounding field. From a near-homogeneous origin, structures grow non-linearly under gravity to form the cosmic web observed in the Universe today (e.g. Frenk & White, 2012). Near-future radio surveys will be able to map neutral hydrogen (using the HI 21cm hyperfine emission line) out to z $\sim$ 30 (Furlanetto & Briggs, 2004) and next generation telescopes (JWST, LSST) will, with the aid of gravitational lensing, probe galaxies to similar distances. For now, the first stars and galaxies in the Universe remain tantalisingly out of reach.

To gain some insight into how the first stars in the Universe formed, and how they may be unique, it is instructive to first consider how stars form in the local Universe. Molecular gas comes together to form immense clouds, as gas contracts under gravity it heats up (gravitational potential converted to kinetic energy) until it comes to hydrostatic equilibrium (Jeans, 1902). External shocks may instigate further collapse, for example from distant supernovae. To continue to collapse the gas must radiate away its thermal energy, but as it becomes denser the cloud becomes opaque to its own radiation. A highly efficient means of removing further heat is by line cooling: metals in the cloud have many possible electron transitions, including emission lines in the infrared to which the collapsing cloud is transparent. If the gas pressure is unable to balance the force of gravity the cloud is Jeans unstable and will continue to collapse. This happens if the mass is particularly high for a given temperature, or the gas is particularly cool for a given mass (Jeans, 1902). Stars in the local Universe tend to form from cool giant molecular clouds, since cooling is efficient due to the relatively high density of metals (relative to the early Universe). Clouds fragment to produce thousands of stars with a broad distribution of masses, empirically determined to scale with the stellar mass like log(N) $\sim M^{-2.3}$ (Salpeter, 1955; Kroupa, 2001): very few high mass stars form. The early Universe tells a different story entirely. Gas is pristine (H, He and Li only) and so cooling is extremely inefficient, with most cooling coming from molecular Hydrogen. Though far from certain, some models suggest that the first stars may consequentially be extremely massive (as required by the high temperatures) (e.g. Omukai & Nishi, 1998). Such stars, known as Pop III stars, would be extremely luminous, ionise bubbles of Hydrogen around them, and would drive winds loaded with the heavy elements forged in their interiors. The enriched gas would mix with the surrounding pristine gas and seed the Universe with the heavy elements needed to form the long-lived low mass stars of today. Many pockets of these first stars may form from the gas as it flows in to form protogalaxies (Tegmark et al., 1997).

Since the late stages of stellar evolution are important for gas dynamics in galaxies and black hole formation, some core features are outlined here. In low mass main sequence stars (core Hydrogen burning) the internal gas pressure remains approximately constant through the lifetime of the main sequence phase (as does the gravitational force) (Laughlin et al., 1997); when the core Hydrogen is depleted fusion of other elements (e.g. Helium) can begin to dominate, proceeding in shells surrounding the core (T. & D., 2001). This severely alters the balance of pressure and gravity in the star, leading to the red giant phase. The star expands to hundreds of times its original size; outer layers are now weakly bound and can easily escape in the form of winds; the core contracts to leave a dense, hot metallic remnant: a White Dwarf with $M_{WD} \sim 0.5 - 1.4 M_\odot$ and R $\sim R_\oplus$ (Chandrasekhar, 1931). The White Dwarf is supported against collapse by electron degeneracy pressure. The Pauli exclusion principle states that two identical fermions cannot exist in the same state at the same time, and so further contraction of the stellar remnant forces electrons into successively higher energy levels. Or, to consider the issue more heuristically, as the gravitational force increases the electrons accelerate, increasing the pressure. At some threshold gravitational force the electrons would have to exceed the speed of light in order to support the remnant against collapse. Beyond this threshold (the Chandrasekhar limit) the electrons combine with protons to form neutrons, this source of pressure is removed and the White Dwarf collapses into a Neutron Star, similarly supported by neutron degeneracy pressure ($M_{NS} > 1.4 M_\odot$ and R $\sim$10km). Massive main sequence stars (M $\gtrsim 8 M_\odot$) have a broadly similar evolution to lower mass stars (though on a much shorter timescale), but in their final stages the core becomes massive enough and without a sufficient opposing pressure (since fusion beyond Iron costs energy) that it collapses directly into a Neutron Star. The process of rapidly converting $\sim 1 M_\odot$ of matter into neutrons releases a huge flux of neutrinos which impact and blast away the outer layers of the star in a supernova explosion, ejecting relativistic material into the cosmos. Even more massive stars (M $\gtrsim 25 M_\odot$) may directly form BHs in their collapse, while ejected material may fail to escape the potential and flow back onto the compact object (Zhang et al., 2008).

Approximately contemporary to the first stars is the formation of massive black holes (MBHs), the seeds which will grow into SMBHs. Though the question of how the first MBHs formed is still open, quite stringent constraints come from observations of quasars at z>6 (Mortlock et al., 2011). These BHs have masses $M_{BH} \gtrsim 10^8 M_\odot$; accreting at the Eddington rate (the approximate maximum rate, section 1.3) these BHs would most likely need seed masses $\gtrsim 10^4 M_\odot$ to reach such masses within a billion years of the Big Bang (Volonteri, 2010). No such BHs are known to form in the local Universe. Intermediate mass black holes are expected to exist in low-mass galaxies or globular clusters, but as yet there is only very tentative evidence of their existence (e.g. Maccarone et al., 2007). Three major hypotheses for the formation of MBH seeds are: that they form from the merging remnants of Pop III stars (Tegmark

et al., 1997); from direct collapse of massive gas clouds (Loeb & Rasio, 1994); or from primordial BHs formed during inflation (Khlopov et al., 2005). No such processes could occur in the Universe today.

If Pop III stars are massive enough or form in dense enough clusters they may leave massive remnant black holes ($M_{BH} \sim 100 M_\odot$) in a gaseous environment (or they may expel the gas to remove angular momentum from the BHs). Through many-body interactions and friction they could come close enough to emit gravitational waves and merge (Tegmark et al., 1997). A merger of $\sim$100-1000 of these would be sufficient to form a seed BH massive enough to grow to the size of those observed at z$\sim$7. Alternatively, huge primordial clouds of gas could collapse to form MBH seeds. In such models, gas pours along filaments into protogalactic cores, but star formation is suppressed due to the supersonic gas motions left over from the Big Bang except in the deepest wells. Once a cloud of $\sim 10^5 M_\odot$ congregates within a dark matter halo it may undergo catastrophic direct collapse into a BH (Hirano et al., 2017). Finally, some density fluctuations during inflation may be sufficiently large for gravity to overcome pressure, triggering whole regions to collapse and form primordial BHs (Khlopov et al., 2005; Volonteri, 2010). Depending on the stage of inflation in which these black holes form they can take on a wide range of masses, from $\sim 10^{-5}$g to $10^5 M_\odot$ (Garriga et al., 2016), since later in inflation density perturbations have stretched to larger sizes, hence enclose larger masses. Though much of this region is excluded by gravitational lensing surveys and searches for Hawking radiation, unexplored regions of parameter space could yet yield exciting discoveries and, in principle, primordial BHs could make up the seeds for SMBHs and even constitute some or all of the dark matter (Volonteri, 2010; Garriga et al., 2016).

Surveys of the low redshift Universe contain many millions of galaxies and map out cosmic structure with exquisite detail. Observations are consistent with galaxies congregating over time since z$\sim$10 (and from density perturbations implied by CMB anisotropies at z$\sim$1000; Vogelsberger et al., 2014a), as expected, growing larger clusters and superclusters right up to the present day through gravitational collapse. As gas flows into galaxies more is available for star formation, but gas also becomes locked up in stars (primarily long-lived, low mass stars with weak winds), contributing to the shape of the Madau plot, shown in figure 1.1. Inferences from simple dark matter simulations, assuming baryonic matter follows the dark matter, suggest the space density of both high and low stellar mass galaxies should be higher than is observed (Silk & Mamon, 2012). Hydrodynamical cosmological simulations (and basic theoretical arguments) suggest the growth of galaxies is inhibited by feedback from supernovae and AGN (e.g. Vogelsberger et al., 2014a), depending on the galaxy mass, also contributing to the downturn at late times in figure 1.1 (as will be discussed in more detail in section 1.4). At late times dark energy slows the assembly of structure (relative to if it were absent) but, as yet, does not disrupt structures already formed. In the following section we describe the
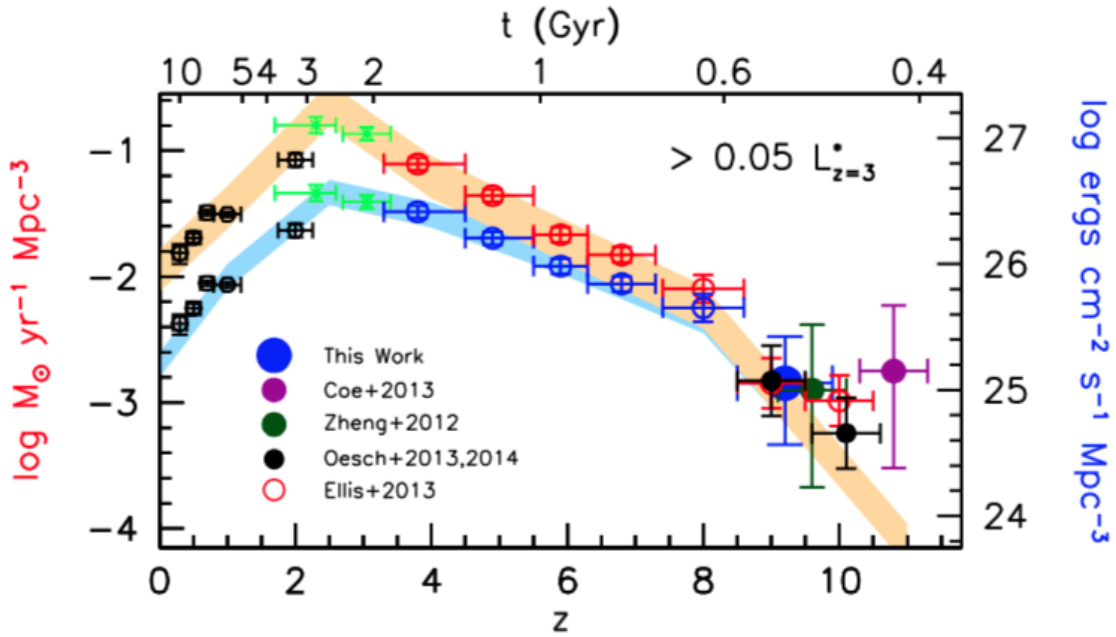
Figure 1.1: Madau plot, taken from (Bouwens et al., 2014)

properties and physics of AGN, before discussing their possible role in galaxy evolution and the role of galaxies in in the evolution of AGN in section 1.4.

1.3 Active Galactic Nuclei

AGN, first discovered in 1909 (Faith, 1909), with spectra partially characterised by Seyfert in 1943 (Seyfert galaxies Seyfert, 1943) and luminosities measured in 1963 (Schmidt, 1963), are amongst the most luminous phenomena in the Universe. The source of their extreme luminosity is a hot accretion disk, from which gas accretes onto a central SMBH. AGN emit powerfully across the spectrum, from radio jets to UV/optical disk emission to X-rays. They can influence gas throughout the galaxy and beyond, with ionising cones of radiation (Pogge, 1988; Tadhunter, 1989) and some with megaparsec-scale radio jets (e.g. Nesvadba & Lehnert, 2008), possibly making AGN vital cosmological ingredients. In this section some details about BHs are outlined (1.3.1) before accretion onto BHs and detection of AGN (1.3.2) and the resultant feedback (1.3.3) are discussed.

1.3.1 Black Holes

The intrigue of black holes lies in their unknowability; their secrets shielded by the event horizon, their very existence paradoxical. The question of how SMBHs come to grow is interesting from a cosmological perspective - the implications for structure, stars and galaxies - but also from a fundamental perspective. What happens to accreted matter? What are the properties of black holes, how do their properties change through time and how do BHs influence

the Universe? To fully motivate a discussion of the feeding mechanisms of black holes their fundamental properties and some of their more interesting features are first outlined here.

The Schwarzschild solution to the Einstein field equations describes the spacetime metric external to a static, spherically symmetric mass, M. The presence of mass causes deviation from the flat spacetime metric and is shown below in spherical coordinates with G = c = 1:

$$ds^2 = \left(1 - \frac{2GM}{c^2 r}\right)c^2 dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 d\theta^2 - r^2 sin^2\theta d\phi^2 \qquad (1.4)$$

which reduces to the flat metric in the limit r → ∞. This metric reveals a peculiarity: where c^2r = 2GM, ds^2 = 0. This implies the existence of a horizon for which particles appear (to an external observer) to slow to a stop as they approach it. For a sufficiently high concentration of mass, spacetime becomes so curved that even light cannot escape from within a radius corresponding to the event horizon (named an "event horizon" since no events are observed beyond this point): this is a black hole. A simple Newtonian derivation is sufficient to calculate the location of this horizon, by equating the kinetic energy of a particle to the gravitational potential with the escape velocity equal to the speed of light:

$$R_S = \frac{2GM}{c^2} \qquad (1.5)$$

Many physicists thought this solution may just be a mathematical quirk of general relativity until observational evidence of black holes began to emerge in the 1960s. First, in 1963 Schmidt reported evidence of a quasi-stellar object with a redshift of ∼0.16 and an implied luminosity of 10^{59}erg s^{-1} (Schmidt, 1963); as discussed further in the following section, the most plausible mechanism to generate such extreme luminosities was accretion onto a SMBH. Over the next decade the X-ray source Cygnus X-1 garnered much attention. From the rapidity of X-ray fluctuations the source was known to be extremely compact, while the X-rays implied a gas temperature of millions of degrees; a radio counterpart was detected and used to pinpoint the source to a supergiant star; measurements of the doppler shift of that star showed it had an unseen massive companion. The radius implied by X-ray observations was found to be comparable to the Schwarzschild radius for the mass calculated from the orbit of the massive star, implying this was truly a BH. Over the following decades hundreds more such objects were discovered, and as the evidence mounted the common source was gradually accepted to be black holes. Today, SMBHs are thought to exist in the centres of most massive galaxies (Ferrarese & Merritt, 2000), stellar-mass BHs are a standard ingredient of stellar evolution, and the existence of BHs is established from detection of gravitational waves (Abbott et al., 2016). As such, there are likely ∼10^{20} BHs in the observable Universe.

Though the Schwarzschild metric is the simplest mathematical solution, being static and symmetric, astrophysical black holes are imbued with the properties of that which falls into them: they are extraordinarily unlikely to be static entities. Due to conservation of momentum and charge, real black holes can be charged and spinning. The mass of a BH impacts on the temperature of accreting gas since smaller BHs are surrounded by a stronger acceleration field (so temperature is inversely proportional to mass). BHs which form from the collapse of stars will tend to have high spin, since stars rotate and the BHs they evolve into are orders of magnitude smaller in radial extent. The spin of a BH affects the location of stable orbits (see section 1.3.2), in turn influencing the temperature of the inner disk (Reynolds, 2013). BHs are expected to have near neutral charge, since stars are approximately neutrally charged. Additionally, consider the process of a charged BH accreting gas: a significantly charged BH would attract particles of the opposite charge preferentially until neutrality is restored (Reynolds, 2013).

Classically, these three quantities - mass, charge and spin - entirely define a black hole. This fact unearths an apparent paradox: if a black hole is completely described by just three numbers, and whatever has fallen into it is inaccessible due to the existence of the event horizon, then its history is indeterminate; it has swallowed all information related to its formation. Whether the black hole formed from the collapse of a star, a cloud of gas or a herd of elephants is, although perhaps apparent, technically unknowable, complicating the process of discovering the origins of SMBHs. One of the core postulates underpinning quantum mechanics is that information is a conserved quantity. If information is lost inside black holes, unitarity is violated (probabilities no longer sum to one) and it can be argued that the very notion of causality breaks down.

One possible saving grace for commonly-held notions of reality is the theorised existence of Hawking radiation. In the early 1970s the laws of BH thermodynamics were developed. Bekenstein conjectured that since matter has an associated entropy, that entropy should be lost as matter falls into a BH, therefore to satisfy the 2nd law of thermodynamics BHs must have entropy (Bekenstein, 1973). An object with absolute zero temperature has zero entropy, and so it follows that BHs must have a finite temperature. Warm objects emit radiation; in 1974 Hawking calculated the temperature of radiation emitted from near the horizon of BHs (Hawking, 1974), emitted due to quantum fluctuations in a curved spacetime (a special case of the Unruh effect). Virtual particle-antiparticle pairs created from the vacuum near the horizon experience differing accelerations: one falls in while the other can escape. As the BH radiates it (extremely slowly) loses mass and the entropy stored at the horizon escapes back into the Universe, until eventually the BH may evaporate away entirely. The rate of radiation is negligible on the timescales under consideration in this thesis (a SMBH would take $\gtrsim 10^{100}$ years to evaporate). The Hawking radiation is emitted in a thermal mixed state, so seems at first

glance to be unable to solve the information paradox. However, since the BH entropy scales with the area, not the volume (as expected from intuition and studies of thermodynamics on Earth), and the notion of entropy is deeply connected to information about the microstates of a system, it is conjectured that information about the interior of the BH is stored on the event horizon. This information may be carried away in a scrambled state by small corrections to the quantum state of the Hawking radiation. The hypothesis that information is contained on the surface of BHs is generalised to arbitrary spacetime volumes in the holographic principle. The best studied form of the holographic principle is the AdS/CFT correspondence (Anti deSitter/Conformal Field Theory), where a direct mapping can be done between an interior gravitational spacetime and a gravity-free field theory bounding the spacetime (Maldacena, 1999). This may have huge implications for cosmology, particularly for understanding the nature of dark energy and/or dark matter (Easson et al., 2011; Verlinde, 2017). One intriguing consideration is the impact of accretion onto BHs on the interior spacetime, particularly in the context of BH cosmologies (e.g. Smolin, 1991; Easson & Brandenberger, 2001; Garriga et al., 2016).

Regardless of whether information is conserved by BHs, it is possible to learn about the growth of BHs through time by studying the light emitted as gas falls into them, and even to infer coarse details about the accretion histories from the spectral properties of that radiation. In the following section (section 1.3.2), details of the emission from infalling gas are presented.

1.3.2 Accreting Black Holes

When the first AGN were discovered there was great debate about what could cause such extreme observed luminosities (10^{59} ergs; Schmidt, 1963). The source could be a single supermassive star, but then emission would be expected to be isotropic and as a black body; it could be a dense cluster of accreting stellar black holes or neutron stars, but this could not explain the observed collimated jets. However, accretion onto a single black hole can satisfy these requirements and release energy:

$$L = \eta \dot{M} c^2 \tag{1.6}$$

where $\dot{M}$ is the mass accretion rate onto the BH and η is the radiative efficiency, related to the BH spin and typically of order $\eta \sim 0.1$ (see below). This efficiency is greater than 10 times the efficiency of nuclear fusion and implies the release of $\sim 10^{45}$ erg s^{-1} for an accretion rate of $1 M_\odot$ yr^{-1} - enough to explain the observations for reasonable accretion rates. The energy derives from the conversion of gravitational potential energy as gas falls from great distances toward the BH. As more and more AGN were discovered alongside apparent black holes in other guises (e.g. X-ray binaries), this explanation for the emission process slowly became

accepted.

SMBHs have grown to masses measured today predominantly via accretion of gas (Soltan, 1982). As gas flows toward the BH, as in many astrophysical scenarios, it tends to collapse to form a disk around the BH, conserving angular momentum. Although accretion proceeds via a disk an approximate upper limit to the rate of accretion can be estimated, called the Eddington limit or Eddington luminosity, by assuming spherically symmetric accretion. Photons emitted by infalling gas carry momentum and scatter off charged particles; this momentum is transferred to surrounding infalling gas and limits its inflow rate. If the radiation pressure exceeds the gravitational force this stops accretion onto the central object. An estimate of the luminosity at which this occurs can be derived by considering a spherically symmetric inflow and emitted photons scattering off electrons with the Thomson cross-section for interaction, σ_T:

$$F_{rad} = \frac{L_{Edd}\sigma_T}{4\pi r^2 c} = F_{grav} = \frac{GM_{BH}(m_p + m_e)}{r^2},$$

(1.7)

$$L_{Edd} \sim \frac{4\pi GcM_{BH}m_p}{\sigma_T}$$

where L_{Edd} is the Eddington luminosity, r is the distance from the BH and m_p is the proton mass ($m_p \gg m_e$, the electron mass). This gives $L_{Edd} \sim 10^{46}$erg s^{-1} for a $10^8 M_\odot$ BH. Although the Eddington luminosity describes a simplified scenario for accretion and therefore does not represent a true upper limit, the vast majority of BHs are observed to have sub-Eddington luminosity (Steinhardt & Elvis, 2010).

Despite their small angular sizes the structure of AGN is now relatively well understood and is depicted in figure 1.2 (Barthel, 1989). At the centre of the system lies the BH with typical mass $\sim 10^8 M_\odot$ (Soltan, 1982) and an event horizon large enough to engulf much of the Solar System ($\sim$2AU at this mass). Encircling the BH is a thermally emitting accretion disk, the engine of the AGN. Close to the BH the disk is extremely hot and blue, with temperatures of order 10^{5-6}K (Bonning et al., 2007), cooling with radial distance from the BH from UV to IR wavelengths. The electrons, being less massive than ions, are thought to scatter from the disk to form a corona of highly excited electrons. Photons from the disk then scatter off these electrons via inverse Compton scattering, meaning the photons extract some momentum from the high velocity electrons and are boosted to higher energies. This leads to an approximate power-law in the X-ray spectrum from AGN between $\sim$ 1 - 200keV (Lusso et al., 2009). Orbiting the BH beyond the accretion disk are clouds of gas which absorb and reemit light from the disk. They are distributed around the BH and are close enough to have rapid orbital velocities, of order 1000 - 25,000kms^{-1} (Peterson, 2006), so their emission lines are Doppler broadened,

hence this is called the broad line region (BLR). The X-ray corona illuminates the disk and BLR, this intense and hard radiation field produces exotic ionisation states, liberating even the most tightly bound electrons (creating a complex spectrum of iron emission lines for instance). Surrounding all this is a thick torus of dust which can absorb radiation from the disk, BLR and X-ray corona. Optical radiation is absorbed and reemitted in the infrared (IR) by the torus. Viewed through the torus, the disk and BLR are obscured from view (type 2 AGN), while observing from higher inclinations reveals the full thermal contribution of the disk and emission lines from the BLR (type 1 AGN). Self-gravitating dust clouds within the torus may obscure different parts of the emitting regions on short timescales, so rapid evolution can be seen in the observed spectral features of AGN (Nenkova et al., 2008). Though the existence of a torus is inferred from spectropolarimetry (light from the central source observed in polarised reflected light beyond the torus, with a spectrum akin to type 1 AGN; Antonucci & Miller, 1985) and multi-wavelength studies (such as Nenkova et al., 2008) the processes from which it forms are uncertain. It may be that radiation from the disk destroys (by photodissociation) dust in cones above and below the disk, leading to a torus-like dust structure. Ionisation cones in the narrow line region (see below) are seen emanating perpendicular to a dusty torus from some AGN (Pogge, 1988; Tadhunter, 1989). Alternatively, dust could be produced in the disk or BLR and blown by radiation pressure to greater radii (Elvis et al., 2002). In addition, there may be a correlation between the torus radius and AGN luminosity since photons from the AGN can destroy dust grains; a higher luminosity then destroys grains within a larger radius (receding torus, e.g. Lawrence, 1991; Arshakian, 2005). The receding torus model makes it likely that there is some redshift evolution in the relative populations of type 1 and 2 AGN with redshift, as the average luminosity of AGN increases with increasing redshift.

Radiation which escapes the immediate vicinity of the BH is then free to travel through and interact with gas throughout the galaxy. In comparison to the fast-orbiting gas making up the BLR this gas is slow, so when it absorbs and reemits AGN light the emission lines are narrow, hence gas throughout the galaxy makes up the narrow line region (NLR). High energy photons create high ionisation states of gas, so forbidden transitions such as [OIII] are observed in large quantities in the low-density ISM (Mullaney et al., 2013). Since NLR emission is driven by the central engine the luminosity of narrow emission lines provides a rough estimate of the luminosity of the central source (e.g. $L_{OIII} \sim \frac{1}{600} L_{AGN}$, Kauffmann & Heckman, 2009). NLR emission also contains information about the dynamics of gas within a galaxy from the shape of the lines (e.g. asymmetric lines can be indicative of outflowing gas Mullaney et al., 2013). Finally, $\sim$10% of AGN launch powerful radio jets on scales of parsecs to megaparsecs (though the fraction is mass- and redshift-dependent; Nesvadba et al., 2006; Jiang et al., 2007). Radio jets are thought to be launched from SMBHs as a result of a magnetised accretion disk interacting with BH spin via the Blandford-Znajek process (Blandford & Znajek, 1977); magnetic
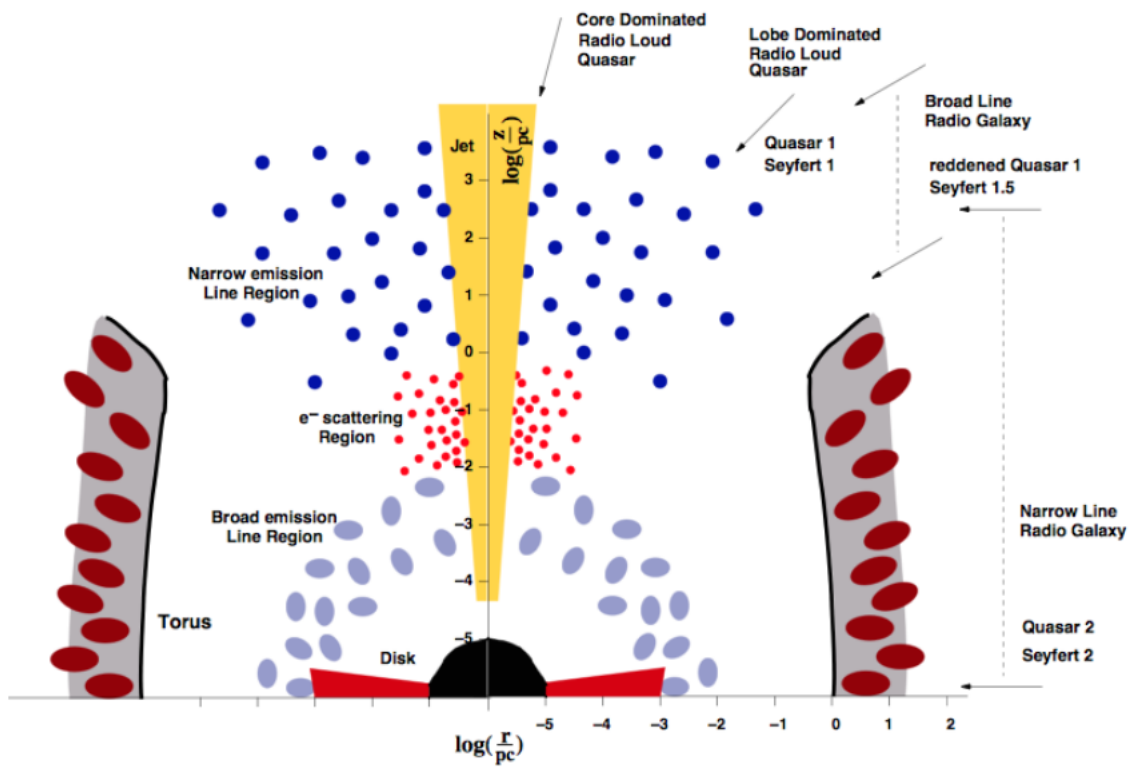
Core Dominated
Radio Loud
Quasar

Lobe Dominated
Radio Loud
Quasar

Broad Line
Radio Galaxy

Jet

$\log(\frac{z}{pc})$

3

2

1

0

Quasar 1
Seyfert 1

reddened Quasar 1
Seyfert 1.5

Narrow emission
Line Region

e⁻ scattering
Region

Broad emission
Line Region

Torus

Disk

Narrow Line
Radio Galaxy

Quasar 2
Seyfert 2

$\log(\frac{r}{pc})$

−5 −4 −3 −2 −1 0 1 2

Figure 1.2: Depiction of the geometry of AGNs proposed by Barthel (1989) including, from nearest the SMBH outwards, the accretion disk, the broad line region, the torus, the narrow line region and potentially kpc scale jets. The labelled arrows demonstrate the influence of line of sight on the detected emission (image from Biermann & Zier, 2008).

field lines near the BH are twisted up by its rotation and cascading electron-positron pairs ejected along the field lines extract this rotational energy. Relativistic charged particles then spiral along magnetic field lines in tightly collimated bipolar directions, releasing powerful radio synchrotron radiation. The most powerful radio sources tend to have the greatest physical extent (Tadhunter, 2016). Jets can plough into gas throughout the galaxy, shock heating and driving X-ray, UV and further radio emission via Bremsstrahlung radiation (Worrall, 2009). Jets can be relativistically beamed along the axis in the direction of motion, meaning that edge-on jets are much fainter and lobes often look one-sided.

Understanding the physical structure of AGN provides intuition for several techniques for identification of AGN from their SEDs. Some typical AGN SEDs are shown in figure 1.3. Since AGN tend to outshine their host galaxies in the UV, one simple detection method is to use the very blue continuum emission to identify AGN in the UV, though this requires the disk to be unobscured, i.e. for AGN to be of type 1. Similarly, the presence of very broad emission lines from multiple elements implies both the existence of rapidly orbiting gas and a hard ionising spectrum, both indicative of the BLR of an AGN. Broad emission lines can thus be used to identify AGN, but again the BLR needs to be unobscured from the perspective of an observer. Less dependent on viewing angle is X-ray emission. X-rays are only produced in the hottest, most energetic environments; the X-ray flux luminosity due to star formation rarely exceeds 10^{42}erg s^{-1} (Ranalli et al., 2003) and so X-ray luminosities above this level can be used to unambiguously identify AGN. Unlike continuum emission from the disk or emission line diagnostics from the BLR, X-rays can penetrate the dusty torus. If sufficient hydrogen lies between the AGN and observer it can obscure X-rays and AGN may go undetected in the X-rays (Compton thick, with column densities $N_H > 10^{24}$cm^{-2} Kocevski et al., 2015). If X-rays are obscured emission from the dust itself remains as a means of diagnosis. Dust in the torus absorbs optical and UV wavelengths and reemits them in the IR, to which the torus is largely transparent, as an approximate power-law, distinguished from the black body spectrum of stars (Stern et al., 2012). Magnitude differences between IR bands can resultantly be used to identify AGN regardless of obscuration (Stern et al., 2012). AGN can also be selected from optical emission throughout the galaxy: since the hard ionising spectrum will enhance emission from [OIII] and [NII] relative to Hβ and Hα in a galaxy hosting an AGN relative to normal galaxies, a BPT diagram can be used to identify AGN (Baldwin et al., 1981) (or other combinations of lines, e.g. SII, since more highly ionised species are typically present in AGN hosts but not in star forming regions due to the hardness of the ionising spectra). Finally, some AGN can be identified from their extreme radio luminosities from the jet. While star formation has a limited capacity to produce radio radiation, above some threshold luminosity (e.g. L$_{1.4GHz} = 10^{23}$WHz^{-1} Meurs & Wilson, 1984) star formation rarely produces such intensities and the majority of sources are likely to be AGN.
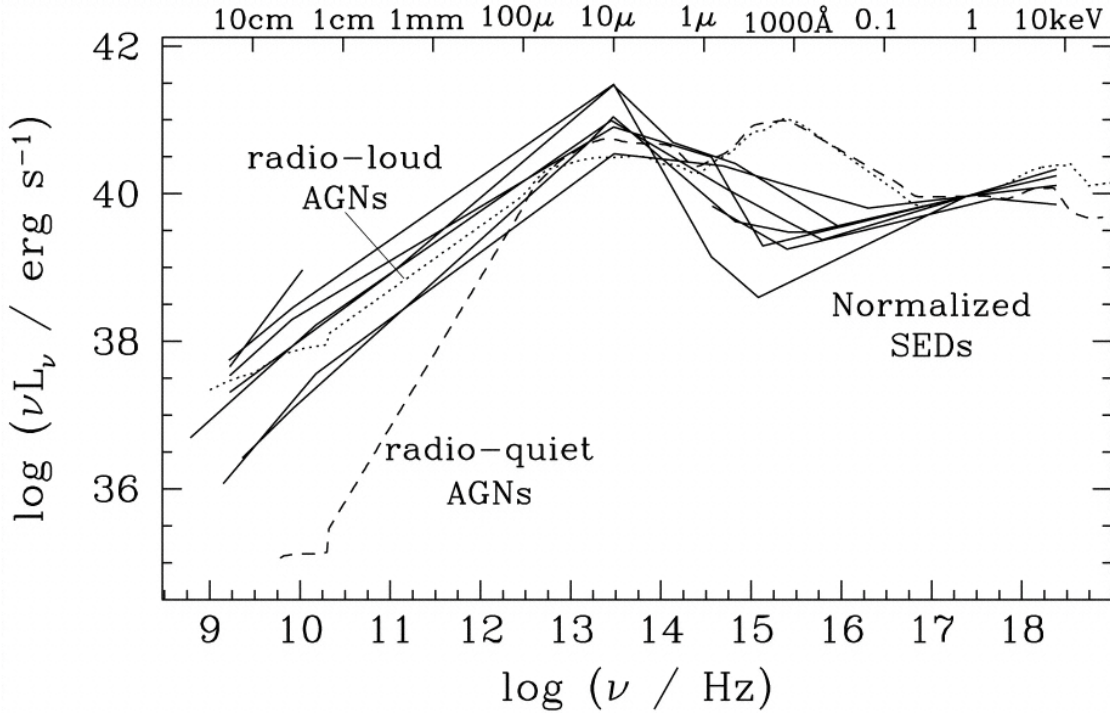
Figure 1.3: Typical AGN spectral energy distributions, taken from (Ho, 1999).

Critical to understanding the growth and evolution of SMBHs, relations between SMBHs and galaxies, and the various fuelling mechanisms for SMBHs is accurate measurement of their masses. In the centre of the Milky Way stars have been observed orbiting an unseen mass. From the velocities of the stars, Newtonian laws of gravitation can be used to calculate the mass of the central object, revealing a mass of $\sim 4 \times 10^6 M_\odot$ (Schödel et al., 2002). Other nearby SMBH masses can be estimated from stellar dynamics in the vicinity of the BH (Tremaine et al., 2002), but SMBHs are only gravitationally dominant within a radius $\sim GM_{BH}/\sigma^2$ ($\sim$10pc for a typical SMBH in a typical galaxy; Peebles, 1972) and so the exquisite angular resolution required for such measurements are unavailable for distant BHs. An alternate approach is to use time-resolution as a proxy for spatial resolution: reverberation mapping does this, assuming that variations in flux from the BLR are driven by variations in the continuum emission from the accretion disk. Continuous observations of the continuum emission reveal rapid changes in the luminosity, variations on the order of days to months, presumably driven by variation in the accretion rate onto the central BH. Light emitted from the disk takes a time, t, to travel to the BLR. The luminosity of broad emission lines are also observed to vary, with a time lag behind variation in the continuum on the order of days. The time delay between a change in continuum emission and the response from BLR clouds gives the distance from the black hole to the BLR (i.e. distances of light days). The BLR line width gives the velocity of the gas and so an accurate BH mass can be derived using Newton's laws ($M_{BH} \sim fct(\Delta v^2)/G$, where Δv is the line width, t is the time delay and f$\sim$1 is a factor depending on the geometry of the BLR), assuming BLR clouds are isotropically distributed about the SMBH and their motion is

due to the potential of the BH (Blandford & McKee, 1982). Line width (cloud orbital velocity) is found to correlate well with time delay, with $t \propto R \propto \Delta v^{-2}$ as expected from Newton's laws for orbit around a central mass (Peterson, 2006). Reverberation mapping provides some of the most accurate measurements of BH masses but is expensive, requiring long observing campaigns with high cadence.

Large samples of reverberation mapped AGN have revealed an empirical relation between the luminosity of the central source and the distance to the BLR (Kaspi et al., 2000, 2005; Bentz et al., 2006). The relation between BLR radius and AGN luminosity can be understood as due to a balance between ionisation and recombination in the BLR. Optimal conditions for line emission arise when there is a balance between ionisation and recombination for some gas, while the ionisation rate is proportional to the luminosity of the central source at a fixed radius. Suppose the continuum luminosity of the AGN increased suddenly - the BLR gas previously in ionisation equilibrium would now be highly ionised, electrons too dynamically hot to recombine, and so at this radius the emission line luminosity would be diminished. At a greater radius, where the flux is less, the gas would now be in ionisation equilibrium and would dominate the line emission luminosity, with a lesser FWHM due to the increased radius (Peterson, 2006). Since the AGN flux decreases with r^2, the distance between the BLR clouds which dominate the line emission and the SMBH is expected to scale with $\sim L^{0.5}$, approximately as observed (Kaspi et al., 2000, 2005; Bentz et al., 2006). Thus, the orbital radius of some broad emission line can be estimated from the luminosity alone. This has allowed for many BH masses to be estimated from measurement of the continuum luminosity and the width of some line in the BLR (e.g. Mg II; Wu et al., 2015), since a distance to the BLR can be estimated directly from the continuum luminosity instead of waiting days-months for a BLR response as per reverberation mapping. This is a far more feasible measurement than reverberation mapping of all AGN (though intrinsically dependent on those reverberation mapping results).

SMBHs acquired most of their mass (with measured masses in the range 10^6-$10^{10} M_\odot$; Ferrarese & Merritt, 2000) through accretion of gas (Soltan, 1982) and accretion can be heavily influenced by the BH spin. Frame dragging (the Lense-Thirring effect) is the phenomenon whereby a spinning BH seems to twist nearby spacetime around with its motion, such that stationary reference frames near a BH co-rotate with the BH. A distant observer would see the rotation of matter align with the direction of the BH spin and orbits precess around spinning BHs. The effect is analogous to swimming into a whirlpool: swimming against the vortex costs energy, swimming with the current allows one to approach the epicentre. Similarly, matter co-rotating with the spin of a BH can reach a stable orbit much closer to the centre of mass than matter orbiting a stationary BH or counter-rotating with respect to BH spin. Co-rotating gas moves faster than the converse case and has a correspondingly higher temperature, boosting

Relative spin	η	Last stable orbit
1	0.42	R_S
0	0.1	$3R_S$
-1	0.04	$9R_S$

Table 1.1: Table of BH spin relative to the direction of motion and the resultant radiative efficiency and location of the last stable orbit.

the blue thermal emission and emitting a greater fraction of the rest mass of the infalling matter in the form of photons (i.e. the radiative efficiency is a function of BH spin). As will be shown below, this can have an extraordinary effect on the growth of SMBHs.

The radiative efficiency varies from around $\eta = 0.04$ in the case of retrograde accretion onto a maximally spinning BH to $\eta = 0.42$ for prograde accretion onto the same BH. The radius of the last stable orbit and the radiative efficiency for 3 BH spins are shown in table 1.1. Since angular momentum is a conserved quantity, BHs retain some memory of their formation/accretion history in the form of spin. BHs are likely to form with some spin (in the collapse of a star or cloud, some small prior rotation will result in large spin velocities for a small remnant). Prograde accretion will spin up BHs and retrograde rotation will spin them down, but the relationship is non-linear: the radiative efficiency of the infalling gas is a function of spin, so retrograde accretion is more efficient at spinning down BHs than prograde accretion is at spinning them up. One dramatic consequence of the spin-dependent radiative efficiency of accretion onto BHs is on the mass growth rate (e.g. Volonteri, 2010):

$$M(t) = M_0 \exp\left(\frac{1-\eta}{\eta} \frac{t}{t_{\text{Edd}}} \right) \tag{1.8}$$

where, as in equation 1.6, η is the radiative efficiency, t_{Edd} is the e-folding timescale for a BH accreting at the Eddington rate, M_0 is the initial mass and M(t) is the mass after a time t. Massive BHs are already in place in the Universe by z∼7 (Mortlock et al., 2011): how did they come to be so massive so rapidly? Plugging maximum and minimum values of η into equation 1.8 (corresponding to prograde and retrograde accretion onto a maximally spinning BH respectively) for a BH with an initial mass of $10 M_\odot$ accreting at the Eddington rate (with $t_{Edd} = M_{BH}/\dot{M}$ ∼0.45Gyr) gives masses of ∼$21 M_\odot$ or $10^{23} M_\odot$ after ∼1Gyr. Since the spin-state of BHs can change as they grow such a range is implausible (as evidenced by a lack of observed BHs with such high masses), spins are likely to take more moderate values on average and accrete with an average radiative efficiency closer to η=0.1 (Volonteri, 2010; Fanidakis et al., 2012).

Though SMBHs seem to be near ubiquitous (Kormendy & Richstone, 1995; Ferrarese & Merritt, 2000), only a small fraction, dependent on redshift and luminosity, of SMBHs are

24

accreting at any moment ($\sim$1% in the local Universe; Fanidakis et al., 2012). The Soltan argument provides a simple estimate of the mass growth of AGN through time. From equation 1.6 it is clear that by measuring the luminosity function of AGN and assuming a value of η one derives an estimate of the mass accretion rate. The implied mass of SMBHs is around $10^8 M_\odot$ per massive galaxy. Since this is of order the measured mass of active BHs this implies accretion episodes are sporadic, have total durations of $\sim 10^8$ years per galaxy, and that all massive galaxies host SMBHs but only a small fraction of them are active at any one time (e.g. Soltan, 1982). The probability of a SMBH accreting at any one time is called the duty cycle. The duty cycle varies with the BH mass and redshift, peaking around z$\sim$2-3 for high mass galaxies, and with both large-scale and galaxy-scale environment (e.g. Fanidakis et al., 2012). As the Universe evolves gas fractions decline, structures grow to greater masses, the galaxy number density and cosmic matter density declines; it should come as no surprise that as the physical conditions of the Universe change, so too do AGN statistics. Another factor which limits and possibly regulates the growth of SMBHs is feedback from AGN. This is discussed in the following section before connections between AGN and galaxies are laid out in section 1.4.

1.3.3 Feedback

Cosmological simulations can be used to identify previously unknown processes influencing galaxy evolution. Extrapolating from dark matter only simulations, assuming dark matter directly traces the distribution of baryonic matter, predicts the existence of a higher space-density of massive galaxies than is observed (Frenk & White, 2012). This most likely implies that some baryonic physics limits the growth of massive galaxies. It is now widely thought that feedback from supernovae and AGN limits the growth of galaxies and causes deviation from the simple dark matter extrapolation. In the Milky Way young massive stars are observed to create cavities in the clouds they form within, blowing gas away from the star by radiation pressure as fusion ignites in the core. Analogous processes could occur on galaxy length-scales, driven by the extraordinary energy input from AGN and supernovae. Additionally, SMBHs are found to have masses correlated with their host galaxy bulge masses or velocity dispersions (Kormendy & Richstone, 1995; Ferrarese & Merritt, 2000; Gebhardt et al., 2000, M$\sim \sigma^4$, see section 1.4). The cause of this relationship is unknown, but one appealing possibility is that the energy released during periods of BH growth couples to surrounding gas, limiting the growth of the galaxy and itself simultaneously (e.g. Di Matteo et al., 2005).

For a $10^8 M_\odot$ BH the energy liberated in its growth is $\sim 10^{61}$erg (from equation 1.6), while the binding energy of the associated galaxy is $M_{bulge}\sigma^2 \sim 10^{59}$erg (assuming $M_{BH} \sim 10^{-3} M_{bulge}$ and a velocity dispersion, σ=300kms^{-1}; Silk & Rees, 1998). The energy released by accreted gas is sufficient to destroy a galaxy, so long as the liberated energy has some way of coupling to the galaxy. Of course, most of the gas in a typical galaxy is locked up in stars and clouds,

hence is afforded some shielding from the radiation, while the energy output is unlikely to be spherically symmetric and so galaxies survive. The influence on the galaxy and surrounding environment depends on whether the outflow is momentum-driven or energy-driven: if outflowing gas cools efficiently the gas will exert negligible thermal pressure on the ISM but can still drive outflows by transferring momentum to the ISM (momentum driven); if outflowing gas is unable to cool it will exert thermal pressure on the surrounding ISM, expanding adiabatically (energy driven).

Consider a spherically symmetric inflow of gas onto a SMBH. Gas flowing in encounters an ionising flux from gas in the vicinity of the black hole; ionised gas resultantly becomes optically thick (as seen as P-cygni profiles in X-ray absorption spectra; Pounds et al., 2003) and the momentum of continuum photons is entirely transferred to the gas (King, 2003):

$$\dot{M}_{out} v \sim \frac{L_{Edd}}{c} \tag{1.9}$$

where $\dot{M}_{out}$ is the rate of mass outflow, flowing with velocity v, due to an AGN accreting at the Eddington rate. Following the arguments of King (2003), the mechanism by which electrons in the outflow cool is due to Compton scattering off of photons from the AGN with a cooling time:

$$t_c \propto \frac{1}{U_{rad} v^2} \tag{1.10}$$

where U_{rad} is the radiation density (dominated by thermal disk emission), proportional to L_{Edd}/R^2. Somewhat counter-intuitively this means that the more luminous the AGN and the closer the outflowing gas is to the BH, the faster the outflowing gas cools. Since the Eddington luminosity of an AGN is proportional to the BH mass, this directly ties the cooling rate to the BH mass with an inverse relation:

$$t_c \propto \frac{R^2}{M_{BH} v^2} \tag{1.11}$$

for a sufficiently small BH, gas can sink back toward the BH to be accreted or reheated, fuelling continuous episodes of growth. To calculate the point at which gas is expelled (i.e. v>v_{esc}) we need to derive the velocity of a momentum-driven shell of gas. The mass of the shell is M = $4\pi R^2 \rho$, where ρ is the density; the mass flux is $\dot{M}_{out} = 4\pi R^2 \rho$v. Under the assumption that all the momentum of continuum photons is transferred to the gas we can use equation 1.9:

$$\dot{M}_{out} v = 4\pi R^2 \rho v^2 = \frac{L_{Edd}}{c} \tag{1.12}$$

differentiating the momentum yields the force:

$$\frac{d}{dt}[M(R)\dot{R}] \sim M(R)\ddot{R} = 4\pi R^2 \rho v^2 = \frac{L_{Edd}}{c} \qquad (1.13)$$

where the term from the differentiation, $\dot{M}\dot{R}$, is omitted, since if significant ISM is entrained (at large t) $\dot{M}\dot{R} \ll M(R)\ddot{R}$. Integrating twice gives:

$$M(R)R \sim \frac{f_g \sigma^2 R^2}{G} = \frac{L_{Edd}}{c}t^2 \qquad (1.14)$$

where the gas mass M(R) has been substituted for the product of the gas fraction f_g and the Newtonian mass for the galaxy potential. Now we find the velocity:

$$\frac{R^2}{t^2} = v^2 = \frac{G}{f_g c}\frac{L_{Edd}}{\sigma^2} = \frac{4\pi G^2 m_p}{f_g \sigma_T}\frac{M_{BH}}{\sigma^2} \qquad (1.15)$$

where in the final form L_{Edd} has been replaced by the expression in equation 1.7. Once the BH grows to some critical mass the winds may entirely escape the BH potential, entrain the ISM and drive galactic-scale winds of the kind seen commonly (Mullaney et al., 2013). To see how the M-σ relation may be established we can consider at what point a momentum-driven shell of gas approaches the galaxy escape velocity in equation 1.15. Given a large enough supply of gas the wind velocity, v, will approach the mean velocity of the stars in the galaxy, σ, as the BH grows. Replacing v with σ in equation 1.15 gives:

$$M_{BH} = \frac{f_g \sigma_T}{4\pi G^2 m_p}\sigma^4 \qquad (1.16)$$

directly relating the BH mass to the galaxy velocity dispersion (and therefore indirectly to the mass), as is observed (Kormendy & Richstone, 1995; Ferrarese & Merritt, 2000; Gebhardt et al., 2000). Shells of swept-up matter may go on to form stars in the bulge (King, 2003) or, if v is high enough, escape the galaxy entirely. It is interesting to speculate that this could contribute to the inside-out growth of galaxies - where the SFR density within galaxies is proportional to radial extent (e.g. Ellison et al., 2018). As the BH attains mass successive wind shells are expelled to greater radii, where they form new stars. Bulges would be built up over cosmic time from the inside out. However, there are numerous mechanisms driving star formation, and cooling flows, central starbursts following a merger, or disk instabilities could well explain such inside-out growth. If the momentum-driven shell escapes the BH vicinity before significant cooling occurs it can expand adiabatically in efficient energy-driven winds which may entirely escape the galaxy (King, 2003). If sufficient gas is expelled the SFR will be depleted, severely impacting galaxy evolution.

Jets launched from AGN (via the Blandford-Zjanek process, section 1.3.2) escape the BH potential at near light-speed, effectively uncooled by the central source. These can plough into the surrounding ISM and impart energy, shock-heating the gas and driving a thermal pressure which inflates hot bubbles of gas (e.g. Fabian, 2012). Jets can extend from parsec-scales to megaparsec-scales and represent a highly efficient heating mechanism, heating gas through the ISM and intergalactic (IGM) or intercluster medium (ICM). The energy of jets can be estimated from the PdV work done: the product of the size of the bubbles and the pressure exerted by the surrounding gas (e.g. Rafferty et al., 2006). Hot, low density bubbles rise through the cool and dense surrounding medium, like air bubbles in water, spreading out and dissipating energy throughout the ICM, slowing the rate of gas cooling. Cool gas will fall back onto galaxies to form new stars; this heating effect greatly reduces the SFRs of some radio-loud galaxies (Fabian, 2012). In the following section the significance of feedback on galaxy formation and evolution is expounded upon, before finally coming to the explicit role of galaxy mergers in SMBH growth.

1.4 Links between AGN and Galaxies

1.4.1 AGN and Galaxies

Gas in the early Universe collapses under gravity from near homogeneity to form the galaxies seen today. These galaxies exhibit a diverse range masses, morphologies, colours and sizes; interdependent properties which evolve over cosmic time. Their diversity reflects the complexity of the evolution of galaxies. What factors determine galaxy properties? What role does the AGN play in galaxy evolution, and what role do galaxy properties play in fuelling the SMBH? There exist a set of (related) relations between galaxies and the SMBHs they host: the M_{BH}-σ relation (a log-log linear correlation between the BH mass and the velocity dispersion of the stars in the bulge; Ferrarese & Merritt, 2000; Gebhardt et al., 2000), the M_{BH}-M_{bulge} relation (as before but directly sampling the bulge mass, instead of indirectly via velocity dispersion Marconi & Hunt, 2003; Haring & Rix, 2004), the M_{BH}-L relation (between the black hole mass and galaxy luminosity Kormendy & Richstone, 1995) and the M_{BH}-M_{halo} relation (correlation with the mass of the dark matter halo Ferrarese, 2002). Given that SMBHs are gravitationally dominant over only a tiny fraction of the size of the galaxy, and gas has to lose an overwhelming fraction of its angular momentum to accrete onto the BH, the origin of these correlations are perplexing. The existence of such relations have been taken as evidence of coevolution between the SMBH and galaxy. In this section the properties of galaxies are established and some of the relations underpinning their evolution elucidated, including the possible role of feedback from AGN. In the following section the processes influencing the growth of SMBHs are discussed, with particular focus on the role of galaxy mergers.

Cosmic structures form and grow hierarchically. Small galaxies in the early Universe collide to form more massive galaxies, which in turn merge to form larger galaxies in an ongoing process. Gas, making up the bulk of baryonic mass in the early Universe (before many stars have formed), is highly dissipative and will tend to collapse into disks to conserve angular momentum. Stars, contrarily, are essentially collisionless. Galaxy mergers are violent events, tending to disrupt morphologies, transforming disky galaxies into irregulars or ellipticals, while supplying gas to induce starbursts and possibly trigger nuclear activity (see section 1.4.2; Toomre & Toomre, 1972; Di Matteo et al., 2005). In the early Universe, where gas makes up the bulk of baryonic mass (before many stars have formed), galaxies are gas rich enough that they tend to settle back into disks following mergers (most galaxies are disky or irregular at high z, with few ellipticals, despite all presumably having undergone many mergers; Robertson et al., 2006b; Buitrago et al., 2013). Galaxies are low-mass and numerous and mergers are resultantly common. Lower mass galaxies are, on average, more gas rich and disk-dominated, have higher specific SFRs (sSFR; the star formation rate divided by galaxy stellar mass) and are resultantly bluer than higher mass galaxies. Later in the history of the Universe much of the gas has become locked up in stars, which do not easily dissipate their angular momentum, and so mergers are more likely to irredeemably destroy disks, particularly for mergers of massive galaxies which tend to be gas-poor. The likelihood of a galaxy being massive, red and elliptical increases as the Universe evolves in time. These factors combine to create mass-morphology-SFR-colour relations, encapsulated in figure 1.4.

Such relationships are of course probabilistic, but they are informative of the processes which transform galaxies. The relationship between mass and colour is exemplified by a colour-magnitude plot, as shown in figure 1.5, with a red cloud, a blue sequence and not many galaxies between (density contours on this figure are on a logarithmic scale). This is often taken as evidence that these galaxies are rapidly progressing from blue star-forming galaxies to red and dead galaxies, i.e. their star formation is shutting off rapidly as if quenched by some process. Hence the conjecture (discussed in the following section, 1.4.2) that major mergers trigger periods of SMBH growth, the resultant feedback drives gas out of the galaxy and quenches star formation, rapidly pushing galaxies from the blue sequence through the green valley to the red cloud. The notion that there is a sharp distinction between blue sequence and red cloud galaxies is called into question by Herschel space telescope IR observations of a continuum of specific star formation rates across morphological classes and stellar masses (Eales et al., 2018), shown in figure 1.4 suggesting instead that galaxies pile up in the red cloud when the mass fraction locked up in low mass stars reaches some threshold (with $M_* \gtrsim 10^{10.5}$), outweighing the blue emission from any newly formed massive stars.

The influx of energy from stars and AGN is extremely important in shaping the detailed evolution of galaxies. Figure 1.6 depicts the stellar mass component of galaxies as a function of

halo mass, predicted from models with and without feedback from stars and AGN, compared to semi-empirical relations (matched to observations). At low halo mass supernovae ejecta have escape velocities exceeding that of the galaxy and gas can be swept out of the galaxy, limiting the SFR. Massive halos have potential wells too deep for star formation to expel gas, requiring the more energetic AGN feedback (section 1.3.3) to match theory with observation. This alters the galaxy mass function, resulting in fewer low mass and high mass galaxies than naively expected by assuming baryons follow the dark matter. One consequence of this feedback is establishment of a galaxy mass-metallicity relation: metal enriched gas from star formation is expelled in supernovae explosions from low mass galaxies, while supernovae ejecta fail to escape massive galaxies and so metals from star formation are retained (Tremonti et al., 2004).

Radio loud AGN are generally associated with massive elliptical galaxies in dense environments (Hickox et al., 2009) and the source of the radio emission is jets. Jets (typically from massive ellipticals Tadhunter, 2016) can drive efficient feedback. The ICM is comprised of hot X-ray emitting gas, gas which has previously been expelled or stripped from infalling galaxies (pressure from the hot ICM strips the ISM from infalling galaxies, resulting in SFR gradients as a function of radial extent in galaxy clusters known as environmental quenching, e.g. Biviano et al., 1997). X-rays carry energy away from the gas, causing it to cool. As gas cools it sinks back toward galaxies in cooling flows where, presumably, it could be recycled to form new stars. The cooling rate implies that of order $\sim 1000 M_\odot \mathrm{yr}^{-1}$ should fall back onto massive elliptical galaxies in clusters, yet the observed SFRs are far below this value (Fabian, 2012). The gas cooling is likely offset by heating by radio feedback. Multiple ripples are seen in X-ray images of the ICM in some massive galaxy clusters, where successive bubbles (X-ray cavities) have inflated due to heating by jets (Fabian, 2012). The effect is to curtail the stellar mass of the most massive galaxies and their BHs, by maintaining a hot ICM which can scarcely feed new star formation or accrete onto BHs.

Feedback is a compelling mechanism of directly linking the masses of galaxies and BHs, but it is possible these relations are acausal. Both SMBHs and star formation rely on a supply of cold gas. If some near-universal fraction of gas which falls into galaxies descends to the centre to be accreted onto the BH, or if some fraction of stellar wind (which is emitted isotropically and so some fraction of wind can be stationary with respect to the SMBH at the point of emission) naturally falls to the centre (e.g. Wild et al., 2010), the masses can be naturally but non-causally correlated. In the following section we explore the hypothesis that galaxy mergers supply gas for both the SMBH and star formation, building the bulge and BH simultaneously.
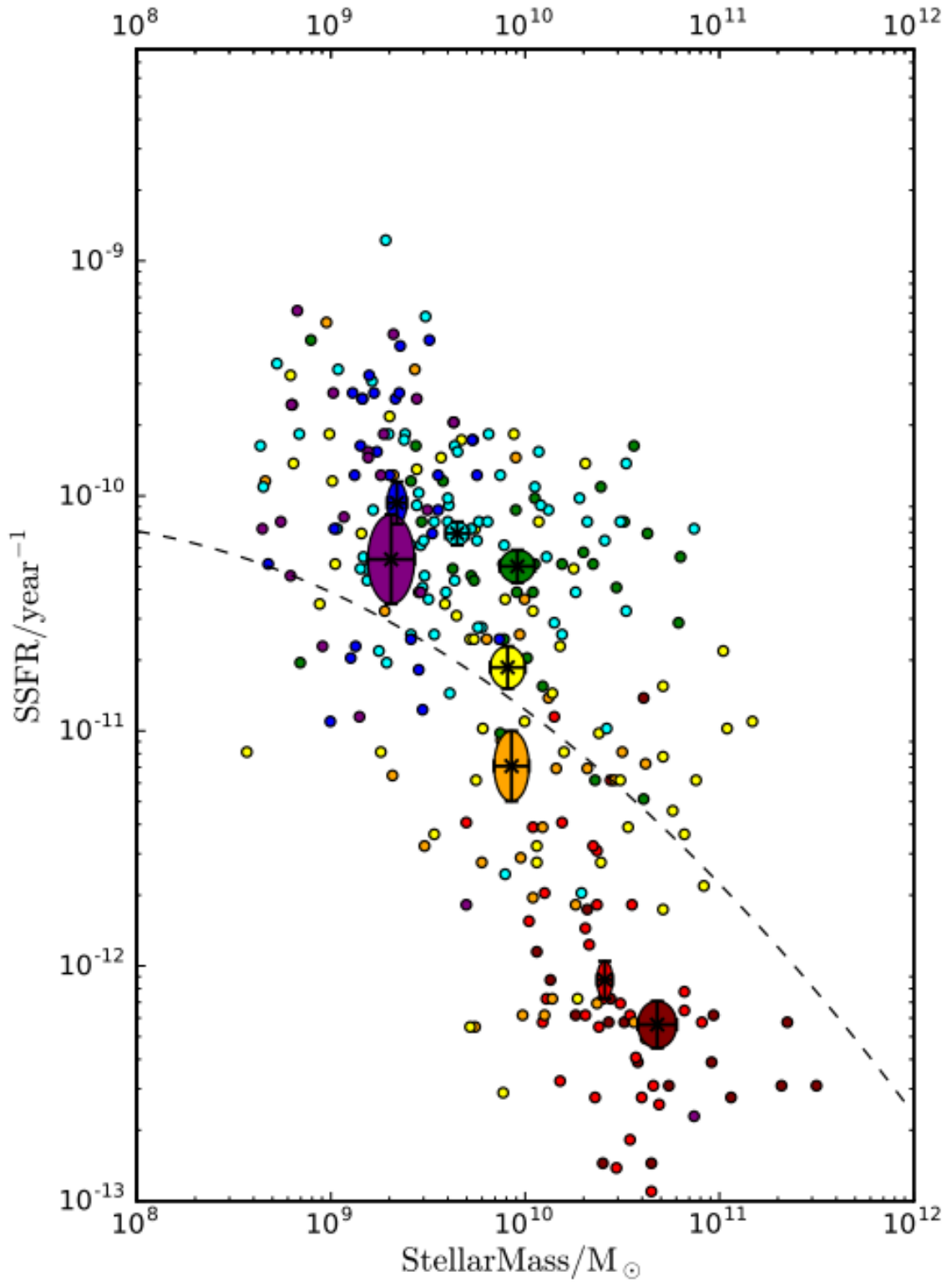
Figure 1.4: The decline in sSFR with increasing stellar mass, split by morphology. Colours indicate galaxy morphology: ellipticals - maroon; S0 - red; S0a and Sa - orange; Sab and Sb - yellow; Sbc - green; Sc and Scd - cyan; Sd and Sdm - blue; Sm, Im, I and I0 - purple. Large symbols represent the 1σ error on the mean morphology class. Taken from (Eales et al., 2018).
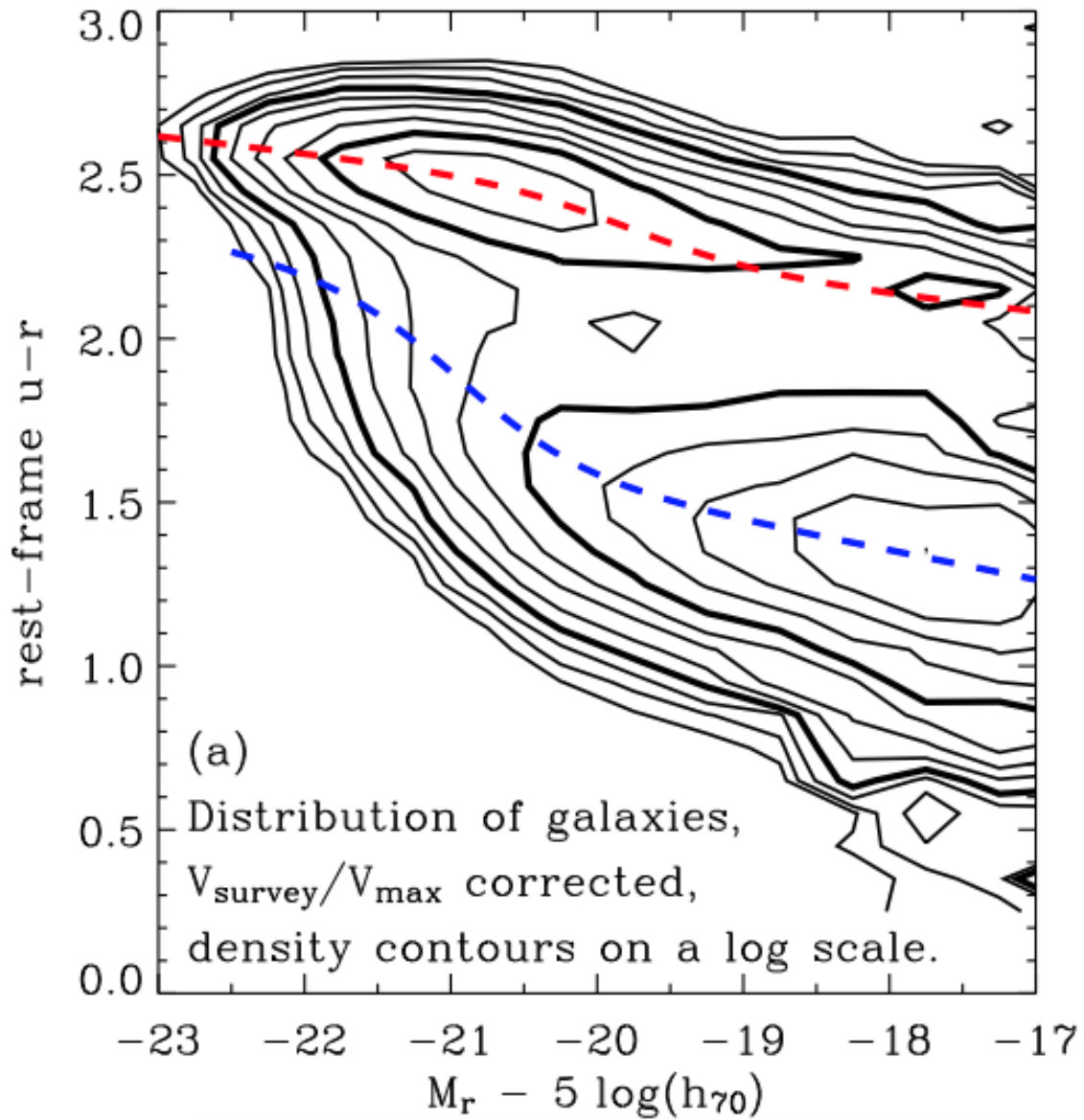
31

Figure 1.5: A colour-magnitude plot of ∼70,000 SDSS galaxies, taken from (Baldry et al., 2004). Density contours are logarithmic and red- and blue-dashed lines show colour-magnitude relations for the red and blue clouds respectively.
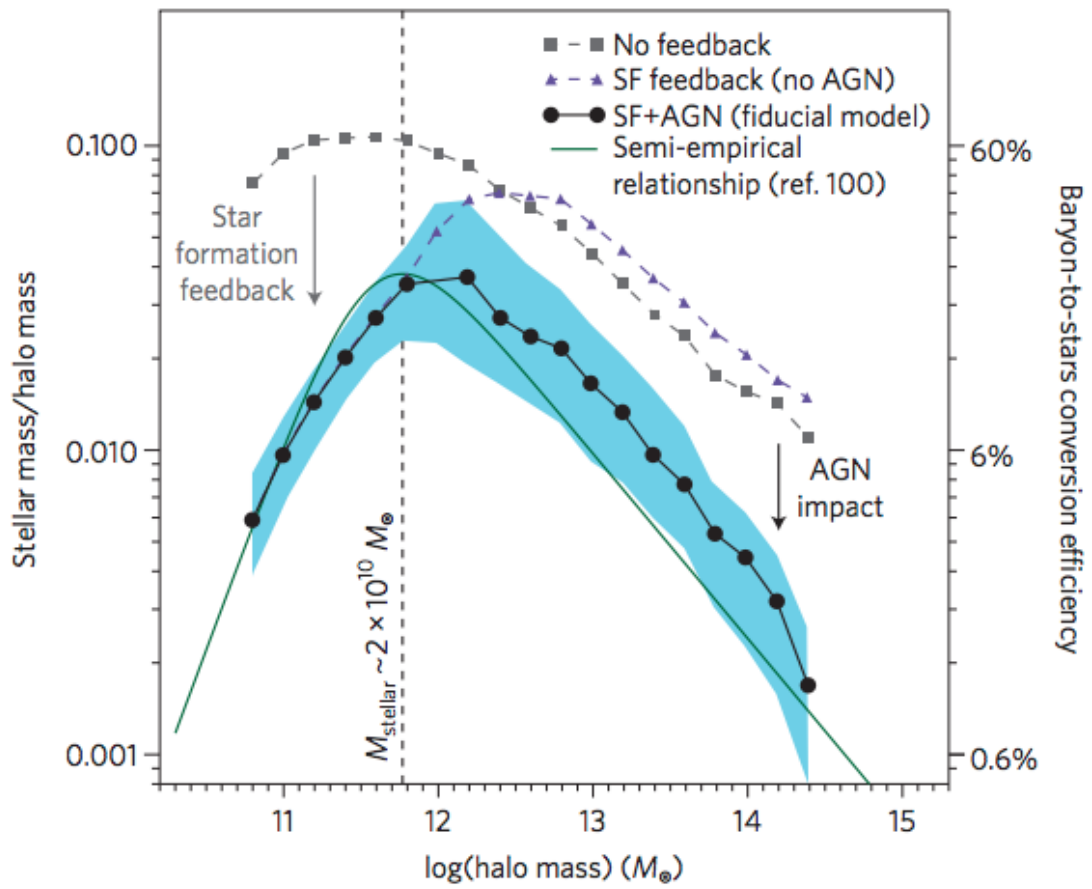
Figure 1.6: The efficiency with which baryons are converted to stars as a function of halo mass, with and without feedback, taken from (Harrison, 2017)

1.4.2 AGN and Galaxy Mergers

Galaxies form hierarchically, meaning galaxies merge to form fewer more massive galaxies through cosmic time. Mergers tend to destroy galaxy disks, morphologically transforming galaxies into ellipticals with little star formation (Toomre & Toomre, 1972; Barnes, 1988, 1991; Di Matteo et al., 2005; Cox et al., 2008), though particularly gas-rich galaxies may reform a disk component (Robertson et al., 2006b). This morphological transformation can take $\gtrsim$ a few 10^8yrs (e.g. Lotz et al., 2010a), offering a window of opportunity in which to detect cosmic collisions. It is clear from electromagnetic signals that most SMBH growth is acquired from infalling gas, despite apparent difficulty in losing such a large fraction of the angular momentum of the gas (Soltan, 1982). Galaxy mergers may facilitate such rapid inflows of gas while establishing the the M-σ relation: mergers drive gas toward galaxy centres where it can be accreted onto the central BH (e.g. Sanders et al., 1988; Di Matteo et al., 2005), trigger nuclear and global starbursts (Mihos & Hernquist, 1994, 1996; Hopkins et al., 2007) and grow the galaxy in tandem with the SMBH. In addition, if these galaxies contain massive black holes then some time after a galaxy merger these BHs are generally expected to fall to the new global centre under dynamical friction: BHs create a wake of stars behind them accelerating them in a direction opposed to their motion until they come to rest in the centre. Through interaction with gas and stars the two massive BHs may come to orbit one another, followed by eventual inspiral via emission of gravitational waves to merge. Thus BHs too are likely to grow hierarchically, with some delay relative to the galaxy mass-growth (due to the additional time required to merge after the galaxies merge). For efficient dynamical friction a SMBH has to be in a relatively dense stellar or gas field, and for gravitational radiation to become significant BHs have to get into a very close orbit, so it is not yet clear how often SMBHs merge. Future gravitational wave detections from LISA and/or pulsar timing arrays should constrain this rate (Abbott et al., 2016; Yunes & Siemens, 2013).

That galaxy mergers drive nuclear activity is an appealing theory since AGN accrete large fractions of their mass in bursts of rapid accretion ($\gtrsim 0.1 M_\odot yr^{-1}$; Croton et al., 2006), requiring rapid inflow of gas from galaxy length-scales. Stripping the gas of enough angular momentum to allow for such rapid accretion, thereby powering the most luminous AGN, proves extremely challenging. Theoretical work suggests major mergers can provide the necessary torque, allowing for the highest accretion rates onto central black holes whilst transforming galaxy morphologies (Toomre & Toomre, 1972; Barnes, 1988, 1991; Di Matteo et al., 2005; Cox et al., 2008). Galaxies are difficult to accurately simulate because of the range of relevant size-scales (sub parsec to megaparsecs) and complexity. The resolving power of astrophysical simulations has been growing exponentially since their genesis, but still the spatial and mass resolution of individual simulated galaxies is far too coarse to allow the detailed physics of
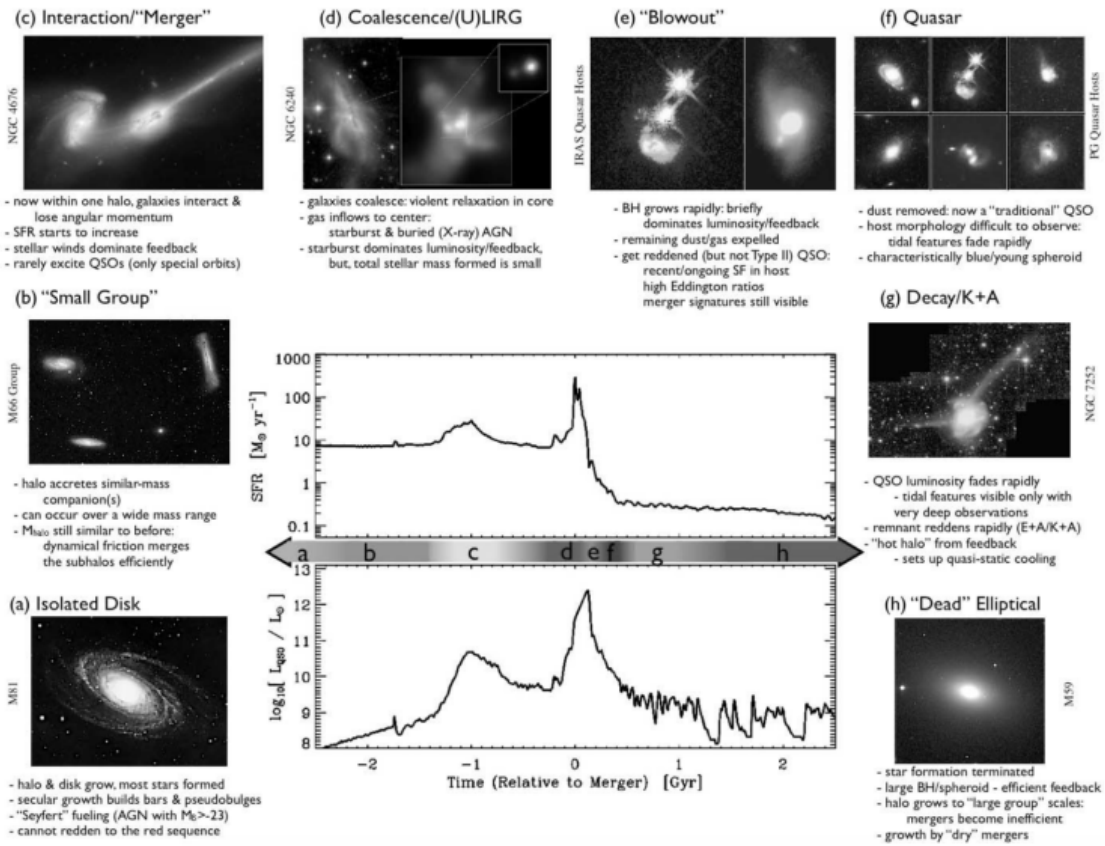
Figure 1.7: Hopkins evolutionary cartoon, taken from Hopkins et al. (2008).

accretion or star formation to emerge, instead the "sub-grid" physics, governing the complex processes within galaxies, is included in the form of analytic prescriptions. Empirical relations can be inserted when known (e.g. the average SFR for a given gas surface density Kennicutt, 1989) and simulations rescaled to reproduce observables in the local Universe, such as the observed stellar mass densities (Vogelsberger et al., 2014a; Schaye et al., 2015). The extent to which this improves understanding of the processes driving evolution of simulated components is somewhat dubious; for instance, often gas in simulations is assumed to accrete onto the SMBH if it crosses into the cell containing the SMBH, while in reality it is likely that some of this gas would escape (Sijacki et al., 2015). Nonetheless, the first simulations of isolated galaxy mergers which included accretion onto SMBHs were instrumental in shaping thought about galaxy evolution and the role of BHs. Di Matteo et al. (2005); Springel et al. (2005) found that SMBHs have greatly enhanced accretion rates following a galaxy merger, acquiring most of their mass in that interval, before driving powerful winds which shut off star formation and further BH growth. Much of this growth may be optically obscured by thick columns of gas and dust (Hopkins et al., 2006) and merger-driven accretion and subsequent feedback can naturally provide a mechanism for the bimodal colour-mass distribution of galaxies, as depicted in the Hopkins cartoon of figure 1.7.

Understanding the significance of galaxy mergers on AGN populations is problematic for isolated galaxy simulations; it is desirable for correlations between SMBHs and their host galaxies to emerge naturally in cosmological simulations, to match observations over a wide range of mass and length scales instead of for artificial initial conditions. Purely gravitational N-body simulations of dark matter (e.g Press & Schechter, 1974; Skillman et al., 2014) have greatly developed understanding of the assembly of large scale structure, while placing constraints on the nature of dark matter and enabling development of the ΛCDM cosmological model. Most astronomy, however, concerns itself with the distribution and properties of baryons which make up the gas, stars and black holes in the universe, and about which dark matter-only simulations can tell us little. Hydrodynamical simulations consider both gravity and baryonic physics, including a wide range of physical processes governing the formation of stars, the growth of black holes, stellar and AGN feedback etc. The cosmological hydrodynamical simulations of today (AGN-Horizon, Illustris, Eagle; Dubois et al., 2014; Vogelsberger et al., 2014a; Schaye et al., 2015) track these dark matter and baryonic components consistently as they interact (through gravitational and feedback processes) and evolve through time. The resolution problem is greatly exacerbated relative to individual-galaxy simulations, with mass resolution elements $\gtrsim 10^{5-6} M_\odot$, and remains insufficient to probe individual star forming regions or the vicinity of black holes. The spatial resolution also suffers, with cells typically $\gtrsim 100$pc. Assuming circular motion at 300km s^{-1} about the galaxy centre, a gas element with unit mass (m=1kg) 100pc from a SMBH has angular momentum $\vec{L} = m\vec{r} \times \vec{v} \sim 10^{24}$kg m^2 s^{-1}, but once it reaches the BH this reduces to L $\sim 10^{20}$kg m^2 s^{-1}. A standard approximation in cosmological simulations is that gas which falls into the same cell as the SMBH is instantly accreted. The calculation above demonstrates that this may not be a realistic assumption, since to do so it first must lose $\gtrsim 99.99\%$ of its angular momentum. Models or empirical relations governing the sub-grid physics are used instead of direct simulation, and massive BH seeds are inserted by hand. The trade-off is that these simulations are now large enough to simulate representative volumes of the universe (up to halos around $10^{14} M_\odot$ Vogelsberger et al. (2014a)), containing 10,000s of galaxies for statistical comparison to real data.

Modern cosmological simulations broadly reproduce many of the observed trends seen in the universe. The cosmic evolution of star formation rate density, the galaxy luminosity function and the proportions of early- and late-type galaxies all match well to data (Vogelsberger et al., 2014a,b). In the Illustris simulation galaxy mergers are found to boost BH accretion rates for some time after onset of the merger (Sijacki et al., 2015), but the enhancement is weak and related to the star formation tracing the availability of cold gas, rather than being directly merger-triggered. In fact, significant evolution is seen in the M-σ relation through cosmic time, with BH growth preceding galaxy growth (Sijacki et al., 2015), contrary to expectations of merger-driven accretion where the BH grows in tandem with the galaxy or slightly

delayed relative to stellar mass growth. At low redshift the merger of quenched galaxies, "dry mergers", leads to a steepening of the M-σ relation, growing the stellar and BH mass without significantly increasing the velocity dispersion (Sijacki et al., 2015). However, it is less than clear whether this holds true in the real Universe. Simulations couple a fine-tuned fraction of AGN luminosity to the interstellar gas (typically around 5%) to recreate the SFR density or stellar mass function in the local Universe (Vogelsberger et al., 2014a; Schaye et al., 2015), as required to get the correct fit but without a great deal of theoretical justification. This reduces the accretion onto SMBHs in simulations in turn.

To gain insight into the populations of SMBHs in the real Universe and their relation to larger structures (galaxies, local galaxy environments) semi-analytic models (SAMs) can be employed. SAMs alleviate much of the need for expensive calculations by combining observed relations, such as the SFR density (Kennicutt, 1989), with theoretical considerations, such as energy conservation. Typically dark matter halos are modelled with the expected halo mass function and galaxy formation modelled within the associated merger trees (e.g. Somerville & Primack, 1999). The evolution of baryons are modelled to describe star formation, feedback, gas cooling etc. Those parameters in the model which are poorly-constrained can be left as free parameters of the model, adjusted to match observational data (Somerville et al., 2008). Not all features of the Universe are required to emerge from fundamental processes in SAMs, freeing up computational power to focus on particular aspects of galaxy evolution. For instance, the growth rate of SMBHs as a function of mass and time is reasonably well constrained from the observed AGN luminosity density, but the fuelling process is not yet understood. By considering merger-induced fuelling of AGN with a physically motivated distribution of accretion rates and radiative efficiencies (as functions of mass and redshift), Shankar et al. (2013) show that including BH mergers (assuming a galaxy merger leads to a BH merger) over-predicts the number density of high mass BHs. This perhaps suggests mergers of massive galaxies do not lead to BH mergers, or else that massive BHs have very high radiative efficiencies (thus grow more slowly from accretion, see section 1.3.2). Such an approach can also demonstrate whether mergers alone can account for all nuclear activity, and what is required of mergers in order to do so. There are insufficient major mergers to fuel the low luminosity AGN population in the local Universe and so minor mergers or non-merger processes must be invoked to explain the observations (Shankar et al., 2013). Modelling a range of fuelling processes to fit the observed parameters can provide insight into possible mechanisms in a physically well motivated and computationally efficient manner (Croton et al., 2006; Somerville et al., 2008; Shankar et al., 2009; Draper & Ballantyne, 2012; Shankar et al., 2013), whilst informing observers about where in the parameter space significant processes may be occurring, i.e. what to observe (e.g. minor mergers).

Low levels of accretion onto BHs are expected without galaxy mergers. There are too

many low luminosity AGN to be plausibly triggered by major mergers, but minor mergers remain as a largely unexplored triggering mechanism due to the difficulty of detecting them. Major mergers and minor mergers are generally separated by a (arbitrary) mass ratio cut: typically mass ratios <4:1 are called major mergers and mass ratios >4:1 are called minor mergers (e.g Kaviraj, 2013; Rodriguez-Gomez et al., 2015a) Minor mergers are far more common than major mergers (Rodriguez-Gomez et al., 2015a) (due to the shape of the galaxy mass function, where galaxies become increasingly rare with increasing mass) and may drive much of the low Eddington-ratio nuclear activity, or even rapid accretion (Kaviraj, 2013), in addition to star formation throughout the Universe. Otherwise disk instabilities (Bournaud et al., 2011), bars (Knapen et al., 2000; Oh et al., 2012) and accretion of cold gas from the cosmic web (Sabater et al., 2013) or cooling flows (Fabian, 2012) may provide SMBHs with the necessary supplies of cold gas to account for the high observed fractions of low luminosity AGN. However, the majority of SMBH mass is thought to be accrued in short bursts of rapid accretion ($\gtrsim 0.1 \mathrm{M}_\odot \mathrm{yr}^{-1}$; e.g. Croton et al., 2006); this requires the presence of large reservoirs of gas falling toward the BH and it is not at all clear that non-merger processes can consistently deliver such quantities of fuel.

To investigate the existence of some connection between galaxy interactions and nuclear activity two things are required: robust detections of AGN and of galaxy interactions. Detection of AGN is described in section 1.3.2. Galaxy interactions can be directly inferred by visual inspection of images or from structural parameters describing the light distribution in images. Merger rates can be constrained by searches for pairs of galaxies with small separations on the sky; from spectra the radial distance can be inferred to ensure the galaxies are truly close, rather than a chance alignment on the sky. While it is not guaranteed that close pairs of galaxies will merge, this can set an upper limit on the merger rate. Post-coalescence galaxy mergers can be visually identified from their highly non-equilibrium state (Toomre & Toomre, 1972) which gives rise to low surface brightness features such as tidal tails - streams of stars ejected from galaxies during the merger - or multiple cores, where the bright central bulges are amalgamating. Merger features fade with time after the merger and the higher the merger mass ratio the faster features tend to fade (Lotz et al., 2010a), so some mergers are likely to go unidentified in these searches. Structural parameters give an automatic indication of morphology without the subjectivity of human visual inspection. These can take the form of parametric indicators, such as the change in light intensity with radial distance from the galaxy centre as described by the Sersic profile (e.g. Peng et al., 2002), or non-parametric indicators such as the asymmetry (Abraham et al., 1996) of galaxies or the distribution of light over galaxy pixels. Humans arguably endure as the best suited to identifying the varied and low surface brightness features from galaxy mergers, but visual classification is highly subjective and prone to identifying features unrelated to mergers. Machine learning techniques of

galaxy classification (e.g. Hocking et al., 2017; Goulding et al., 2017) show promise. However, supervised machine learning algorithms must first be 'taught' what a merger is, with training samples typically drawn from visual classification catalogues, and so retain the same biases as humans, albeit at a much lower human cost. Unsupervised machine learning does not suffer from this, instead collating what the algorithm judges to be similar galaxies into groups, but it is not necessarily clear what the different groups classified by the machine correspond to. These methods are elaborated upon in chapters 2 and 3.

At low redshift mergers are more easily identified, but looking to higher redshifts faint merger features may be of too low surface brightness to reliably detect, even soon after the merger, and galaxies are often poorly resolved. Because of the difficulties inherent in detecting mergers at such high redshift many studies thus far have focussed on either low luminosity sources (e.g. Schawinski et al., 2011; Kocevski et al., 2012; Villforth et al., 2014) so the AGN does not drown out galaxy features, or low redshift AGN (e.g. Heckman et al., 1986; Sanders et al., 1988; Schawinski et al., 2010). It is not enough to show that AGN do or do not live in merging galaxies; the crucial comparison is to an inactive sample of galaxies matched in stellar mass, redshift and star forming properties to determine whether the merger rate is enhanced in active galaxies relative to "normal" galaxies.

The possibility of a connection between nuclear activity and galaxy mergers can be investigated observationally by comparing the morphologies of AGN hosts and comparable inactive galaxies. Heckman et al. (1986) find that very radio-loud AGN much more frequently inhabit mergers compared to less powerful radio AGN and normal ellipticals. Sanders et al. (1988) find that all 10 of their sample of ultraluminous IR galaxies are likely to be mergers, while 9 contain AGN. They propose that AGN are initially obscured during a merger, only later being revealed as optical quasars. Ellison et al. (2013) find four times as many AGN in post-merger galaxies in the SDSS at low redshift, with an increasing incidence of AGN activity and luminosity with decreasing separation between galaxies. Similarly, Satyapal et al. (2014) find a >10 times enhancement of AGN in mergers by selecting SDSS AGN using IR colours. Treister et al. (2012) found a strong correlation between AGN luminosity and galaxy mergers, increasing from 4% of low luminosity AGN to ∼90% of high luminosity AGN in mergers. Whilst this appears to support the notion that the highest luminosity AGN are triggered by mergers, a range of AGN selection criteria were compiled into the study and in some cases the true AGN luminosities are uncertain. Additionally, the lack of a control group makes interpretation of these results problematic. Kocevski et al. (2015) find a correlation between the fractions of AGN in mergers and the degree of obscuration of the AGN, a result echoed by studies of dust reddened quasars (Urrutia et al., 2008), consistent with mergers triggering BH growth but also hiding the AGN behind thick columns of gas and dust, biasing studies aimed at discerning the driver of nuclear activity. Conversely, Grogin et al. (2005), Gabor et al. (2009), Cisternas

et al. (2011), Schawinski et al. (2011), Kocevski et al. (2012), Villforth et al. (2014), Mechtley et al. (2016) and Villforth et al. (2017) compare AGN hosts to matched inactive galaxies and find no significant difference between them either in disk fractions, asymmetry distributions or visual classification. For AGN to be significantly triggered by mergers, explanation as to why the connection remains hidden in each of these studies must be offered. The variance within the literature clearly indicates the need for more research.

AGN flicker on short timescales while galaxies transform much more gradually, so one can make inferences about the processes driving nuclear activity from galaxy morphologies. Specifically, major mergers are expected to destroy disks and so if AGN are typically or commonly triggered by major mergers disks would be expected to be less prevalent in AGN hosts. Schawinski et al. (2011, 2012), and Bruce et al. (2015) find that most AGN live in disk galaxies with no significant differences to control groups, suggesting major mergers are not responsible for the nuclear activity, though without ruling out minor mergers as a contributing cause. Environmental factors broadly define the supply of cold gas from which BHs grow. Allevato et al. (2011) and Karouzos et al. (2014) find no evidence that AGN are preferentially triggered by interactions once environmental biases, such as the evolution of halo density, are considered. Sabater et al. (2013, 2015) find that once stellar mass and nuclear star formation rates are controlled for the effect of galaxy interactions and the density of galaxy environments has minimal effect on nuclear activity. Mullaney et al. (2012) find a universal relation between star formation rate and BH growth, while some studies find a delay between a nuclear starburst and accretion onto the SMBH (Wild et al., 2010), perhaps suggesting that mergers can trigger star formation and thus cause an incidental relationship between mergers and AGN. The extent to which galaxy mergers drive nuclear activity remains an open question. Developing techniques to unambiguously identify mergers, controlling for galactic factors such as star formation, understanding the role of AGN selection, and constraining the influence of time delays on possible nuclear activity triggers would be instrumental in answering the question: what role do galaxy mergers play in the growth of SMBHs?

In the following chapters AGN hosts are compared morphologically to matched controls across half the history of the Universe (chapter 2); simulations are used to analyse how the changing merger rate through cosmic time, galaxy mass ratios, time elapsed since mergers, and star formation influence observations and their interpretations (chapter 3); and the triggering processes of different classes of AGN, selected using different methods, are investigated (chapter 4).

2

Redshift Evolution & Luminosity Dependence of Merger Induced Nuclear Activity

As the gas distribution of the Universe evolves - falling into galaxies, becoming locked up in stars and black holes, ejected from galaxies and dispersed throughout clusters and groups - the mechanisms by which black holes are fed may be expected to change. When galaxies were gas rich and rapidly star forming one might not expect extreme dissipative processes such as mergers to be necessary to fuel AGN; in massive clusters (forming at late times) perhaps the virial temperature of gas is too high to allow much gas to fall towards SMBHs without galaxy mergers; conversely, galaxy mergers may be sufficiently rare at late times that secular processes dominate the BH feeding mechanisms. One aim of this study is to see if triggering mechanisms changed over cosmic time, both during and outside the peak of mergers/AGN activity. Another was to probe the triggering mechanisms of the highest luminosity AGN to see if there was some threshold accretion rate above which galaxy mergers became the dominant fuelling mechanism. A sample of 106 luminous X-ray selected type 1 AGN were compiled from the COSMOS survey. These AGN occupy a large redshift range ($0.5 < z < 2.2$) and two orders of magnitude in X-ray luminosity ($\sim 10^{43}$ - 10^{45} erg s^{-1}), probing the upper end of the AGN luminosity function. Looking to earlier epochs means looking toward the peak of AGN activity ($z \sim 2$-3), where the most rapid growth of SMBHs occurred (Ueda et al., 2003; Aird et al., 2010). This makes $z \sim 2$ a natural place to search for an AGN-merger connection,

since z∼2 also coincides with the predicted peak of galaxy mergers in hierarchical galaxy evolution (Kauffmann & Haehnelt, 2000; Volonteri et al., 2002; Conselice et al., 2003). AGN hosts were carefully mass and redshift matched to 486 control galaxies and a novel technique for identifying and quantifying merger features in galaxies was developed, subtracting GALFIT galaxy models and quantifying the residuals. Comparison to visual classification confirms this measure reliably picks out disturbance features in galaxies. No enhancement of merger features with increasing AGN luminosity is found with this metric, or by visual inspection. We analysed the redshift evolution of AGN associated with galaxy mergers and found no merger enhancement in lower redshift bins. Contrarily, in the highest redshift bin (z∼2) AGN are ∼4 times more likely to be in galaxies exhibiting evidence of morphological disturbance compared to control galaxies, at 99% confidence level (∼2.4σ) from visual inspection. Since only ∼15% of these AGN were found to be in morphologically disturbed galaxies, it is implied that major mergers at high redshift make a noticeable but subdominant contribution to AGN fuelling. At low redshifts other processes dominate and mergers become a less significant triggering mechanism.

The sample of AGN is introduced in section 2.1; the methodology in section 2.2; the results are presented in section 2.3 and discussed in section 2.4 before conclusions are presented in section 2.5. AB magnitudes are used throughout and, where needed, a ΛCDM cosmology with $h_0 = 0.7$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ is assumed. It is worth noting that with this cosmology the angular scale is approximately constant at ∼8.5 kpc/" between z=1-2.2, dropping off to ∼6.2 kpc/" at z=0.5. The fraction of galaxies in this work with z<1 is small enough to state here that working with pixel scales throughout is approximately equivalent to working with physical scales and will not affect the conclusions.

2.1 Sample and Data

The COSMOS survey contains multi-wavelength coverage over a region of sky ~2 deg^2 (Scoville et al., 2007). This includes X-ray data from XMM-Newton at a depth of ~50Ks (Hasinger et al., 2007; Cappelluti et al., 2009), allowing for identification of a sample of AGN using X-rays, and HST F814W data (near IR) for optical followup. X-rays are used for identification of AGN because it has been argued that they provide the cleanest AGN sample compared to other selection methods (e.g. optical, IR, radio), with very few non-AGN sources with $L_X >$ 10^{42} erg s^{-1} (Ranalli et al., 2003). Only Compton thick sources ($N_H > 10^{24}$ cm^{-2}; Risaliti et al., 1999) are likely to obscure X-ray emission from AGN.

HST imaging in the F814W band has a point source limiting AB magnitude of 27.2 at 10σ and a surface brightness limiting magnitude of 26.1 mags/arcsecond^2 at 3σ (calculated from image noise), otherwise stated as a 5σ depth in a 3" aperture of 25.3 (Capak et al., 2007). Point source depths drop off with $(1+z)^2$, while extended sources drop off with $(1+z)^4$ (an inherent difficulty in detecting extended, low surface brightness features in mergers at high redshift). This high sensitivity would seemingly make the HST ideal for identifying low surface brightness features such as tidal tails, hence was used in this investigation in the search for evidence of mergers. However, Bennert et al. (2008) combine 5 orbits of observations of quasar hosts to achieve the highest signal/noise ratio images possible. Comparison between these very deep images and single-orbit observations of the same galaxies demonstrates how low surface brightness features go undetected, and serves as a cautionary tale in interpreting the observations presented in this chapter. Some merger features may go undetected, though the comparison to a control group 2.2.2 renders this qualitatively inconsequential. HST images in COSMOS have been corrected for sky background and instrumental noise, and drizzled from an inherent resolution of 0.05"/pixel to a resolution of 0.03 "/pixel in post-processing (see Koekemoer et al., 2007; Massey et al., 2010, for details). The Wide Field Channel is comprised of two 2048 × 4096 pixel CCDs, giving it a ~200" field of view. From the 545 luminous X-ray detected type 1 AGN in the COSMOS field in Lusso et al. (2009), 106 were selected for analysis here over a redshift range z=0.5-2.2. A wide range of redshifts were selected in order to investigate if triggering mechanisms varied with time. We stipulated the need for pre-existing SMBH mass measurements, as required for development of the control sample (section 2.2.2).

The 106 AGN selected are spectroscopically confirmed to lie between z=0.5-2.2 (Lusso et al., 2009, and references therein) and have estimates of the SMBH masses from line widths of the Mg II lines (Peterson et al., 2004). Few AGN beyond z=2.2 have black hole mass measurements due to the difficulty of obtaining high quality spectra, hence we limit our sample to AGN with z<2.2. Figure 2.1 shows the complete Lusso et al. (2009) sample of AGN with
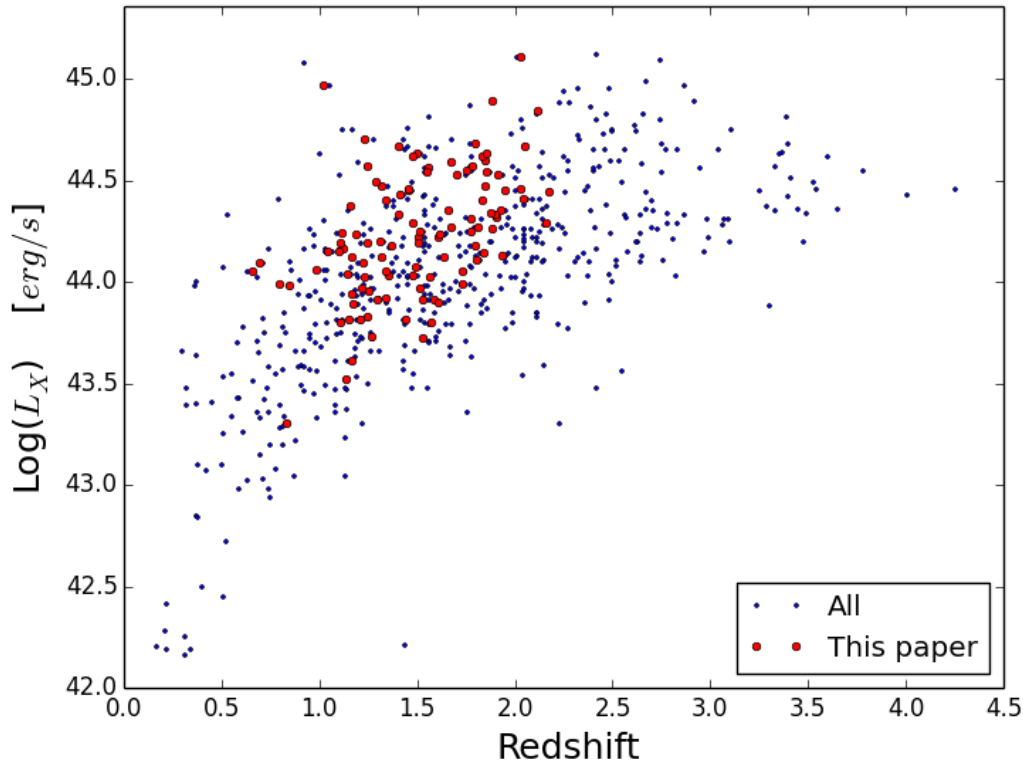
Figure 2.1: Larger red points represent the X-ray luminosities and redshifts of AGN in this sample; blue points show the total sample they were selected from. X-ray luminosities and redshifts are taken from Lusso et al. (2009).

those selected for this study in red. Control galaxy selection is discussed in section 2.2.2.

2.2 Methodology

For the sample we now aim to derive galaxy properties, such as the dominance of disks or bulges, by fitting to the light profiles; we build a sample of control galaxies with similar properties to AGN hosts; and subtract models of the light profiles, using the residual flux and visual classification to measure how disturbed the galaxies are. Use of GALFIT to construct 2D galaxy light profiles is described below in section 2.2.1; the selection and alteration of control galaxies to ensure they are equivalent to AGN hosts is described in section 2.2.2. Details of human visual classification are shown in section 2.2.3. The residual flux method is described in section 2.2.4 below, along with potential biases associated with it in section 2.2.5.

2.2.1 2D Decomposition

In this section we describe the fitting procedure with GALFIT. We derive galaxy properties such as Sersic indices and effective radii, magnitudes of galaxies and point sources. This provides a diagnostic of the evolutionary state of galaxies; for instance, a galaxy with Sersic index

n>4 is most likely an evolved elliptical, n~1 corresponds to a disk. Subtracting best fitting models from images reveals previously hidden structure, and measurement of the residual flux provides a quantitative measure of disturbance (sections 2.2.4). However, in order to accomplish any of this effectively a good PSF must first be constructed.

We constructed a PSF from field stars. 13 stars with a range of apparent magnitudes (m_{star} ~ 20 - 23) were initially selected for this by identifying stars in the SDSS field which overlapped with the COSMOS field; 5 were rejected due to possible blended sources, leaving 8 stars. These were combined and normalised to a total flux of 1 within a circle of radius 50 pixels, slightly larger than a typical galaxy radius (with half light radii typically <10 pixels). This avoids artificial structure in the residual images due to over-subtraction in an arbitrarily small region and resulting in a smooth subtraction of the point source, despite only small fractions (<1%) of point source light lying beyond a radius of 30 pixels. In the outer wings of the PSF the signal to noise ratio is low, around 1.5, with most flux concentrated in the centre. Application of this PSF to each of the stars used to construct it demonstrated reliable subtraction, leaving residuals consistent with noise, and so 8 was deemed a sufficient number of stars to accurately build the PSF. It was not clear that more stars would have improved the accuracy of the PSF, given its efficacy in tests.

Models of the light profiles of galaxies were made with the two dimensional (2D) fitting program GALFIT (Peng et al., 2002, 2009) and subtracted from the images. In the models galaxy surface brightness varies smoothly and was constrained to be symmetric about the centre. GALFIT operates through a χ^2 minimisation algorithm to find the best fit parameters for point sources, Sersic profiles, spiral arms and more complex components as necessary (Peng et al., 2002, 2009). Models in this study were kept as simple as was required to accurately recreate light profiles and features such as spiral arms or bars were not fit; models consisted only of point sources and Sersic profiles.

It was already known that these galaxies contained a point source (the AGN) and so the initial model in each case was a point source only. Figure 2.2 demonstrates how removing the point source can reveal a previously hidden underlying galaxy by adding a point source to a normal galaxy (creating an image comparable to an AGN host) and then re-subtracting the best GALFIT model of the point source. A Sersic profile was introduced to the model where necessary, with models kept as simple as required to accurately recreate the light profile. Nearby galaxies in the image, if showing no sign of association to the galaxy of interest, were masked out to improve fits. Multiple models were made for each galaxy and discriminated between based on the value of the reduced χ^2 and the physical plausibility of parameters. Specifically, since especially large Sersic indices or small half light radii are degenerate with point sources, and small Sersic Indices or large half light radii tend toward fitting noise, it was stipulated

that only Sersic models with $0.6 < R_e$ [pixels] < 70 (~ 0.15-18kpc) and $0.3 < n < 10$ would be accepted, and only then if the reduced χ^2 was improved relative to a point source only model. If model parameters failed to converge the fitting was run twice more with the Sersic index held fixed at n=4 (classical bulge) and n=1 (exponential disk). Fixing values reduces the number of parameters to be fit in GALFIT and so can aid in the construction of models. If convergence was still not achieved, or the reduced χ^2 was not improved, the point source only model was selected. The high luminosity of the AGN in this sample makes fitting in GALFIT especially difficult, and parameters often had to be fixed for fits to work. When fixed, there may be a bias toward fitting exponential profiles (n=1). Since compact bulges are degenerate with point sources (e.g. Bruce et al., 2015), exponential profiles may more smoothly capture the outer regions of a galaxy while the point sources recreate the central regions. Getting the combination of the correct luminosity and radial extent of a classical bulge fit in combination with a bright point source is extremely challenging. We note that fixing parameters introduces some subjectivity, and this makes direct comparison to the fits of other lower luminosity, lower redshift samples problematic (e.g. Gabor et al., 2009). For $\sim 50\%$ of galaxies a point source alone was sufficient to model the light profile, since the underlying galaxy was too faint to be detected/modelled. The other $\sim 50\%$ were fit by a point source + Sersic profile.

2.2.2 Control Galaxies

Selecting a group of control galaxies similar to the AGN hosts was imperative since the enhancement of merger rates in AGN hosts relative to inactive galaxies was the quantity of interest, rather than the absolute value of merger rates. Similar galaxies may have comparable merger rates but no AGN, which would suggest AGN are triggered by something other than mergers, and any presence of mergers is coincidental. AGN accretion rates are expected to be a function of stellar mass and redshift (Netzer & Trakhtenbrot, 2007), as are major merger rates (Cedric & Shaun, 1994; Fakhouri et al., 2010), so galaxies were matched in stellar mass and redshift. By comparing to a carefully matched control group the relative levels of disturbance for different AGN properties (e.g. luminosity) could be compared.

Control galaxies were selected from the COSMOS Photometry Catalogue January 2006 (Mobasher et al., 2007). Masses and redshifts in this catalogue were determined from photometry (M/L$_V$ $\propto$ B-V; mass to light ratios are proportional to colour, so masses can be estimated from measurements of V band luminosity and B-V colours). There is remarkable agreement between spectroscopic and photometric redshifts, with an rms scatter of $\Delta(z) = (z_{phot} - z_{spec})/(1+z_{spec}) = 0.031$ between them (see Mobasher et al., 2007, for details).

AGN host galaxy masses were estimated from the SMBH mass using the M-M$_{bulge}$ relationship from McConnell & Ma (2013) (equation 2.1). This was done because AGN hosts lacked

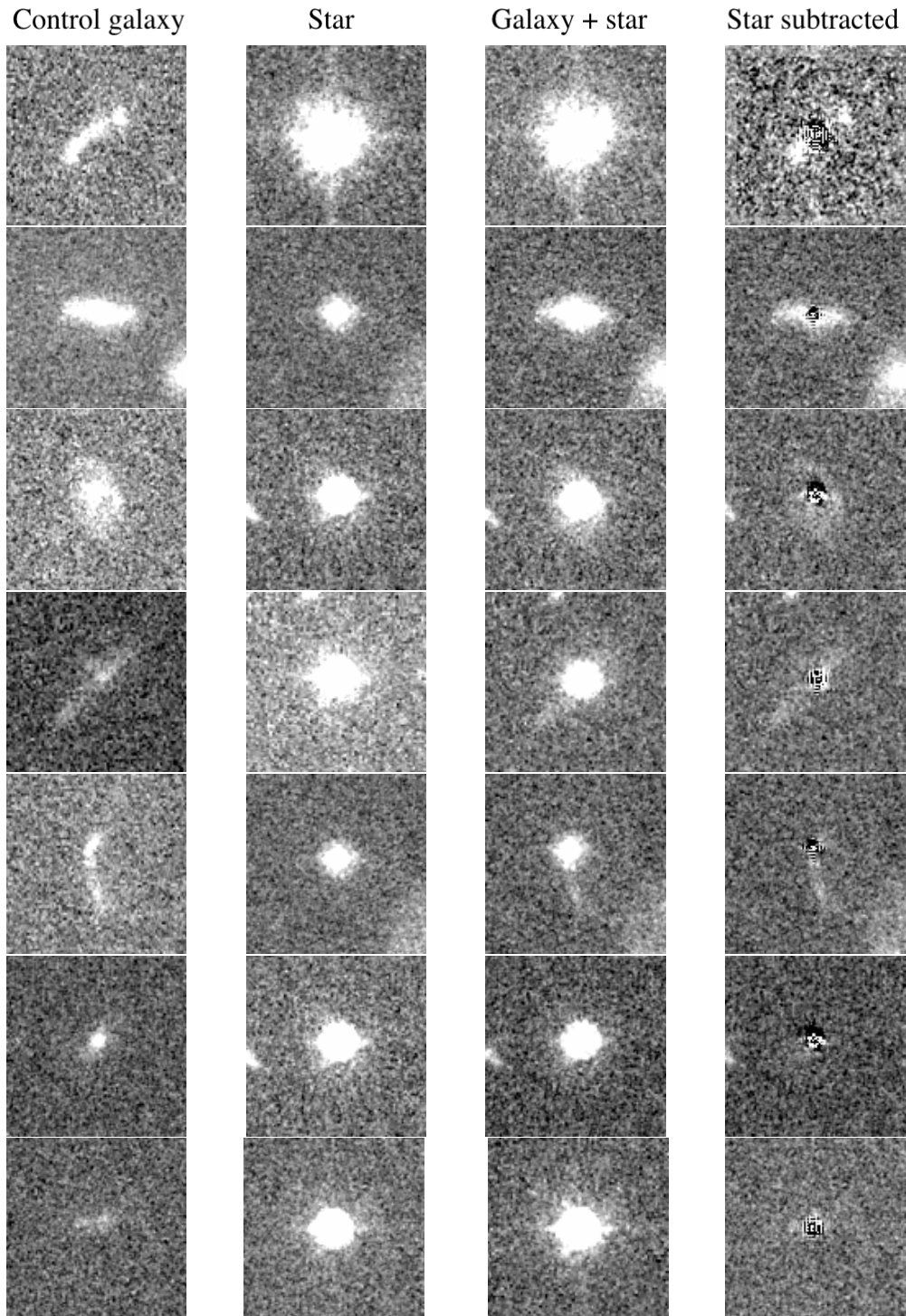| Control galaxy	Star	Galaxy + star	Star subtracted

Figure 2.2: ~3" cutouts of control galaxy (left) plus star (centre left) equals final image (centre right) for seven example galaxies. Far right panels show the stars being re-subtracted from these images, using GALFIT, to recover the host galaxy profile. From top to bottom: ID=2381537, z=1.56; ID=3024848, z=1.34; ID=2747981, z=1.24; ID=2479182, z=1.56; ID=688479, z=1.55; ID=3038344, z=2.06; ID=427649, z=1.53.

stellar mass estimates and obtaining stellar masses is especially difficult when the light is so dominated by AGN emission. Equation 2.1 provides a consistent mass estimate for all AGN hosts.

$$\log\left(\frac{M_{\rm bh}}{M_\odot}\right) = (8.46 \pm 0.08) + (1.05 \pm 0.11)\log\left(\frac{M_{\rm bulge}}{10^{11}M_\odot}\right) \qquad (2.1)$$

A Gaussian distribution around each AGN host galaxy mass was created, with width defined by the combined error (combined using standard formulae, under the assumption that the errors are independent) in equation 2.1, and five mass values were picked at random from this Gaussian. These values were matched as closely as possible to masses and redshifts in the COSMOS Photometry Catalogue leading to 5 matches per AGN (4 in one case due to a lack of 5 matching galaxies) and an initial control sample of 529 galaxies. Figure 2.3 demonstrates the outcomes of the matching procedure. To avoid AGN contamination in the control group the 43 that were X-ray detected (a possible signature of nuclear activity) were removed from further analysis, leaving 486 control galaxies remaining in the sample with a minimum of 3 controls per AGN, the majority of AGN still matched to 5 control galaxies. The X-ray detected galaxies were normally distributed across the full parameter space, with no trend with mass or redshift, so no bias was introduced due to some AGN hosts matching to more control galaxies than others. The maximum separation in redshift was $\Delta z = 0.3$ for rarer high mass galaxies, with the majority matched within $\Delta z = 0.1$. There is a larger range in mass due to the scatter in equation 2.1, with the median mass range about the assumed host mass $\Delta\log(M/M_\odot) \sim 0.4$ (with all $\Delta\log(M/M_\odot) < 0.7$), but since the true mass of the AGN hosts is unknown, merely estimated from equation 2.1, and there is large uncertainty in the stellar masses from the B-V colour, the intention was to compare to a range of control galaxies, each of which may represent the true form of the AGN host. As long as the average masses are similar, and the M-M$_{bulge}$ relation approximately holds to z∼2 (Robertson et al., 2006a; Lauer et al., 2007; Shankar et al., 2009; Shen et al., 2015), the control galaxies should, on average, represent the AGN hosts. We appreciate this is a mildly controversial assumption; however, current evidence suggests relations between the mass of the SMBH and the galaxy it inhabits must be slowly evolving, if at all (Robertson et al., 2006a; Shen et al., 2015), and so uncertainties in control galaxy mass will dominate. Projected histograms in figure 2.3 show the distribution of redshifts and (in the case of AGN hosts, assumed) masses for AGN and matched control galaxies. As expected, the match in redshift is extremely tight while there are longer tails in the control galaxy masses due to selecting masses randomly from a gaussian distribution around the expected value of the AGN host mass. There is a slight bias in the matching because lower mass galaxies are more numerous, so in matching to the first available galaxy within some bounds (e.g. $\pm\log(0.1)$) it will more likely be of slightly lower mass than the AGN host.

The effect is small however, with mean masses $\log(M_{control}) = 10.9M_\odot$ and $\log(M_{AGNhost}) = 11.0M_\odot$.

The F814W band samples the near UV emission at z~2, and optical emission (λ~5000) at z~0.5, so will predominantly image star forming regions. Star formation is clumpy at high z (Cowie et al., 1995; van den Bergh et al., 1996) so this may produce patchier light profiles at high redshift, contributing to the residual flux without any morphological disturbance. The dominance of AGN light in the spectra makes controlling for star forming properties extremely difficult. AGN hosts have a range of star formation rates (e.g. Mullaney et al., 2012). Rest frame U-V and U-B colours reveal the same is likely true for the control galaxies here, with colours representative of the underlying population, and so it is likely that the star forming properties of the control galaxies match reasonably well to the AGN hosts. It is noted that star formation rates are not directly matched between AGN hosts and control galaxies, but that the control galaxies are representative of galaxies of similar mass. It will be shown in chapter 3 that matching star formation rates between AGN hosts and controls is crucial for fair morphological comparison, but such matching was not attempted here due to the inherent difficulties.

Many images of AGN + hosts are dominated by light from the point source, making distinguishing features by eye difficult. To make control galaxies actually resemble AGN hosts, 26 stars were identified from the SDSS, selected in COSMOS, matched as closely as possible in apparent magnitude to AGN and added to the associated control galaxies as AGN-like point sources. Figure 2.2 shows control galaxies before and after adding stars; the structure of the underlying galaxy is largely obscured by even a relatively dim point source, so it was imperative to add realistic point sources to the control galaxies before any further analysis. Figure 2.2 also demonstrates the efficacy of removing the point source in GALFIT, successfully recovering much of the underlying galaxy structure, albeit with less certainty in the central regions. There were no adequately faint stars in the field for the faintest AGN. In this case, the faintest star available was used as a surrogate. Potential bias introduced by this is discussed in section 2.2.5.

Some (~40, 8%) control galaxies were undetected at F814W. Their reported magnitudes in the catalogue were in the dim tail of the whole distribution of galaxies used here, ~25th magnitude, so if slightly extended would be hard to pick out visually. Since stellar masses were required for selection, objects were selected from the COSMOS Photometry Catalogue 2006 Mobasher et al. (2007) instead of the ACS catalogue, and resultantly there are some cases where the galaxy is missing in ACS. We wished to avoid biasing the results, since it is likely that in some AGN images the host is also undetected in F814W with only the AGN itself seen, so a point source was added to empty sky in these cases. This allows testing of the reliability
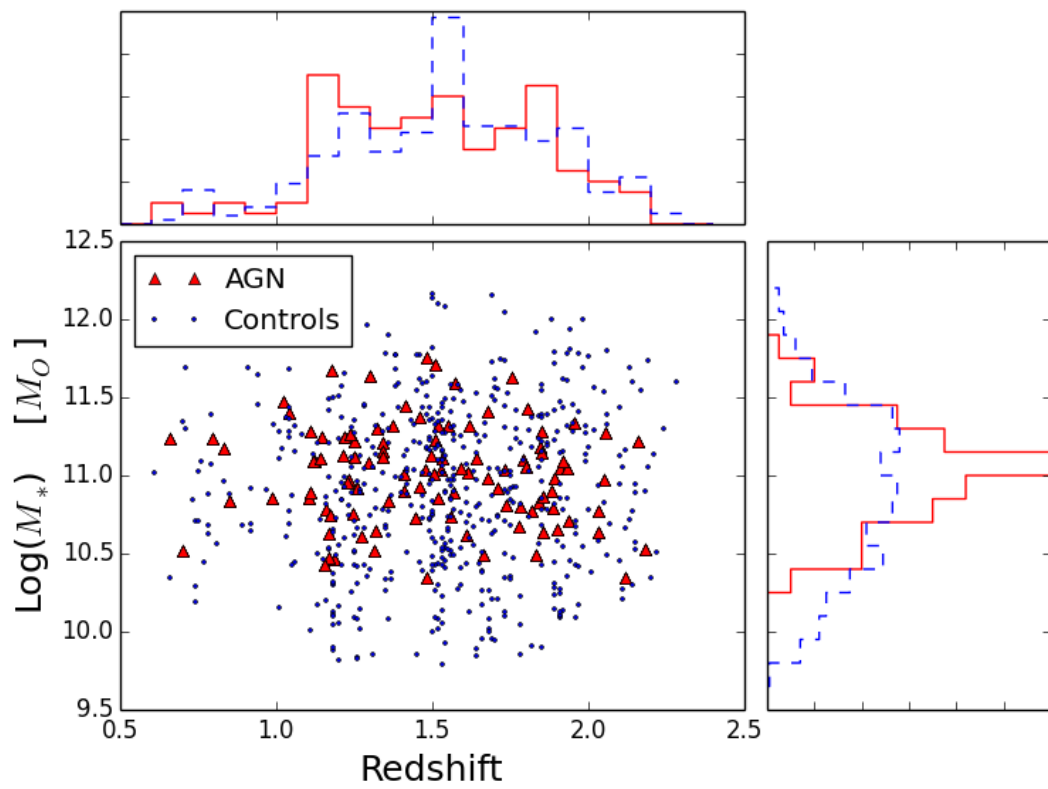
Figure 2.3: Estimated stellar masses of AGN hosts (red triangles) and the matched control galaxies (blue circles). Projected histograms show the distribution of AGN and control galaxies in stellar mass and redshift; blue dashed histograms correspond to control galaxies, red solid histograms to AGN hosts.

of host galaxy reconstruction in GALFIT, since if more than a point source is recovered where only a point source was added this provides an estimate of the number of false positive galaxy detections. $\sim$1/4 of these had false galaxy detections with a mean magnitude m_{Gal}=24.4 in F814W (far fainter than the mean point source magnitude for these galaxies, m_{PS}=21.6), and radii of 2 pixels (just over the threshold radius for the model to be accepted). This implies GALFIT is ascribing some point source light to a non-existent galaxy, hence the dim and compact galaxy. Recall that $\sim$50% of all images, AGN and controls alike, were best fit by point sources alone, implying galaxies were often dim compared to the central point source.

2.2.3 Visual Classification

Three experts classified the galaxies into five main groups both for comparison to the residual flux method and for independent science analysis. The five main categories of classification (see figure 2.4 for example classifications) are:

i) PSF dominated - for cases where the image is dominated by poor PSF subtraction: for instance, large symmetric diffraction spikes.

ii) Undisturbed without residuals and iii) undisturbed with residuals - classes intended to discriminate between galaxies that the residual flux method might pick out as merger candidates, but humans reveal to have highly symmetric residuals due, for example, to marginal galaxy detections.

iv) Disturbed galaxies - where there is some asymmetry or otherwise disturbance in the light profile, but the source of the disturbance may be unclear, for example with asymmetric regions of over- or under-subtraction but no clear multiple cores. This may correspond to the late stages of a merger, more minor mergers, or possibly in some cases to clumpy disks.

v) Clear mergers - a category for galaxies deemed to be clear mergers with multiple cores, tidal tails, or otherwise excessive asymmetry.

Classifiers were asked to put each of the 592 model-subtracted galaxies into one of the five classes, with additional optional columns for features such as spiral arms or foreground objects. Classifiers were in broad agreement with one another: when one classifier categorises a galaxy as a merger $\sim$85% of the time the other two classifiers have classified it as disturbed or a merger.

How does the detectability of merger features vary with redshift? Kaviraj et al. (2013) analyse merger remnants from a cosmological hydrodynamical simulation (Peirani et al., 2010)

and find that at redshifts of ∼1.25, only mergers with mass ratios of 1:5 are detectable with HST imaging at a depth of 26 mag arcsec^{-2}. At z∼3 this detectability reduces to real train wrecks, with mass ratios of 1:2. This suggests that those identified as mergers here are likely to be major mergers, and at higher redshifts the fraction of galaxies identified as mergers will drop off (unless balanced by an intrinsic increase in the fraction undergoing major mergers). Constraints on the detectability of merger features through cosmic time from cosmological simulations are established in chapter 3. For this study it suffices for the control galaxies and AGN hosts to be comparable across the redshift range, giving rise to similar biases, allowing like-for-like comparisons to be made.

2.2.4 Residual Flux

The primary question under investigation was whether the brightest AGN were in merging galaxies and how this changes with redshift. The residual method utilised here is conceptually simple. Merger signatures were quantified by subtracting GALFIT models from images and measuring the remaining flux as a proxy for morphological perturbation. Models of the light profiles of galaxies were made with the two dimensional (2D) fitting program GALFIT (Peng et al., 2002, 2009) and subtracted from the images. In the models galaxy surface brightness varies smoothly and was constrained to be symmetric about the centre; this means that significant flux remaining in the residuals of an image will be due to the galaxy itself having disturbed, clumpy or asymmetric light profiles, which may be indicative of morphological disruption. Subtracting models from images leaves a profile consistent with noise if the fits are good; if the light profiles are discontinuous or asymmetric however, such as expected in merging systems, GALFIT will be unable to make good fits, leading to areas of over- or under-subtraction in the residuals. Summing the residual flux in the under-subtracted regions and the modulus of the residuals in over-subtracted regions, for values greater than the error on each pixel, gives a direct numerical measure of disturbance. This method was developed in part because at z=2 with $L_{X,AGN} \sim 10^{45}$ erg s^{-1} visual classification becomes difficult in many cases, is highly subjective, and other automated methods were unlikely to be sensitive to the features searched for. In addition, this may be useful for future studies involving very large samples for quickly identifying those galaxies most likely to be mergers.

Methods in the literature for measuring the symmetry or internal disorder of a galaxy include the rotational asymmetry (Shade, 1995; Rix & Zaritsky, 1995; Abraham et al., 1996; Villforth et al., 2014); the clumpiness parameter (Isserstedt & Schindler, 1986; Conselice, 2003); the Gini index (Lotz et al., 2004) or the shape asymmetry (Pawlik et al., 2016). The problem with applying these measures to the study of AGN lies in the low surface brightness of the features being searched for compared to the high intensity emission from the AGN. For example, the rotational asymmetry parameter is dominated by the bright central regions
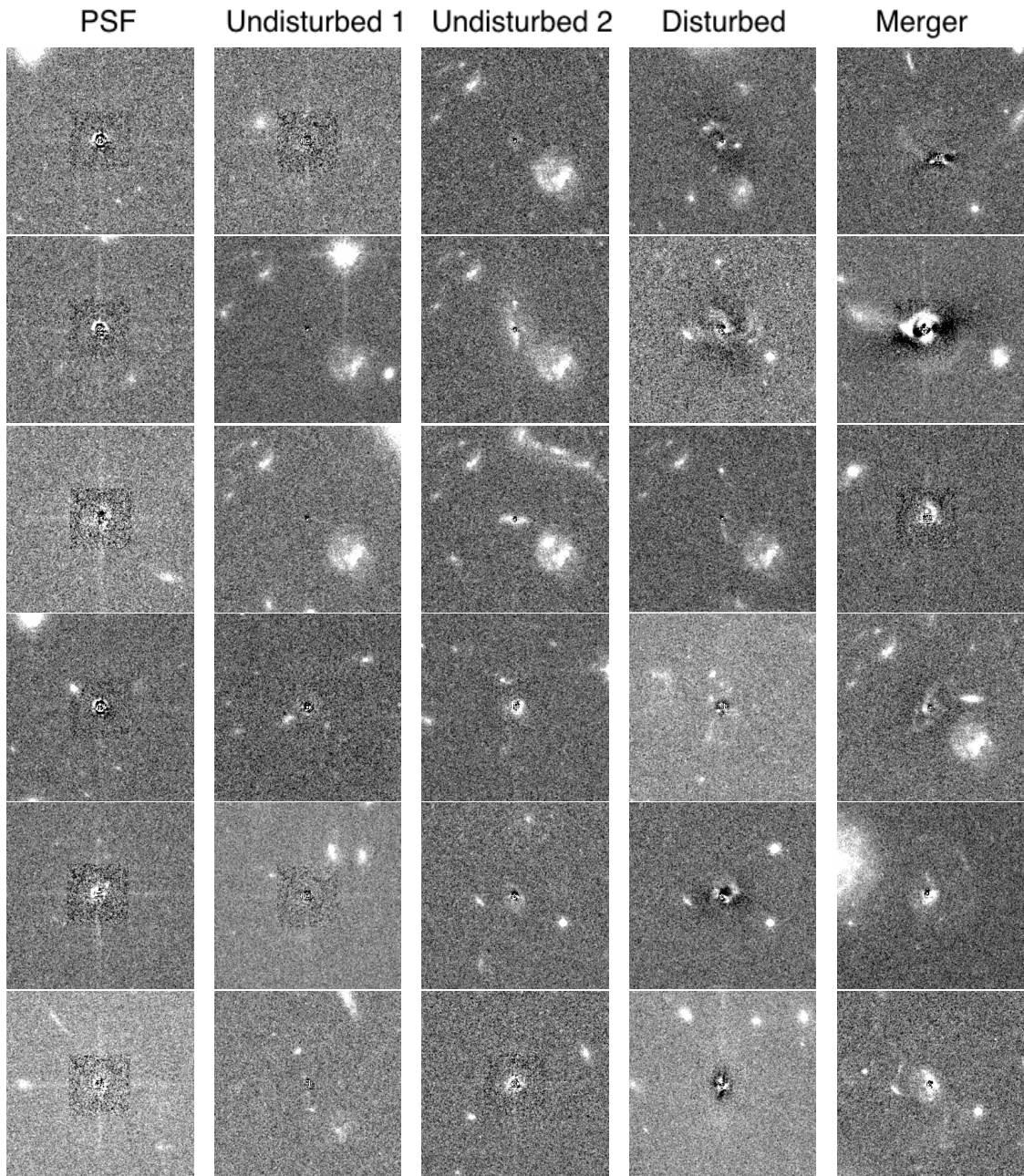
Figure 2.4: Example classifications of different galaxies. Columns from left to right: PSF dominated; undisturbed without residuals; undisturbed with residuals; disturbed; merger. Although subjective, the merger is distinguished in particular from its extended tidal features.

and not by extended tidal tails; the clumpiness parameter falls short if star forming regions are differently resolved at different redshifts; the shape asymmetry works exceptionally well at detecting low surface brightness features but is poorly suited to detecting multiple cores. The residual flux method shares in these challenges, but was born of an attempt to create a measure which simultaneously measures the internal disorder; the extended emission and the presence of multiple nuclei; and features which are otherwise outshone by the AGN, giving a continuous scale of disorder. Residual flux is not a measure that is directly comparable to other measures of disturbance such as those discussed above. In the future, the residual flux method may be utilised to pick out likely merger candidates for further inspection and analysis: since mergers are rare (Bundy et al., 2009) gaining large samples of mergers can be challenging, particularly at high redshift where galaxies are most poorly resolved.

Figure 2.5 shows the residual flux across the centre of a disturbed galaxy, exposing clearly a merger feature and demonstrating that merger features may very plausibly show up as residuals in fits: the more disturbed the galaxy the more residuals remain. With this motivation, the residual flux in annuli about the point source was calculated. The central circle with radius r_{mask}=15 pixels was masked out because this region was PSF dominated, with much of the residual flux here being attributed to under-sampling of the PSF (see figure 2.5). Through fitting to stars it was found that beyond 15 pixels this effect was largely mitigated, with the mean residual flux for stars with faint apparent magnitudes dropping close to zero. Annuli were at 30, 40 and 50 pixel radii in order to capture different regions of the galaxy (R $\sim$ 7.5, 10 and 13 kpc radii at z=2, and is relatively constant over z=0.5-2.2). It was found that the inner 30 pixels (minus the central 15) were sufficient to qualitatively compare disturbance between galaxies, with broad results remaining the same for all annulus radii, independent of redshift.

High residual flux remaining after the model subtraction is highly suggestive of some non-equilibrium state (a galaxy in equilibrium will relax into a symmetric state; Robertson et al., 2006b): e.g. when galaxy morphologies are perturbed by major mergers. Residual flux is not just an on-off switch signifying mergers however, but a sliding scale of disturbance. This may help to distinguish between major and minor mergers, or recent and old mergers. The more morphological disturbance the less smooth the light profile and the more residual flux remaining post model-subtraction. Merger features are expected to fade away on relatively short timescales (totally faded within $\sim$1Gyr, typically less Tinsley, 1978; Kennicutt et al., 1987; Lotz et al., 2010a,b; Ellison et al., 2013) and so moderate values of residual flux may indicate some relaxing state after a merger, or some other form of disturbance such as minor mergers, while the highest values of residual flux should belong to major mergers. Thus, this has the power to differentiate somewhat between triggering mechanisms, or possibly to constrain any possible time delay between mergers and AGN activity. Occasionally host galaxies are only

marginally resolved. This can result in an incomplete GALFIT model and an overestimation of the residual flux as due to disturbance. In these cases the residual flux is generally low, as a bright underlying galaxy will be fit better by GALFIT, but visual classification by 3 expert classifiers was used to identify marginally detected hosts, so that they could be excluded from analysis to determine the significance of this factor (see section 2.2.3).

One motivation for developing the residual flux method was to create a more continuous measure of disturbance than is easily achieved by visual inspection (although see Mechtley et al. (2016) for a creative continuous method of visual classification of disturbance), so it is not expected to yield identical results to visual classification (section 2.2.3). Correlation between the methods, however, is required for credibility in the residual flux method. Figure 2.6 shows the residual flux for the different morphological categories. There is no particular difference between classifications of undisturbed galaxies with or without residuals, suggesting the residual flux method is similarly good at picking out true disturbance as humans. The residual flux is not heavily influenced by non-axisymmetric features such as spiral arms: omitting objects flagged by visual classifiers as having such features did not influence results. The mean residual flux rises for disturbed and again for merger classes: on average this method is picking out similar features to humans. Although there is considerable scatter within any classification, a galaxy with residual flux > 2 has a >50% chance of being classified by humans as a merger or disturbed, despite these constituting <20% of the sample. This highlights the power of this method for identifying large samples of mergers in future studies.

2.2.5 Accounting for Potential Bias

To ensure that results were not biased significant effort was made to control for factors that may influence the measurement of the residual flux: poor point source subtraction; brighter galaxies having higher residual flux just because they are brighter to start with; and relations between residual flux and the signal to noise ratio (SNR). The most important consideration is for any biases to be equal between AGN hosts and controls, since enhancement of one over the other would be the significant discovery.

First, to control for poor point source subtraction, point sources were fit in GALFIT to the 26 stars used as artificial point sources in the control galaxy group (section 2.2.2) and subtracted from the images. Figure 2.7 shows their residual flux as a function of apparent magnitude. Brighter point sources have greater residual fluxes, unrelated to the galaxy morphology. The tight correlation between residual flux and magnitude can, however, be used to correct for point source light contamination. The magnitudes of AGN and artificial point sources were measured in GALFIT and the associated residual flux from the best fit line in figure 2.7 was subtracted from the residual flux value obtained for the AGN+host or artificial
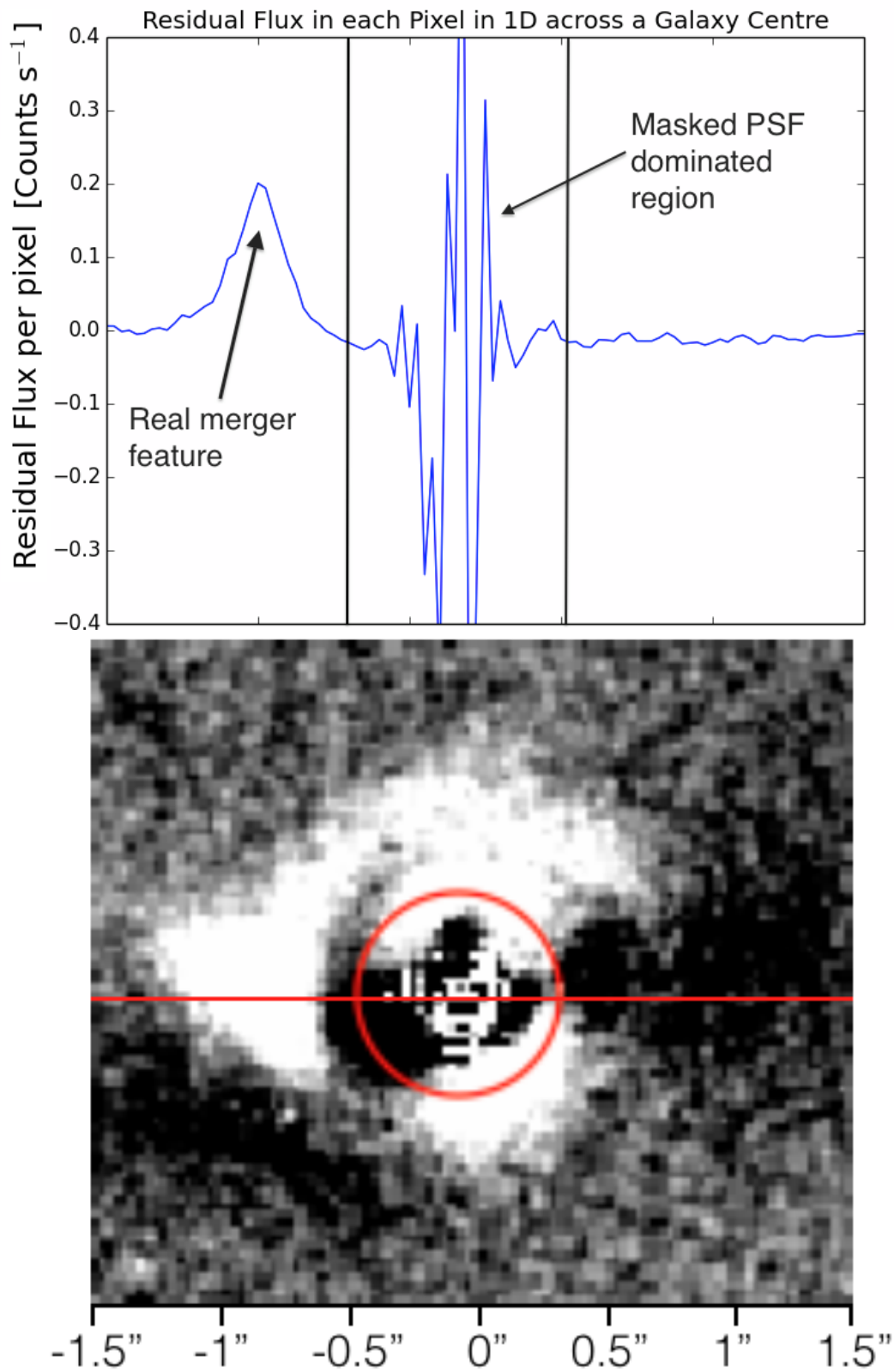
Figure 2.5: Upper panel: residual flux in a one dimensional horizontal line across the centre of the disturbed AGN host in the lower panel. Vertical lines denote the region masked out due to diffraction spikes in PSF residuals, seen as a mosaic-like region in the centre below, while a merger feature is highlighted at around -1". The image below is aligned with the upper panel and the angular distribution on the sky is displayed in arcseconds, with the horizontal line and masked out region, with $r_{mask} = 15$ pixels, denoted by the red line and circle. X-ray ID=5133, z=0.66
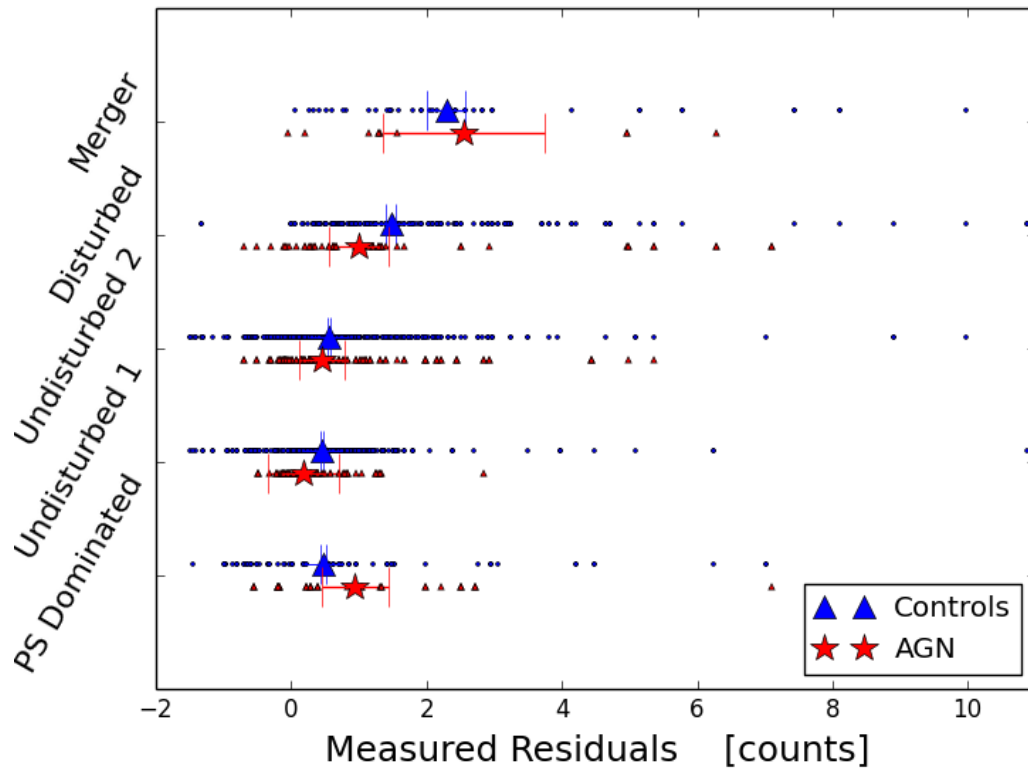
Figure 2.6: Human visual classification of morphology vs residual flux. Classifications are point source dominated galaxies (PS dominated); undisturbed without residuals (undisturbed 1); undisturbed with residuals (undisturbed 2); disturbed galaxies and clear mergers. Small red and blue points are the individual galaxy classifications while the big red stars and blue triangles are median values of residual flux for each classification (AGN and Controls respectively). Errors are found by bootstrap resampling. 1/3 of the data is randomly sampled (and replaced) and the mean found; this is repeated 1000 times and the standard deviation on those mean values provides the final error. Some values of residual flux are negative since the point source contribution is estimated and subtracted using the trend-line in figure 2.7, while the true residual flux due to the point source may scatter beneath the trend-line (see figure 2.7 and text for details). This means the point source is over-compensated for and the residual flux value becomes negative (though note this small error is far less than if the point source contribution is not corrected for).
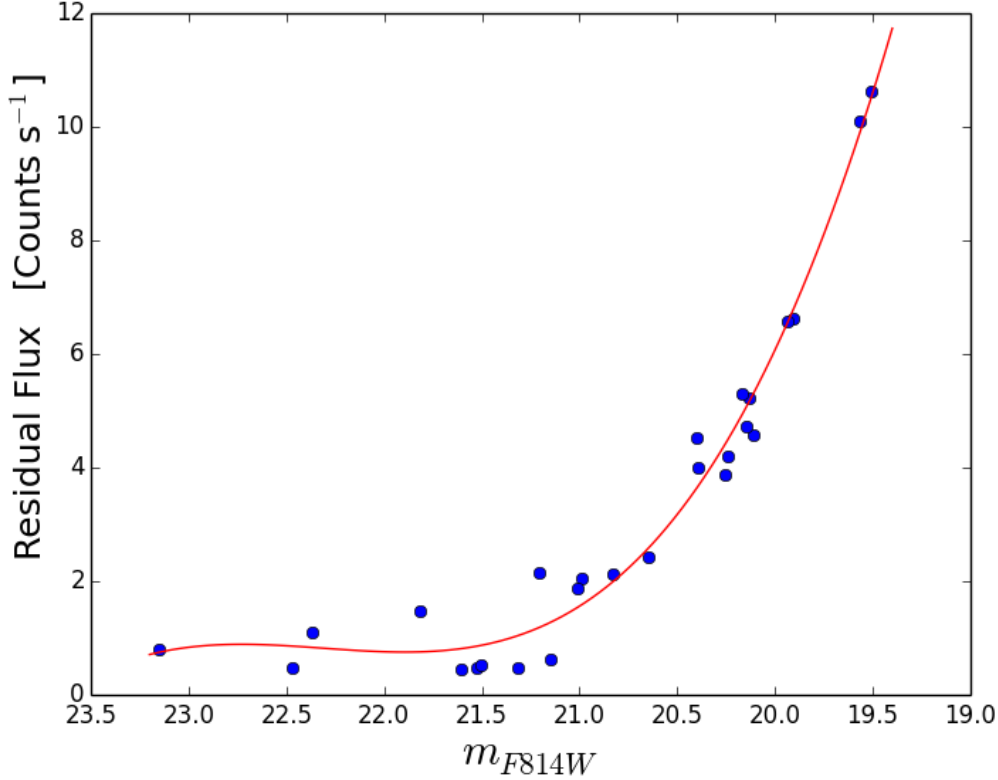
Figure 2.7: Residual flux of stars in annuli from 15-30 pixels as a function of magnitude. The tight correlation can be used to control for poor point source subtraction.

point source+galaxy system, meaning any remaining residual flux should be due to the galaxy itself and not poor point source subtraction. Since figure 2.7 shows how the trend of residual flux vs magnitude flattens at magnitudes fainter than ∼21, and there is no reason to expect this trend to change at even fainter magnitudes, using the faintest star as a surrogate for even fainter AGN (section 2.2.2) should not bias the values of residual flux. We chose to correct for point source light contamination by subtracting residual flux, since there was such a tight relationship between magnitude and residual flux, instead of scaling down point source flux to exactly match the AGN, since this could have introduced new and unknown sources of bias. The brightest stars were comparable to the brightest AGN. Some variation in the PSF over the field may be expected, however no trend between residual flux and distance from the centre of the chip was found, and no obvious asymmetries were detected in the majority of the stars from visual inspection. This suggests the overwhelmingly dominant factor influencing the residual flux from the point source is its magnitude.

Second, no correlation between control galaxy stellar masses and residual flux was found. Mass was used as a proxy for magnitude as a check that is independent of the GALFIT models, since model magnitudes are clearly degenerate with the residual flux, and F814W magnitudes are contaminated with AGN light for the AGN hosts. Since we found no correlation between

stellar mass and residual flux this strongly suggests the brightness of the host is not a significant factor in producing high residual fluxes.

Finally, since residual flux of a pixel is counted only if its value is greater than the error on that pixel, some correlation between residual flux and the SNR may be expected. Indeed, a correlation between SNR and residual flux was found. AGN and control galaxies are found to be equivalent however, so this will not cause any kind of enhancement in one group over the other. Because in the control group two images have been added together (galaxy and star) the sky flux is correspondingly higher. For this reason 1σ is the correct choice above which to sum flux, since errors take account of the addition of images by summing the image errors in quadrature, before subtracting the point source contribution defined by figure 2.7.

Figure 2.8 shows all galaxies with residual flux > 5 counts, most demonstrating high degrees of disturbance; five galaxies were removed due to non-physical features in the images such as the light from overexposed nearby stars bleeding over the galaxy image. The highest residual flux galaxies appear to have the highest degrees of disturbance: this technique picks these features out and identifies strong merger candidates, in addition to the increased statistical likelihood of a galaxy being disturbed if residual flux is high.

2.3 Results

We will now present the results of the residual flux method alongside visual classifications to determine any difference between morphological disturbance of AGN hosts and control galaxies as a function of: 1) the fraction of control galaxies and AGN in different categories or residual flux bins; 2) the point source or AGN luminosities; 3) redshift. For each question we will first present evidence from the residual flux method, then from visual classification.

The most obvious question we can ask regards statement 1: are AGN more likely to live in disturbed systems? Figure 2.9 shows the comparison between AGN hosts and control galaxies in residual flux. By this metric AGN are no more likely to be found in disturbed galaxies than control galaxies. Similarly, figure 2.10 shows no statistically significant difference between controls and AGN hosts by visual classification, with both merger and disturbed fractions equivalent for both samples. Together these results show that there are not higher numbers of AGN in merging galaxies in this sample.

Through fitting to the light profiles in GALFIT, it was found that the majority of galaxies in this sample are best fit by disks, rather than spheroids, with a mean Sersic index of $\sim$1.15 for both AGN hosts and control galaxies. This is in agreement with some studies (e.g. Cisternas et al., 2011) and in contrast with others (e.g. Dunlop et al., 2003, though these are low redshift galaxies). It might be expected for mergers to destroy disks, or otherwise make
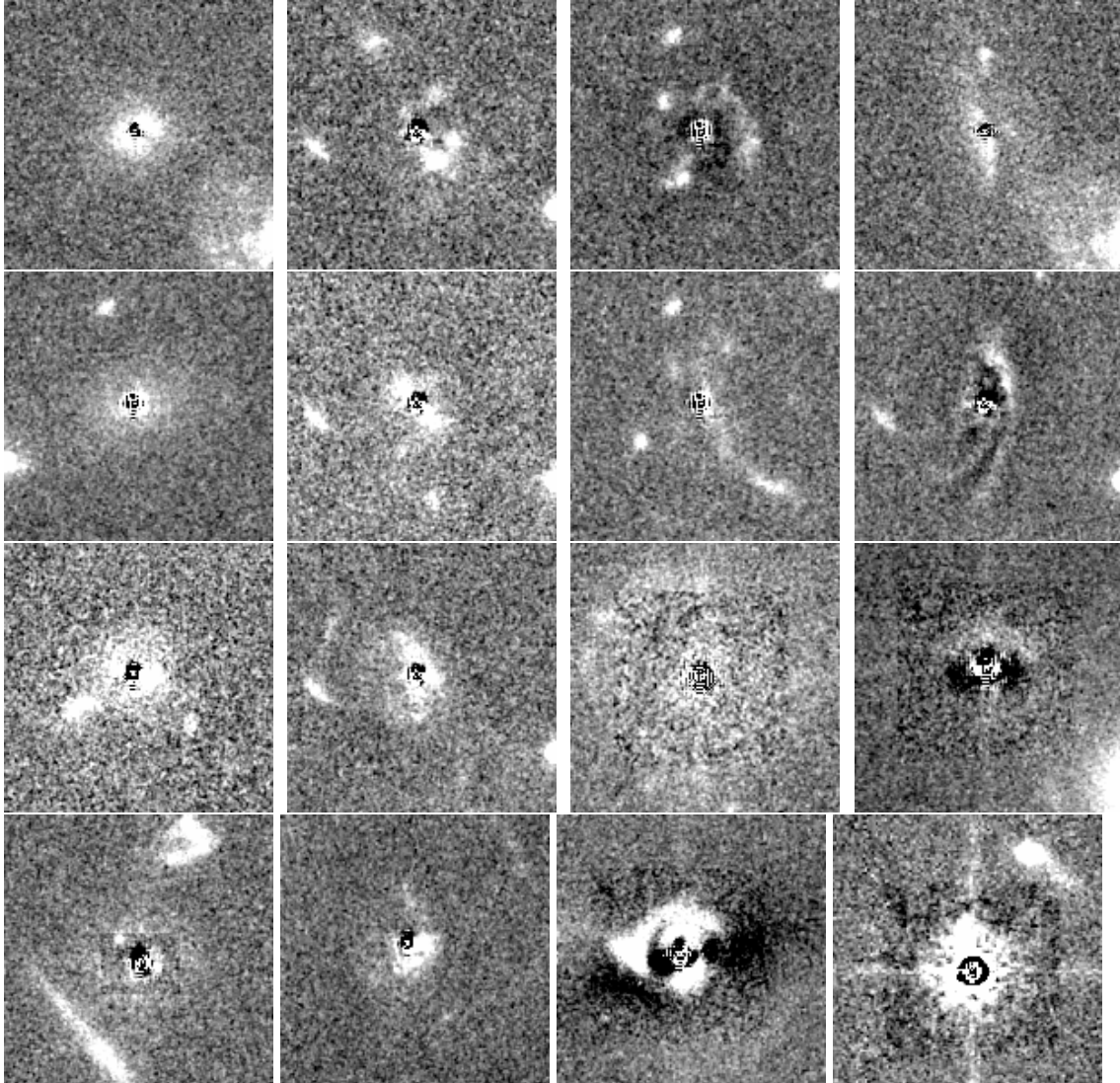
Figure 2.8: All galaxies with residual flux > 5 counts; high degrees of disturbance are seen in almost all cases. The first 11 galaxies (counting left to right, top to bottom and starting in the top left) are controls and the last 5 are AGN hosts. 5 Galaxies (1 AGN and 4 control) were removed due to image defects; see text for details.
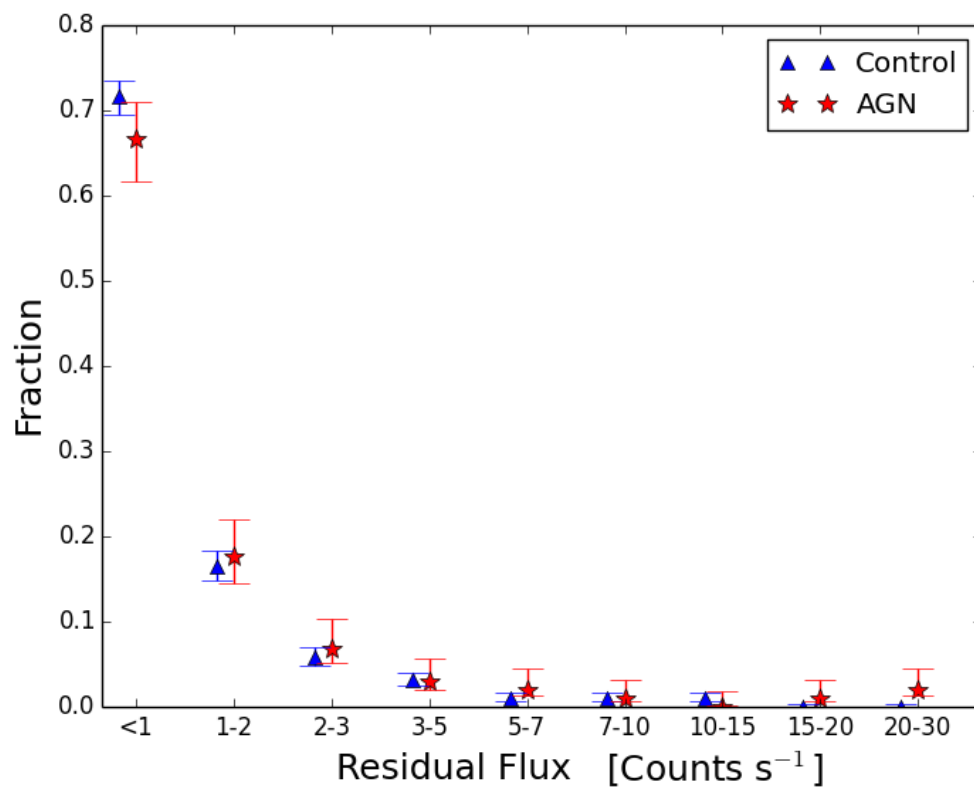
Figure 2.9: Fraction of AGN and control galaxies in bins of residual flux. Blue triangles are control galaxies and red stars represent AGN. 1 sigma confidence levels are calculated according to the prescription of Cameron (2011).
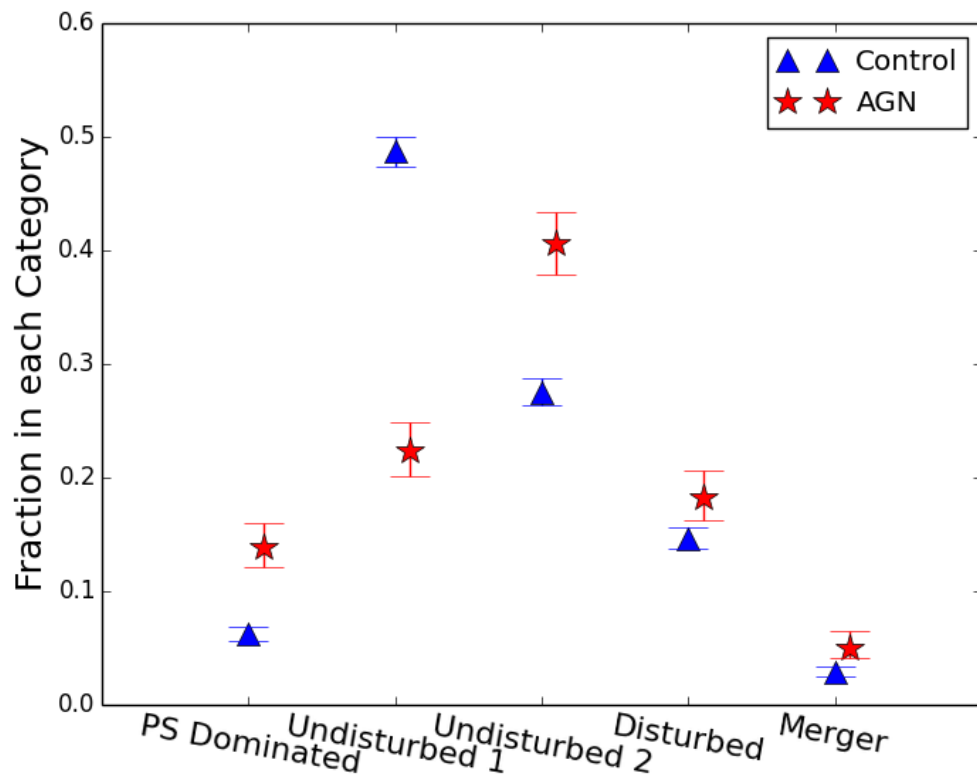
Figure 2.10: The total fraction of AGN and control galaxies classified by humans as mergers, disturbed, undisturbed with residuals, undisturbed without residuals or point source dominated. Blue triangles are control galaxies and red stars AGN. As in figure 2.9 1 sigma confidence levels are calculated according to the prescription of Cameron (2011).
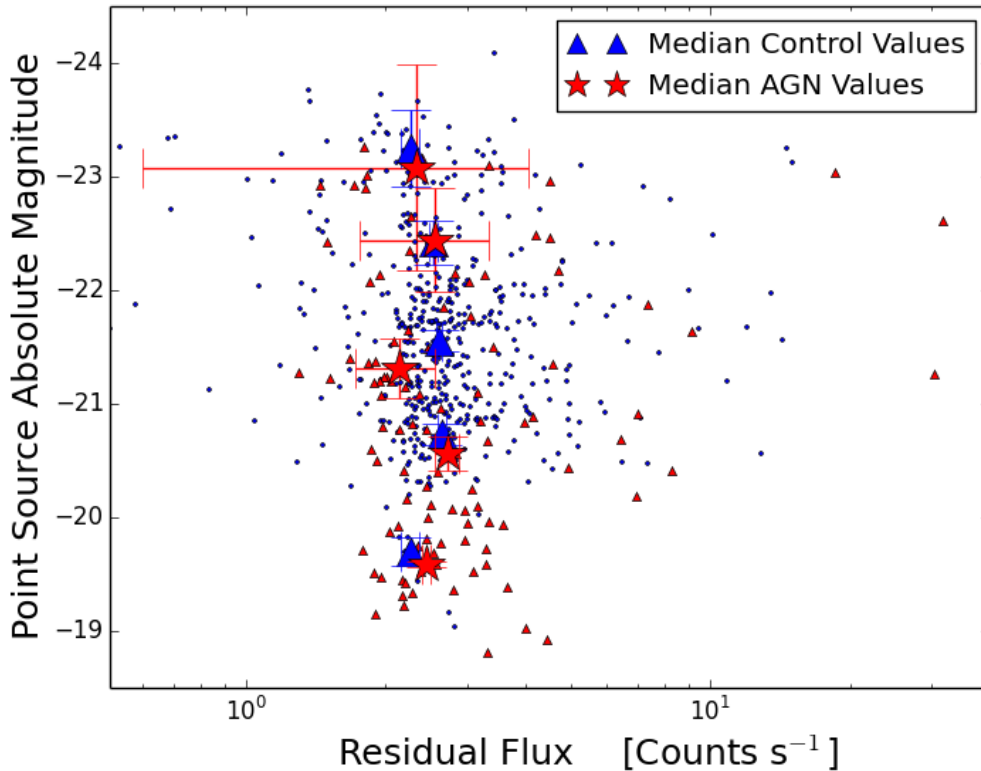
Figure 2.11: Residual flux of AGN + hosts and control galaxies as a function of the point source absolute magnitude, corrected for poor subtraction of point sources. Small blue circles and red triangles represent all values of control galaxy and AGN residual flux; large blue triangles and red stars are the corresponding median values of residual flux in bins of point source absolute magnitude. All residual fluxes values have had +2 added for clarity; this is merely because ~15% of points have negative values once the erroneous point source contribution has been subtracted, and these points do not appear on a logarithmic scale. The significant comparison is between the relative median values, which is unaffected by a uniform addition. Errors are as in figure 2.6.

galaxies more bulge-dominated (Toomre & Toomre, 1972; Barnes, 1988). Finding comparable disk components in AGN hosts and control galaxies suggests that these AGN do not live in especially disturbed or evolved systems. However, as described in section 2.2.1, models were often improved by holding some parameters fixed, so the calculated mean Sersic index may be somewhat biased by this and care should be taken in interpreting the results.

To address point 2, the question of the impact of mergers on AGN luminosity: if the most powerful AGN must be triggered by merging galaxies then in the highest luminosity bins the AGN should be in the clearest mergers (highest residual flux) compared to the control galaxies. Figure 2.11 shows the residual flux, corrected for PSF residuals, of AGN hosts and control galaxies as a function of the absolute magnitude of the point source. Control galaxy artificial point sources have the "absolute magnitude" that the point source would need in order to have the measured apparent magnitude at its host galaxy redshift. No enhancement in residual flux is found with increasing AGN luminosity relative to the control group. Only ~3% of
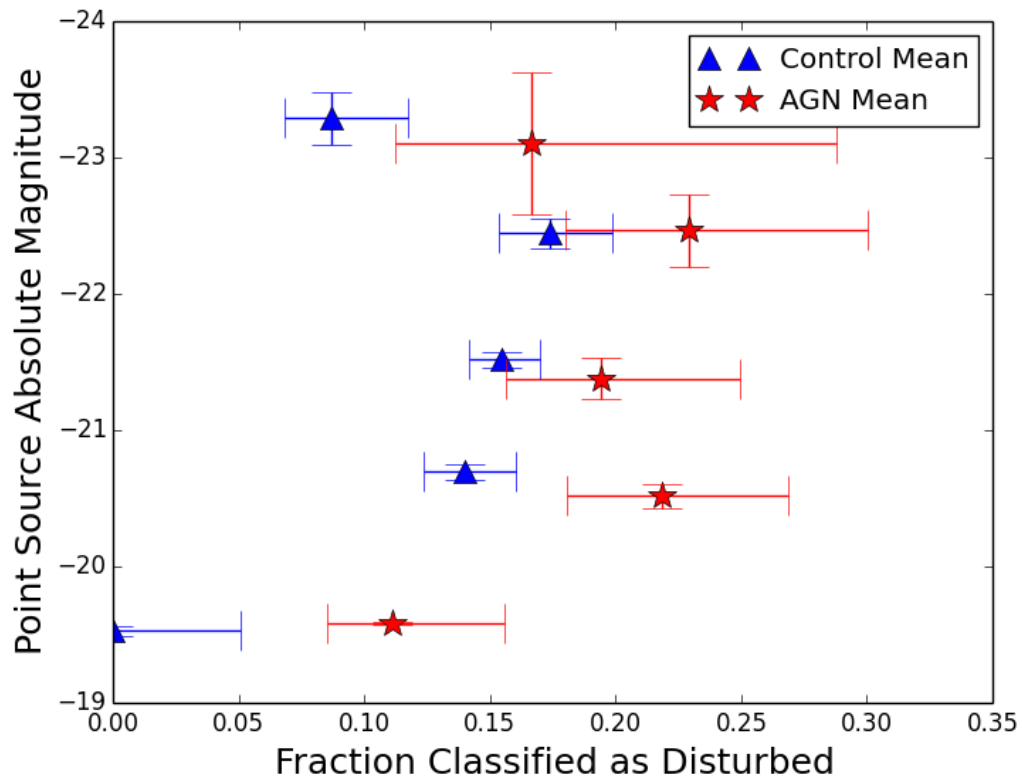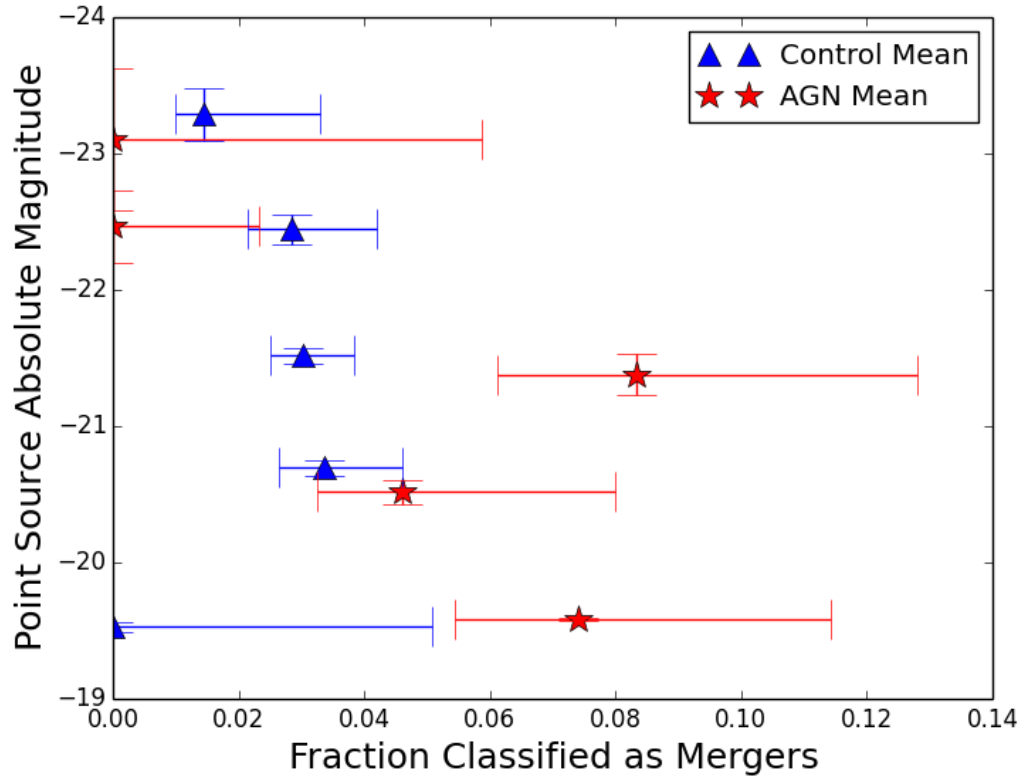
Figure 2.12: The fraction of galaxies classified by humans as mergers (upper panel) or disturbed (lower panel) in bins of point source absolute magnitude. 1 sigma confidence levels are calculated according to the prescription of Cameron (2011).
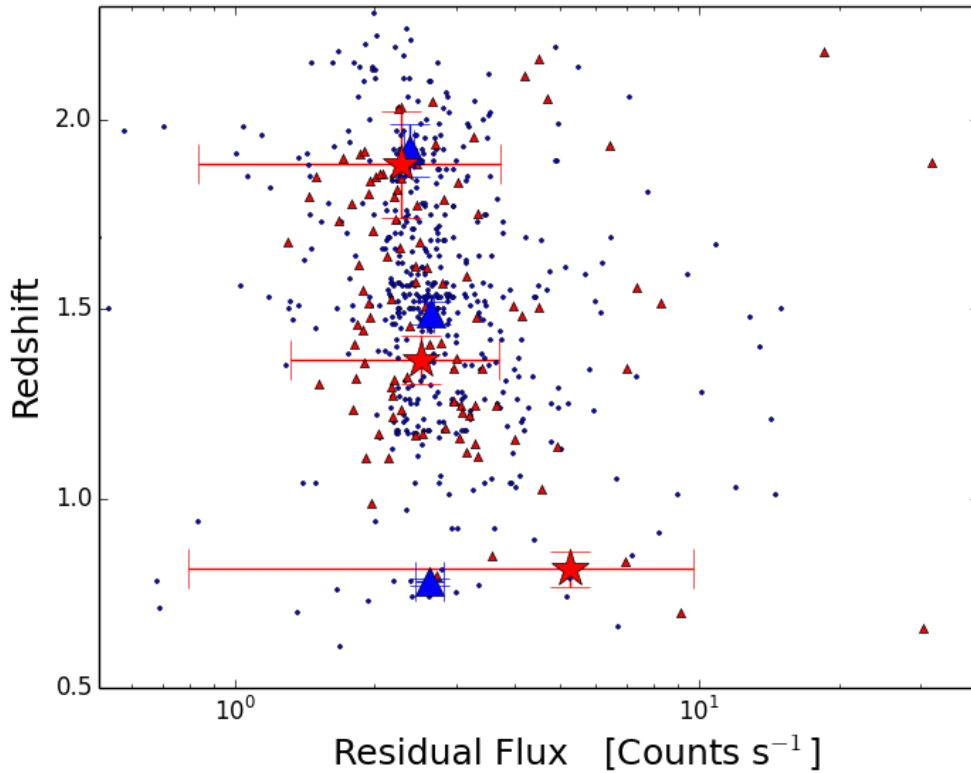
Figure 2.13: Relation between redshift and residual flux. As in other plots blue points are control galaxies; small red triangles are AGN; and the large blue triangles and large red stars represent the median values of controls and AGN respectively, with associated bootstrap errors for different bins of redshift. As with figure 2.11 all residual flux values have had +2 added for clarity; this is merely because ∼15% of points have negative values once the erroneous point source contribution has been subtracted, and these points do not appear on a logarithmic scale. The significant comparison is between the relative median values, which is unaffected by a uniform addition. Errors are as in figure 2.6.

galaxies are found to be disturbed above 5 counts of residual flux, but this is independent of the AGN/artificial point source absolute magnitude. If mergers were triggering the most luminous AGN the expectation would be for an enhancement of residual flux in AGN hosts over their control galaxy counterparts.

Figure 2.12 shows the fraction of galaxies classified as mergers or as disturbed by three human classifiers in bins of point source absolute magnitude. In the highest luminosity bins there is no merger excess, while ∼3% of control galaxies at all luminosities are found in merging galaxies. Both controls and AGN hosts are consistent with a constant fraction of mergers (∼3%) or disturbed (∼15%) galaxies at all point source magnitudes. No link is found between the most luminous AGN and galaxy mergers. For a comparison between residual flux and visual classification see section **??**.

One of the primary aims of this investigation was to see if there was any evolution of triggering mechanisms with redshift. Figure 2.13 shows the correlation between residual flux and
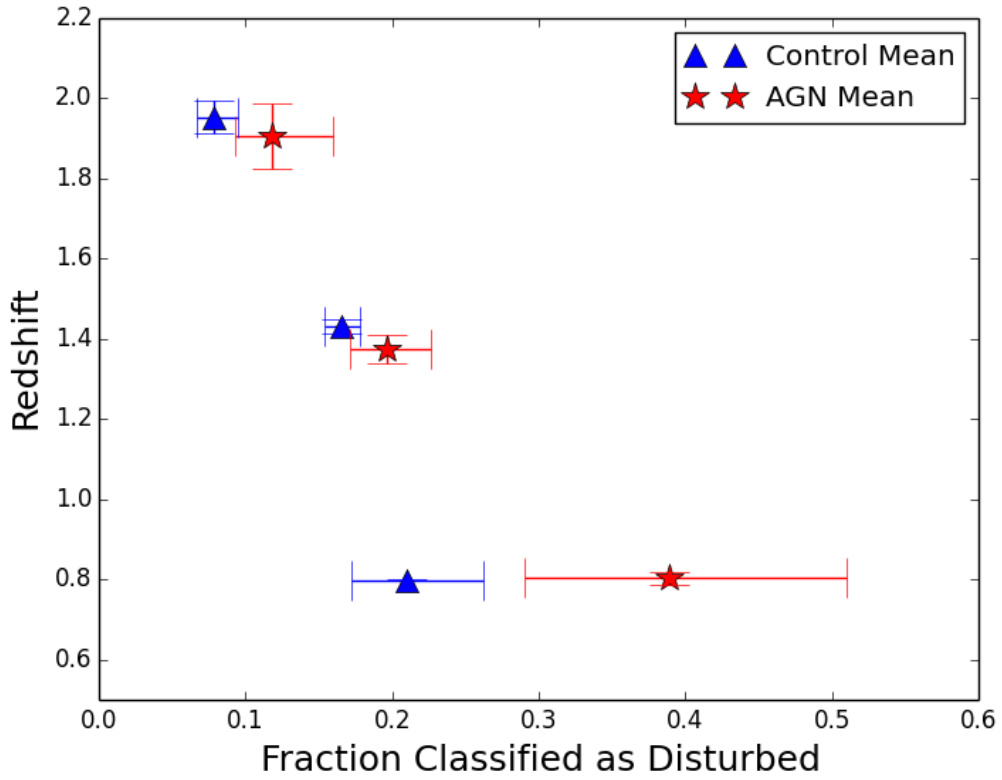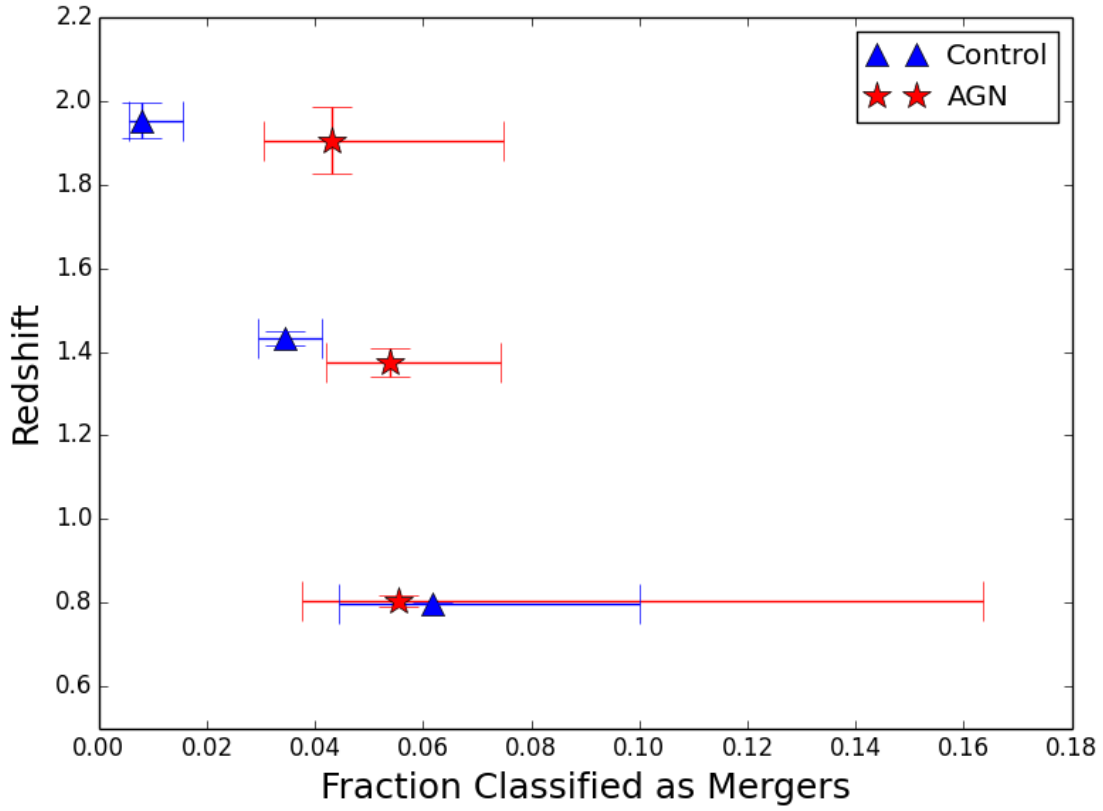
Figure 2.14: The fraction of galaxies classified by humans as mergers (upper panel) or disturbed (lower panel) in bins of redshift. 1 sigma confidence levels are calculated according to the prescription of Cameron (2011).

redshift, with similar residual flux in all redshift bins. Figure 2.14 shows the fraction of galaxies with "disturbed" or "merger" classifications in bins of redshift. A factor of ∼4 enhancement of AGN in galaxies visually classified as mergers relative to the control galaxies is detected at z∼2 at 99% confidence level, 2.4σ. This was determined from the binomial confidence limits as prescribed by Cameron (2011). We note the apparent evolution in the upper panel of figure 2.14, with a smooth transition to increased merger dominance for AGN at higher redshifts, though only the highest redshift bin is statistically significant. Combining the "merger" and "disturbed" classifications, thereby improving number statistics but possibly slightly reducing the robustness of the classification, gives a factor of two enhancement of AGN in morphologically disrupted galaxies at 98.5% confidence level. This provides evidence that AGN are more common in disturbed galaxies at high redshift, where mergers are most common (Cedric & Shaun, 1994; Fakhouri et al., 2010), but at lower redshifts other processes are dominant. This enhancement of merger fraction at z∼2 has not been found in other studies (Gabor et al., 2009; Allevato et al., 2011; Cisternas et al., 2011; Schawinski et al., 2011; Kocevski et al., 2012; Treister et al., 2012; Villforth et al., 2014; Bruce et al., 2015; Villforth et al., 2017). This is the largest study to date of the morphologies of high redshift luminous AGN. The better number statistics may allow for detection of this mild enhancement; this possibility is explored in some detail in the discussion below.

2.4   Discussion

In this study the host galaxies of 106 high luminosity AGN from COSMOS are mass and redshift matched to 486 control galaxies and their morphological properties compared in order to attempt to answer the question: are mergers a dominant AGN triggering mechanism? GALFIT models of galaxy light profiles are made. Models are subtracted from images, and the residual flux serves as a measure of disturbance in the galaxies. Galaxies are visually classified by three experts, identifying likely mergers. Comparison to a group of mass and redshift matched control galaxies allows for analysis of the role of galaxy mergers in triggering AGN. Most galaxies are found to be best fit by exponentially declining light profiles, indicative of disk dominated morphologies (Lin & Pringle, 1987), in contradiction to post-merger morphologies as predicted by some models (Toomre & Toomre, 1972; Barnes, 1988; Di Matteo et al., 2005; Cox et al., 2008). It is found that the fractions of AGN and control galaxies are equivalent for all residual flux values. Similarly, the fractions of AGN and control galaxies in different visual classes are equivalent. In addition, no correlation is found between AGN luminosity and residual flux, and visual inspection of morphology yields no indication that more luminous AGN are hosted by merging galaxies. Residual flux values are approximately constant across the whole range of redshifts, but there is a 2.4σ enhancement of the visually classified merger fraction of AGN hosts relative to control galaxies at z∼2. This may imply some evolution in

the dominance of merger-triggering through time. These results will now be discussed in the context of the wider literature.

First we consider the impact of selection bias on the results. The AGN studied here are X-ray selected. Are these representative of the underlying AGN population? It has been argued that merger-induced nuclear activity is more likely to be Compton-thick (Cattaneo et al., 2005; Kocevski et al., 2015), therefore selecting AGN in the X-ray biases against observing merger signatures contemporaneously with the AGN emission, which is obscured. It is important to bear this in mind when interpreting these results, but also to recall biases associated with other selection methods. Infrared (IR) emission, for instance, allows for selection of AGN by relying on the approximately power-law spectrum of AGN in the IR (Stern et al., 2012). This can, however, bias in favour of finding mergers, since mergers can induce starbursts (Mihos & Hernquist, 1994, 1996; Hopkins et al., 2006) which produce copious quantities of dust and boost the IR signal from AGN, enhancing detection rates (Stern et al., 2012). Indeed, studies of IR selected AGN have found enhancements of AGN fractions in merging galaxies relative to controls (e.g. Satyapal et al., 2014). Selecting AGN using X-rays may bias results away from finding mergers, but certain results remain true regardless: for example, that mergers are not *required* to power the most luminous AGN, since we find so many luminous AGN unrelated to mergers.

What do the GALFIT models reveal about the galaxies in this sample? Some models predict that major galaxy mergers should destroy disks (Toomre & Toomre, 1972; Barnes, 1988; Di Matteo et al., 2005; Cox et al., 2008) and post-mergers are thought to relatively quickly evolve into elliptical galaxies (unless gas-rich enough to quickly reform a disk; Robertson et al., 2006b). Elliptical galaxies have Sersic indices of ∼4 (Mellier & Mathez, 1987) while disk dominated systems have Sersic indices closer to 1 - an exponentially declining light profile (Lin & Pringle, 1987). If the AGN hosts have recently undergone a merger, driving the nuclear activity, the Sersic indices may be expected to tend toward values indicative of bulge dominated systems. GALFIT models reveal an average Sersic index of ∼1.15 with similar Sersic index distributions for AGN and control galaxies, suggesting that most of the galaxies in this sample are disk dominated. As mentioned in section 2.2.1 these models may be biased by the requirement to fit a bright point source, otherwise this provides some evidence that these are not mergers with faint merger features, or late mergers where features have faded, but instead are predominantly secularly evolving galaxies. This is in agreement with Cisternas et al. (2011); Schawinski et al. (2011); Kocevski et al. (2012); Villforth et al. (2014) and Villforth et al. (2017), extending the work to greater redshift and/or luminosity.

Next we shall discuss the relative fractions of AGN and control galaxies in different morphological categories, and their distributions in residual flux. Figures 2.9 and 2.10 demonstrate

the similarity between the AGN host and control galaxy morphologies. Similar fractions of active and inactive galaxies are found in merging or disturbed systems. Such a small fraction of AGN are found to exist in merging galaxies that either the majority of black hole growth must occur via alternative processes, the AGN must be triggered after merger features have faded, or mergers are simply undetectable here (though this is rendered unlikely by comparisons to simulations; Kaviraj et al., 2013).

In addition, it is found here that even the most luminous AGN are no more likely to be found in merging galaxies (see figures 2.11 and 2.12), in contradiction to the conjecture that major merger triggering of AGN should be a strong function of luminosity (Finn & Impey, 2001; Treister et al., 2012) with the most luminous AGN triggered solely by mergers. This is in opposition to some previous studies (e.g. Treister et al., 2012; Ellison et al., 2016) and in agreement with others (e.g. Cisternas et al., 2011; Allevato et al., 2011; Karouzos et al., 2014; Villforth et al., 2017). We cannot conclusively rule out the possibility that a time delay between mergers and AGN switching on is responsible for merger features significantly fading, however the disk-like morphologies suggest major mergers have not recently occurred in the majority of the sample, making this more unlikely (other studies have also come to this conclusion due to the dominance of disks in the samples, e.g. Gabor et al., 2009; Cisternas et al., 2011; Schawinski et al., 2011; Kocevski et al., 2012; Villforth et al., 2014, 2017). At high redshifts, where gas densities are higher, some simulations show that these galaxies may be able to relax into disks again (Robertson et al., 2006b), though since this may require specific orbital configurations, and not all the galaxies investigated here are at especially high redshift, this seems implausible to account for such a large portion of the sample here. Treister et al. (2012) found a strong correlation between AGN luminosity and merging status, with just 4% of AGN with bolometric luminosities of 10^{43} erg s^{-1} in mergers and $\sim$90% in mergers at 10^{46} erg s^{-1}. Whilst this appears to support the notion that the highest luminosity AGN are triggered by mergers, a range of AGN selection criteria were compiled into the study and in some cases the true AGN luminosities are uncertain. Additionally, the lack of a control group makes interpretation of these results problematic.

Through visual analysis of 140 X-ray selected AGN with 1264 control galaxies with HST data in COSMOS, Cisternas et al. (2011) find no merger-AGN connection from z=0.3-1. Similarly to this work, they go further and demonstrate the implausibility of a time delay between mergers and the peak of SMBH accretion being responsible for this null result, since the morphologies are often disk dominated. Allevato et al. (2011) measure the halo mass bias of different classes of AGN with redshift, concluding that the increase in bias with redshift (z<2.25) suggests that different AGN phases (classic Seyfert; quasar mode; radio loud) are purely related to the halo mass. Semi analytic models assuming major merger triggering of AGN cannot reproduce these biases for type 1 AGN, such that alternate triggering processes

(secular; tidal disruptions; disk instabilities) may be more significant than major mergers for moderate luminosity AGN up to z~2.2 (Allevato et al., 2011), a conclusion bolstered by the findings of this study. In short, this work broadly supports the findings of other observational studies, though differing AGN selection methods may obfuscate the issue.

The final piece of evidence is the correlation between residual flux or visual classification and redshift in figures 2.13 and 2.14. First we will discuss the residual flux result. Figure 2.13 shows similarity between AGN and control galaxies in residual flux across the redshift range. One may be concerned that merging galaxies will have similar enhancement of residual flux compared to galaxies with non-axisymmetric features, such as spiral arms and clumpy disks. Over this redshift range the F814W band measures rest frame UV through to visual; since star formation is intrinsically clumpy (Bournaud et al., 2011; Trump et al., 2014) this may manifest itself as excess residuals at high redshift. We did not attempt to fit non-axisymmetric structures (section 2.2.1), instead choosing to compare residual flux to visual classification. Visual class and residual flux correlate well (see figure 2.6).Visual classifiers had the opportunity to highlight non-axisymmetric structures for each galaxy: their exclusion did not affect the result, and further inspection of images with high residual flux reveals tidal tails and structures that do not look like clumpy disks or spiral arms, ruling this out as the probable cause of similarity (See figure 2.8).

Considering visual classification, figure 2.14 shows the fraction of AGN hosts and control galaxies classified as mergers in bins of redshift. AGN hosts are found to have a higher incidence of mergers at high redshift, with a factor of ~4 enhancement of AGN in mergers at z~2 with 99% confidence. Considering disturbed and merger classifications together as one class gives a factor of ~2 enhancement with 98.5% confidence at z~2. We consider whether a mass mismatch between the AGN hosts and control galaxies could lead to a difference in merger incidence between AGN hosts and controls. If, for instance, the M-M_{bulge} relationship evolves with redshift this could lead to a mass mismatch, and galaxies of different masses may undergo major mergers with different frequencies. Stewart et al. (2009) and Rodriguez-Gomez et al. (2015b) predict the merger frequency to be a function of mass, with more massive galaxies having slightly higher merger rates. At z~2 there are approximately twice as many major mergers for galaxies with $M_* \sim 10^{11} M_\odot$ as there are for $M_* \sim 10^{10} M_\odot$ (Rodriguez-Gomez et al., 2015b). This suggests any mass mismatch here would have to be implausibly severe to entirely explain the observed factor of ~4 enhancement. Were the mass mismatch so severe, the model magnitude distributions for AGN hosts and control galaxies would be extremely different to each other: in reality they are comparable. The M-M_{bulge} relation would be expected to evolve slowly, if at all, to have avoided robust detection in previous studies (Robertson et al., 2006a; Lauer et al., 2007; Shankar et al., 2009; Shen et al., 2015), so it is likely that control and AGN host galaxies are similarly massive.

Next, we consider the possibility that this result has remained undetected in previous studies due to poor number statistics. Work by Mechtley et al. (2016) constitutes the best comparison to the study presented here, since they consider a sample of 19 high luminosity AGN at z∼2, compared to the 32 high luminosity AGN at z∼2 here. Mechtley et al. (2016) found a statistically insignificant enhancement in the probability of the fraction of AGN in mergers being greater than the fraction of controls of 0.78σ. Similar to figure 3 in Mechtley et al. (2016), we show the probability density functions of the 32 z∼2 AGN and their respective control galaxies in figure 2.15. Assuming the AGN in Mechtley et al. (2016) are intrinsically drawn from the same underlying population of AGN as those in this work, we can explore the likelihood of finding a statistically significant enhancement of the merger fraction in a smaller sample than presented here. Randomly drawing 19 AGN and 84 control galaxies from the sample (matching to the number of AGN and controls in Mechtley et al., 2016), finding the statistical difference in merger fraction between the two populations, and repeating multiple times yields a wide (Gaussian) distribution of statistical significances. The probability that Mechtley et al. (2016) would find a significance of $<1\sigma$ is ∼24%. Work to date has not found a statistically significant redshift evolution of AGN in mergers (Gabor et al., 2009; Allevato et al., 2011; Cisternas et al., 2011; Schawinski et al., 2011; Treister et al., 2012; Bruce et al., 2015; Mechtley et al., 2016); it will be interesting to see if this elevation persists in future large high redshift samples, since we are now possibly approaching large enough samples to detect such an enhancement.

Assuming this is a real signal and z∼2 AGN are more likely to be merger triggered than their low redshift counterparts the cause of such a redshift evolution should be considered. One naive expectation would be that high redshift AGN, sitting in relatively gas rich environments, would less require catastrophic events such as mergers to drive accretion. However, factors other than gas fractions may play some role: at high redshift galaxy mergers are more common (e.g. Rodriguez-Gomez et al., 2015b); stellar mass is assembled more rapidly (with higher star formation rates); the largest bound structures (massive galaxy clusters) are yet to form and galaxies are more compact (e.g. Ellison et al., 2018). The influence of each of these factors on BH growth can be considered in turn. If galaxy mergers contribute a modest enhancement to the AGN duty cycle their influence could easily be overlooked in low redshift samples of AGN and controls, but where mergers are more common by a factor F this enhancement would be similarly enhanced by F, assuming a constant contribution to the duty cycle. Major mergers are likely ∼10× more common at z∼2 than at z∼0 (Rodriguez-Gomez et al., 2015b) and so the signal from an enhancement to the AGN duty cycle due to mergers could be expected to be ∼10× greater at high redshift than at low redshift. The processes driving enhanced star formation rates at high redshift are likely to be significant to AGN, given studies suggesting universal links between star formation and nuclear activity (Wild et al., 2010; Mullaney et al.,
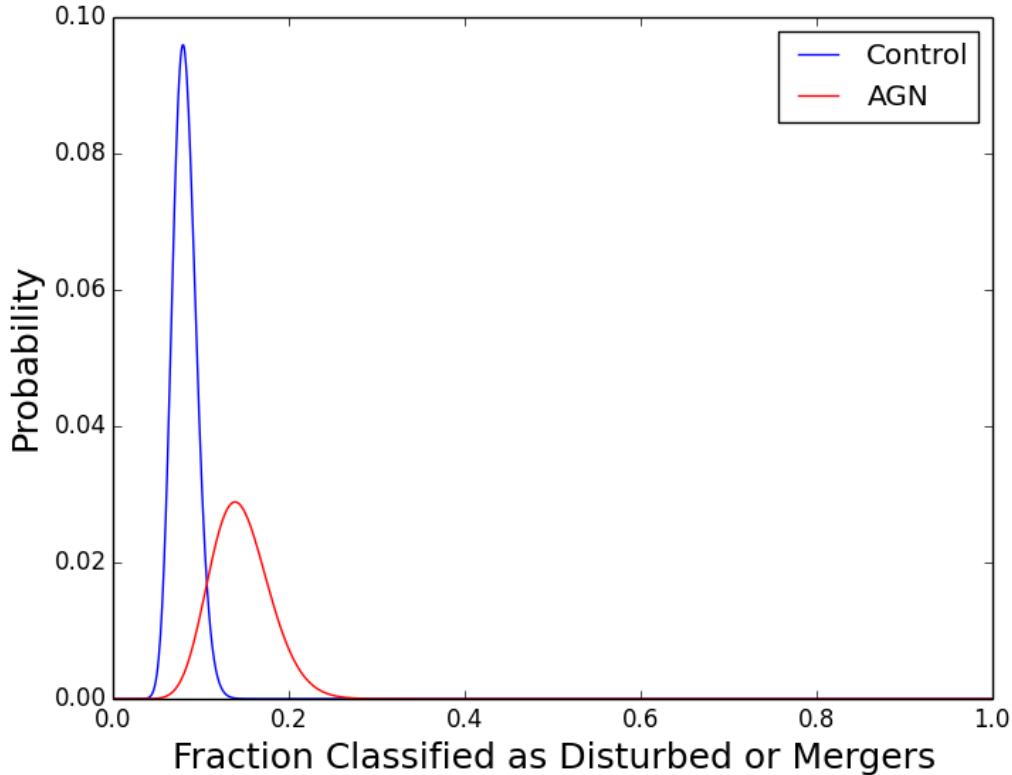
Figure 2.15: Probability density functions, calculated in accordance with Cameron (2011), for the fraction of AGN (wider red line) and controls (narrow blue line) classified as either mergers or disturbed.

2012; Sabater et al., 2013, 2015). If the crucial factor associated with nuclear activity is the rate of nuclear star formation then at high redshift, with greater gas densities and more compact galaxies, mergers may trigger more star formation and boost the probability of black hole growth relative to low redshift examples. In chapter 3 it will be shown that higher star formation rates additionally can make mergers much more easily detected in imaging, possibly biasing the detection of mergers with redshift. As galaxies fall into massive clusters their gas can be removed by ram pressure stripping - pressure from the hot intra-cluster medium - and so mergers within clusters may be gas poor and unlikely to trigger AGN. At high redshift clusters are smaller and the effect of ram pressure stripping is proportionately lessened, leading to a possible small bias in favour of merger-induced nuclear activity at high redshift.

Overall AGN are not found more commonly in galaxies exhibiting signs of recent major mergers, but an excess of mergers is found in the highest redshift bin (z~2). No evidence is found in support of conjecture that the most luminous AGN must be fuelled predominantly by mergers, in contradiction to Treister et al. (2012) but in accordance with other studies (Cisternas et al., 2011; Allevato et al., 2011; Karouzos et al., 2014; Mechtley et al., 2016; Villforth et al., 2017). An enhancement of merger fraction in AGN hosts is found in the highest redshift bin, z~2, despite the results of other studies (Gabor et al., 2009; Allevato et al., 2011;

Cisternas et al., 2011; Schawinski et al., 2011; Bruce et al., 2015; Mechtley et al., 2016). The majority of galaxies here are unlikely to have recently undergone a merger, given their disk-dominated light profiles (Toomre & Toomre, 1972; Barnes, 1988; Di Matteo et al., 2005; Cox et al., 2008; Gabor et al., 2009; Cisternas et al., 2011; Schawinski et al., 2011; Kocevski et al., 2012; Villforth et al., 2014, 2017).

2.5 Conclusions

The aim of this study was to determine if a link exists between galaxy mergers and AGN activity, and whether this link evolves with redshift or AGN luminosity. High resolution HST images of 106 X-ray selected AGN and 486 host galaxy mass and redshift matched control galaxies from the COSMOS survey are analysed. Artificial AGN (stars), matched in apparent magnitude to the relevant AGN, are placed into control galaxies to make images equivalent. A new method for detecting disturbance in galaxies is developed: subtracting 2D photometric decomposition models from images (using GALFIT) and summing the remaining flux yields a numerical measure of merger features, since asymmetries and disturbances in the light profiles leave greater residual flux. The method correlates well with visual classifications.

The key findings are summarised below, relating the residual flux and visual classification by three experts to AGN properties.

i) Neither visual classification nor the residual flux method find evidence of greater fractions of AGN in merging galaxies, with similar fractions of merging control galaxies at all levels of disturbance.

ii) Neither method finds evidence that mergers play a more dominant role at higher AGN luminosities, with similar fractions of merging control galaxies and AGN hosts at all luminosities. This suggests even the most luminous AGN are predominantly triggered secularly.

iii) Moderate evidence for redshift evolution of the role of mergers in AGN triggering is found for the first time, with a factor of ∼4 enhancement of the visually classified merger fraction in AGN hosts at z∼2 relative to control galaxies, with 99% confidence.

iv) The fraction of galaxies exhibiting evidence of recent mergers in this sample is low (≤20%), suggesting that the majority of X-ray selected AGN at high luminosity are likely to be triggered alternatively, perhaps secularly or by minor mergers, at 0.5≤z≤2.2.

Although the single-filter approach to morphological analysis means that star forming regions will be over-sampled at high redshift as the rest frame shifts into the UV, by ensuring

control galaxies are similar to AGN hosts like-for-like comparisons can still be made with validity. The residual flux technique is an effective way of picking out mergers in an automated way, and could be utilised on larger datasets, particularly with the aim of finding if only the most catastrophic mergers cause AGN of all luminosities.

The detection of an enhancement of the incidence of mergers at z~2 is reported tentatively, given its absence in previous studies. Future work will have to use large samples (>30 luminous AGN) and find a similar enhancement of the merger fraction to lend credence to the results here.

3

The Impact of Major mergers on AGN: Interpreting Observations with Illustris

Although integral to many models, the extent to which galaxy mergers trigger AGN remains an open question. Identifying galaxy mergers, either by human visual classification or automated measures, is difficult due to the transience and low surface brightness of features. Measuring other parameters such as galaxy mass ratios or the time elapsed since a merger is even harder, and so robust tests of AGN triggering models remain elusive. Using the Illustris simulation, realistic mock SDSS images of major mergers, minor mergers and isolated galaxies (at z=0) were created to test a range of possible triggering models: assuming mergers trigger AGN with some probability we can uncover what would be detected in the real universe using visual classification or structural parameters, and compare to data. This allows quantification of the increase (henceforth, enhancement) of AGN duty cycles (the fraction of time for which the AGN is "on") as due to mergers as a function of merger mass ratios, time elapsed since mergers, star formation rates etc. The detailed findings of the simulation with respect to the triggering and evolution of black holes are, while interesting, inconsequential to this study since we do not select AGN within the simulation, only galaxies which could host AGN. Our aim is to build samples of galaxies from the Illustris simulation which correspond to different merger-triggering efficiencies for AGN. By constructing fake samples of AGN hosts and control galaxies, with the fraction of mergers in the AGN sample as a free parameter, it is possible to

determine what would be observed in the real universe for a given merger-enhancement to the AGN duty cycle.

The Illustris simulation reproduces many of the observed trends seen in the Universe. The cosmic evolution of star formation rate density, the galaxy luminosity function and the proportions of early- and late-type galaxies all match well to data (Vogelsberger et al., 2014a,b). Whether or not this implies that the physics responsible for producing structures has been understood, or that ΛCDM is the correct cosmological theory, is not of primary significance here. We wish to analyse fake Illustris images similarly to real galaxies, classifying them visually and with structural parameters for comparison to real data. For this purpose all that is required is for Illustris galaxies to resemble their counterparts in the real universe with the correct proportions of different galaxy types, to have realistic star formation histories and for merger histories to resemble the merger histories of real galaxies[1], all of which the Illustris simulation accomplishes. Illustris galaxies suffer from a number of [relevant] minor issues: ∼10% of galaxies with $M_* \sim 10^{10.5 - 11} M_\odot$ have unphysical ring-like structures; the galaxy colour distribution is not clearly bimodal, with an over-populated blue- and green-sequence, and the star formation rate density is too high at low redshift (Genel et al., 2014; Vogelsberger et al., 2014a; Snyder et al., 2015). With these caveats, broadly the galaxy populations observed match well enough to the Illustris simulation for a useful comparison with existing data.

Through comparison to studies in the literature we find that mergers (mass ratios <10:1) may increase the AGN duty cycle by as much as a factor of ∼10-40, with mergers responsible for ∼50±20% of nuclear activity in the local Universe. The detectability of mergers is a strong function of the star formation rate. Intrinsic underlying correlations between mergers and nuclear activity are likely hidden by delays between merger-induced starbursts and AGN activity, causing mergers to become undetectable by most commonly used morphology measures around the peak of merger-induced nuclear activity.

In section 3.1 we describe the relevant details of the Illustris simulation; in section 3.2 we describe the calculation of structural parameters and the process of visual classification; we present results of both the efficacy and the expected observational signatures for different merger-triggering efficiencies of those measures in section 3.3; implications and comparison to real data are discussed in section 3.4 before conclusions are stated in section 3.5.

---

[1] We note that modified gravity theories generically predict fewer, longer lived mergers (due to the lack of dynamical friction from dark matter halos). If MOND was shown to be favourable to dark matter the theoretical underpinning of this work, and so the results derived, would be called into question. However, the number density of mergers in a MOND cosmology would still have to match to observations, as in Illustris, so the influence on our results in this case would most likely be small.

3.1 The Simulation

Here the key features of the Illustris simulation are summarised before the relevant details are outlined below. The Illustris simulation is a large volume cosmological hydrodynamical simulation including a suite of models governing the baryonic physics (Vogelsberger et al., 2014a,b; Genel et al., 2014). The formation and evolution of galaxies in a $(106.5 \text{ Mpc})^3$ volume cube are simulated in five different components simultaneously: dark matter particles (resolution $= 6.3 \times 10^6$ M$_\odot$ per particle), gas particles, gas tracers, stars and stellar wind, and supermassive black holes (baryonic resolution $= 1.6 \times 10^6$ M$_\odot$ per particle Nelson et al., 2015) with $N_{\text{DM}} = N_{\text{Baryon}} = 1820^3$ resolution elements [2]. Particles are distributed at z=127 in correspondence to the density profile expected from observations of the CMB and tracked through their evolution to present day. Six runs of the Illustris simulation were released (three different resolutions, each with a full baryonic model and a dark matter only run): in this work we make use of the highest resolution run, Illustris-1. Henceforth all discussion refers to Illustris-1 only.

To build realistic galaxies for comparison to the real Universe the hydrodynamics governing the baryonic physics must be realistically incorporated. Baryonic physics models in Illustris include sub-grid prescriptions for gas recycling and chemical enrichment, supernovae (SNe) and AGN feedback, metal line cooling, ISM pressurisation due to SNe and stochastic star formation in gas above $0.13 N_{\text{H}}$ cm^{-3} (Springel & Hernquist, 2003). The data are split (with uneven spacing in redshift and time) into 136 snapshots spanning $z = 47$ to $z = 0$, including catalogues at each snapshot detailing halo and subhalo properties and merger trees derived from these catalogues. Multi-band mock images, vital for the purposes here, are available for all galaxies at $z = 0$ with stellar mass $> 10^{10}$M$_\odot$ from four viewing angles. Illustris data is available in full online [3].

To construct mock images the particles associated with a galaxy are assigned a spectrum, including some obscuration and contribution from neighbouring galaxies, using the radiative transfer code SUNRISE (Jonsson, 2006), from which spatially resolved images are constructed (Torrey et al., 2015). In Torrey et al. (2015) stellar population synthesis models were produced for each star particle in massive galaxies (M $> 10^{9.7}$M$_\odot$) in the simulation and convolved with various transmission functions, yielding spatially resolved images in 36 photometric bands. Usefully for this study, AGN emission is not included in the imaging (Torrey et al., 2015). In order to ensure mock images were available for all galaxies selected, only galaxies with stellar mass $> 10^{9.7}$M$_\odot$ were selected in this study, obtained from the online servers. Observational

---

[2]It is interesting to note that since the first computerised astrophysical simulations the number of simulated resolution elements has grown exponentially. If this trend were to persist, a naive extrapolation suggests every atom in the observable Universe could be simulated in just ~300 years.

[3]http://www.illustris-project.org/data/

realism was added using the publicly available code SUNPY (Mumford et al., 2015). We selected mock images in the i band (near infrared, with a limiting i band magnitude of 21.3 mag/", as per the SDSS) to ensure photometric maps were representative of the underlying stellar populations (sampling older stars). In order to resemble low-redshift studies as closely as possible images were then redshifted to z = 0.05 and convolved with an SDSS point-spread-function (PSF). Mock images created in this way are remarkably indistinguishable from real SDSS images.

In this analysis mock samples of Illustris galaxies were compared to observational studies of real galaxies, so it is important for the properties of simulated galaxies to closely match observed properties of real galaxies. Broadly speaking the comparison is favourable, with good general agreement between the simulation and observations, with a few caveats. The cosmic evolution of star formation rate density was found to match quite well to observations, though at z<1 (relevant here) is over-predicted by a factor of ∼2 (Vogelsberger et al., 2014b; Genel et al., 2014). Similarly, the number densities of high ($M_*>10^{11.5}M_\odot$) and low mass ($M_*<10^{10}M_\odot$; not relevant to this work) galaxies are slightly over-predicted (Vogelsberger et al., 2014b; Genel et al., 2014). Simulated galaxy colour distributions fail to show a clear bimodality, as expected, and contain too many blue galaxies at low-intermediate masses ($10^9 < M_*[M_\odot] < 10^{11}$), though the relationship between environment and galaxy colour is recovered reasonably well (Vogelsberger et al., 2014b). The black hole mass function at z=0 matches well to observations in the local Universe, with perhaps a slight overestimation of the density of low mass SMBHs ($M_{BH} < 10^7M_\odot$; Sijacki et al., 2015), though such black holes are rare in this study.

Illustris data is accessible online and downloadable in hdf5 format. The simulation has a total data volume ∼300Tb, but the web-based api is insufficient for some advanced searches and so a combination of the two were used. Data is split into snapshots, group catalogues and merger trees. Snapshots describe the simulation at some point in time and include the number of each particle type, including time-varying tracer particles (as gas is converted to stars for example), split into groups and subgroups. Data is sorted into a mass hierarchy first by the galaxy group number, then the subgroup and finally by binding energy (for those particles associated to a group but not a subgroup) (Nelson et al., 2015).

Group catalogues contain information about halos and subhalos at a given snapshot. Data are organised such that the most massive halo is assigned an ID of 0, with associated subhalos, the second most massive an ID of 1 etc. Subhalos are identified either by a friend-of-friends algorithm or the SUBFIND algorithm (Springel et al., 2005). Qualitatively FoF and SUBFIND algorithms are equivalent: we make use of SUBFIND groupings in this work. The SUBFIND algorithm identifies gravitationally bound substructures and assigns them to subhalos appro-

priately, while the FoF algorithm links substructures by length. For details see Nelson et al. (2015). Group catalogues contain details of subhalo masses, radii and star formation rates, along with many other quantities of interest to this work.

From group catalogues merger trees are constructed, since the association of particles to particular subhalos is known and tracked through time. Merger trees were used extensively here to construct samples of major mergers (mass ratio 4:1 to 1:1), mergers (10:1 to 4:1), minor mergers (100:1 - 10:1) and control galaxies (no merger with a mass ratio < 100:1 within 1Gyr). These four (arbitrary) categorisations are used throughout. The important aspects of merger tree generation from the SUBFIND method are summarised here, for a more comprehensive description the reader should see Nelson et al. (2015) (particularly figure 4) and Rodriguez-Gomez et al. (2015a). Merger trees are constructed for both dark matter particles and baryons: in this project the baryons are the constituents of interest, since it is the baryons we observe. Subhalos are tracked from one snapshot to the next, each subhalo which contains particles from the subhalo in the previous snapshot is given a score based on the binding energy rank the particles had in that previous snapshot (a merit function), and the subhalo with the highest score is assigned as the unique descendant (Rodriguez-Gomez et al., 2015a). This has the advantage of weighting toward the inner regions of the galaxy, where particles are most tightly bound, instead of the relatively easily stripped outer regions. If a galaxy has high enough velocity it may shoot through another subhalo without merging. In this case a timing coincidence could cause a merger to be fallaciously identified, while one halo sits in the other. To avoid this, subhalos are tracked not just between adjacent snapshots but across gaps in time separated by two snapshots. Due to the hierarchical assembly of galaxies a subhalo may have many progenitors but only a single descendant. As such the merger history of some subhalo can be tracked back to some earlier snapshot, by following the merger tree back through time to find all progenitors.

Since merger trees reveal when two subhalos have the same unique descendent, while catalogues contain information about the masses of those subhalos, it is possible to determine the baryonic mass ratio of any mergers that occur. Walking back in time along the merger trees gives the time after which a merger occurred, since it necessarily occurred following the previous snapshot but before the unique descendant is identified in the current snapshot, i.e. a merger is identified with time resolution corresponding to the snapshot interval ($\sim$150Myr). We are interested in mergers where observational features may still be present, so limit the sample to the most recent five snapshots where $t_{\mathrm{merge}} < 800$Myr. In order to examine the ideal case where features are most observable we consider z$\sim$0 images here only. In addition to samples of mergers with a range of mass ratios we also constructed a sample of control galaxies where it was known that no mergers had occurred in the last Gyr with a mass ratio < 100:1, in order to quantify the efficacy of automated measures and human visual classification in
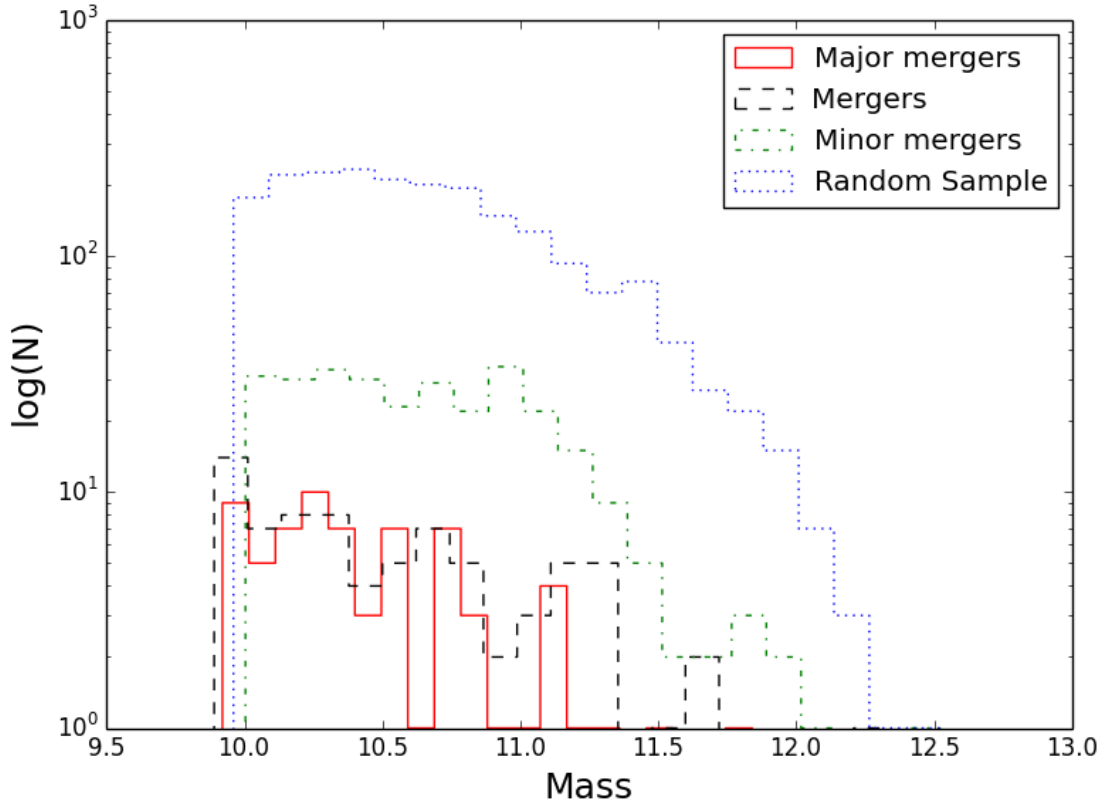
Figure 3.1: Histograms showing the mass distributions of major mergers (defined to have mass ratio 4:1 to 1:1), mergers (10:1 to 4:1), minor mergers (100:1 - 10:1) and a random sample of galaxies.

detecting merger features with a range of physical parameters (time elapsed since the merger, mass ratio etc.). Figure 3.1 shows the mass distribution of the different samples. The sample contains a total of 69 major mergers, 77 mergers and 294 minor mergers, each from four viewing angles, and an additional 2098 galaxies selected at random from the Simulation.

3.2 Methodology

The question "do mergers trigger nuclear activity" is neither precise nor particularly interesting. There are clear examples of AGN in galaxies which show no signs of recent interactions, and of mergers with no apparent AGN. The question should be restated: by what factor is the AGN duty cycle enhanced by mergers? This is harder to quantify due to a combination of observational biases, which we attempt to quantify and interpret in sections 3.3 and 3.4. To interpret existing results in the literature and this thesis we recreate analogues of observational studies, creating mock samples of AGN hosts and matched controls. In these mock samples "AGN hosts" refers to galaxies which we assert host AGN, rather than to AGN in the simulation, since we are not concerned with simulation predictions of where AGN might be, but with the extent to which mergers trigger AGN in the real Universe. The simulated products of interest are the mock images (and data from which mock images are created) and merger histories.

In observational studies images are visually classified or analysed using structural parameters to deduce the merger fraction; however these techniques are fallible, true merger fractions are unknown and so interpretation of results is problematic. Under the assumption that AGN can be triggered by processes unrelated to mergers, if the duty cycle is increased relative to the non-merger duty cycle for some period after the merger by a factor, F, then the fraction of mergers in the AGN sample should be similarly increased by a factor F. The control sample is therefore constructed with a fraction of mergers corresponding to the merger fraction in the local Universe and the AGN sample contains more mergers by a factor F. So long as the duty cycle is <1 and the background fraction of mergers ≪1 it is a reasonable approximation to keep the fraction of mergers in the control sample constant (as opposed to removing some mergers since they might host AGN). By choosing a range of values of F and comparing the measured merger fractions of AGN hosts to controls, determined by analysis of structural parameters and visual classifications as in observational studies, likely values of F can be constrained through comparison to those observed merger fractions. Additionally, F does not have to be just a simple multiplicative factor but can depend on the merger mass ratio, the time elapsed since the merger, the SFR and stellar mass, since these details likely influence the degree of enhancement to the duty cycle from the merger. We now describe the process of calculating automated morphology measures including Gini, M_{20}, concentration, asymmetry and shape asymmetry in section 3.2.1. In section 3.2.2 we present details of the human classification using Zooniverse.

3.2.1 Automated Measures

In order to compare morphologies of selected samples of Illustris galaxies to real galaxies, structural parameters of Illustris images are calculated here. The structural measurements Gini, M_{20} and concentration used here were previously calculated in Snyder et al. (2015), we summarise the calculations and physical significance here and additionally calculate the asymmetry and shape asymmetry. The Gini index, originally a measure used to calculate wealth inequality in economics, is used astrophysically to measure the inequality in the distribution of flux across pixels of a galaxy (Abraham et al., 2003):

$$G = \frac{1}{2\bar{X}|n(n-1)} \sum_{i}^{n} (2i - n - 1)|X_i| \qquad (3.1)$$

where X_i is the pixel intensity in ascending order, $\bar{X}$ the mean pixel intensity and n the total number of pixels. The value of Gini ranges from 0 to 1: equal flux in all pixels to all flux in a single pixel. An old galaxy in equilibrium would be expected to have low values of Gini, while a rapidly star forming or merging galaxy with multiple cores would typically have higher values.

M_{20} provides information about the distribution of light in a galaxy; specifically, M_{20} measures the second order moment of the brightest 20% of the flux, normalised by its total moment (Lotz et al., 2004):

$$M_{tot} = \sum_i^n M_i = \sum_i^n f_i[(x_i - x_c)^2 + (y_i - y_c)^2]$$
$$M_{20} = \log 10\left(\frac{\sum_i^n M_i}{M_{tot}}\right), \quad \text{while} \sum_i f_i < 0.2 f_{tot} \tag{3.2}$$

where f_i is the flux in a pixel, x_c and y_c are the image centre and x_i and y_i are individual pixels. If a galaxy is highly centrally concentrated it will have low values of M_{20}, whereas if bright pixels are distributed through the galaxy, as is typically true in mergers, the value of M_{20} will be higher.

The asymmetry parameter measures the rotational symmetry of the flux distribution (Abraham et al., 1996). To calculate, the image is rotated by 180 or 90 degrees and subtracted from the original image, normalised such that a value of 1 implies total asymmetry and a value of 0 implies perfect symmetry (all pixels with equal flux, for instance):

$$A = \frac{\sum |I_0 - I_\theta|}{2 \sum I_0} - A_{bgr} \tag{3.3}$$

I_0 and I_θ represent the original and the rotated image respectively. A_{bgr} is the asymmetry of a background region of sky, accounting for asymmetry of noise. Galaxies tend to relax under gravity into symmetric shapes, so high asymmetry values are highly indicative of mergers or otherwise morphological disruption. Studies making use of these quantitative measures in the literature (e.g. Abraham et al., 1996; Lotz et al., 2004; Conselice, 2003) typically use combinations of parameters, identifying some quadrant of the phase space in which mergers are expected to lie (as per the function of the parameters, e.g. mergers are expected to have high asymmetry), and are typically typically found to lie by visual inspection.

The shape asymmetry (Pawlik et al., 2016) is additionally calculated here. Shape asymmetry has the same mathematical form as the rotational asymmetry (equation 3.3), but the image rotation and subtraction is performed on binary detection masks. Binary detection masks are computed using an 8-connected structure detection algorithm: the algorithm searches for neighbouring pixels to the central (brightest) pixel which have flux one standard deviation above the noise, and then repeats the process for those neighbours, snailing out to explore all galaxy pixels. All pixels with flux greater than the noise which are connected by an unbroken chain to the central pixel in this structure are assigned a value of 1, while all other pixels are set to 0. Rotation and subtraction from the original image, normalised by the number of pixels the galaxy occupies, gives a number between 0 and 1 giving an indication of morphological
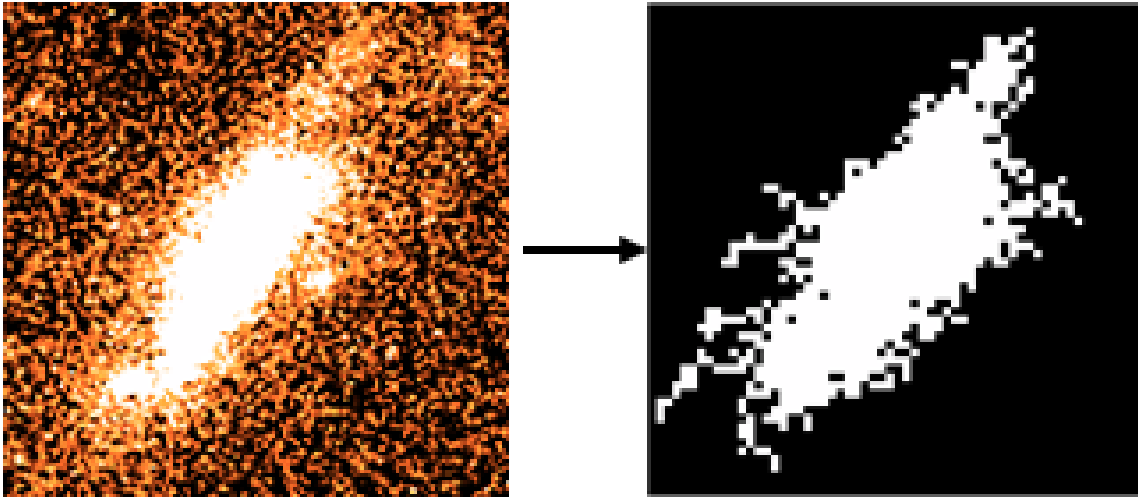
Figure 3.2: Example binary mask for an Illustris mock galaxy image.

asymmetry. Unlike classical asymmetry this has no dependence on the relative brightness of pixels, only the distribution of flux associated to the galaxy, i.e. it is a purely morphological measure. This makes shape asymmetry ideal for detecting low surface brightness features such as tidal tails, associated with galaxy mergers (Pawlik et al., 2016). Figure 3.2 shows an example binary detection mask for one of the Illustris galaxies.

3.2.2   Visual Classification - Zooniverse

A notorious issue with human visual classification of galaxies is its time consuming nature, an issue which automated measures attempt to alleviate. We use Zooniverse to simplify the process as far as possible. Zooniverse uses a user-friendly process where classifiers can click through images as prompted, requiring minimal effort on their part (Christian et al., 2012). To ensure good statistics without fatiguing quantities of work, only a subsample are visually classified here (148 major mergers and mergers, 53 minor mergers and 165 control galaxies) by 5 expert classifiers (researchers in galaxy evolution).

Classifiers were asked first to identify which features from the following list were present in a galaxy image: multiple cores, tidal tails, high asymmetry, expelled gas/stars, close neighbours, patchy light profile, train wreck merger, none of the above. As many or as few options could be selected as the classifier wished. Second, classifiers were asked to identify the galaxy as either a major merger, minor merger, otherwise disturbed, undisturbed or a bad image (unclassifiable). Asking the questions this way round was intended to make participants think through their reasoning before assigning some classification. Each classification was taken independently, with no requirement for agreement between classifiers, while classifiers gave the same galaxy the same classification ~60% of the time. The implied enhancement factors, F, presented in section 3.4 are dependent on the visual classifications of the classifiers used here. To avoid strong bias we used 5 independent classifiers and verified consistency by re-

moving the classifications of one visual classifier at a time, finding that the results presented in this chapter remain largely unaffected by individual bias (i.e. classifiers were in statistical agreement with one another).

3.3 Results

Generally, it is hard to quantify the effectiveness of structural parameters and visual classification in detecting mergers. Structural parameters are often compared to visual classification as a benchmark, but it is commonly unknown whether observers assign galaxies to the correct category, since there is no way of travelling to or rotating the galaxy to check the result. Here we make use of the fact that a simulation allows for precisely this analysis, since four viewing angles are available for each galaxy and the merger history of each galaxy is known ab initio. The efficacy of combining various structural parameters to identify mergers is analysed in section 3.3.1 and the ability of humans to recognise mergers is analysed in section 3.3.2. In both cases the dependence on galaxy masses, mass ratios, the time elapsed since mergers and star formation rate are studied.

Understanding the biases inherent in observations is critical for reasonable interpretation of astronomical data. In sections 3.3.3 and 3.3.4 the expected observational signatures from structural parameters and visual classification for a range of triggering models are created, specifically the inferred merger fractions from some true underlying distribution of mergers. Different triggering models assume mergers increase the duty cycle by some factor which is a free parameter. The inferred difference between "AGN-hosts" and control galaxies is highly dependent on the background merger rate (which varies across cosmic time) and, study to study, on the sample size. Otherwise stated: for some sample size, what underlying merger fraction and AGN duty cycle in mergers (relative to relaxed galaxies) are required for humans or automated methods to detect an over-density of mergers in the AGN sample? Comparison to appropriate studies in the literature (section 3.4) will allow for observational biases to be corrected for and true AGN merger fractions ascertained.

3.3.1 Efficacy of Automated Classification

We wished to use both structural parameters and visual classification to analyse the morphologies of the Illustris galaxies in order to compare to real observational studies, but first must quantify the efficacy of each in identifying mergers. We find that shape asymmetry (Pawlik et al., 2016) is the most effective structural parameter for robustly identifying mergers here. An overview of results from the structural parameters and comparison of identifying mergers with different parameters is presented below.

Gini and M_{20} are often used together to identify mergers (e.g. Lotz et al., 2004). Figure 3.3

shows the distribution of galaxies used here alongside the fractions of each which lie above the red line (Gini > (-0.115 × M20) + 0.384; Lotz et al., 2004), indicating a merger. It is found that only mergers with a mass ratio < 4:1 are identified as mergers more commonly than controls, with a high contamination rate (∼25% false positive rate). Perhaps surprisingly, this remains true even if only young mergers are considered (though ∼10% more major mergers are identified).

Comparison between asymmetry and shape asymmetry is shown in figure 3.4, and normalised histograms of the shape asymmetry for each galaxy category in figure 4.3. Galaxies with shape asymmetry > 0.2 are considered as merger candidates, as in Pawlik et al. (2016); we use the same criterion here (since both studies consider low redshift SDSS-like imaging the comparison is fair). Though similar trends are (unsurprisingly) observed in both shape asymmetry and asymmetry, with minor mergers indistinguishable from controls, shape asymmetry clearly has more success in identifying more significant mergers. Shape asymmetry finds ∼10% more mergers and major mergers than its less sophisticated counterpart and identifies mergers down to mass ratios of 10:1 almost as effectively as 1:1 mergers (figure 4.3). The false positive rate is significantly improved compared to Gini-M_{20}, dropping from ∼25% to ∼10%, though the asymmetry values for control galaxies are anomalously high. The right-hand panels of figure 4 show the efficacy of these measures as a function of time since the merger occurred. Asymmetry values drop rapidly as the time since the merger increases such that after 300Myr asymmetry fails to detect mergers more commonly than in the random galaxy sample. The same is true for shape asymmetry for all but the most major mergers, which remain detectable at a ∼30% detection level for at least 800Myr, the entire range probed here. Lower mass galaxies tend to have higher values of shape asymmetry, as demonstrated by figure 3.6, possibly due to the mass-dependence of morphological bipolarity.

3.3.2 Efficacy of Visual Classification

Visual classification, though subjective, remains perhaps the most effective technique for identifying mergers. 5 volunteers classified a subset of the total population described in section 3.1 and each classification is taken independently. Figure 3.7 demonstrates a low level of false positives, but also a low identification rate for mergers. If, instead of finding the fraction visually classified as major mergers we find the fraction classified as major or minor mergers, the fraction in each category increases by ∼10%, so including minor mergers is more akin to adding noise than helping to discriminate between mergers and non-mergers. Figure 3.8 shows the proficiency of humans at classifying mergers some time after coalescence, showing that humans are significantly better at detecting recent mergers than older mergers, with a strong dependency on mass ratio. Figure 3.9 shows the fraction of major and minor mergers which are identified as having some feature, prior to categorisation as either major or minor
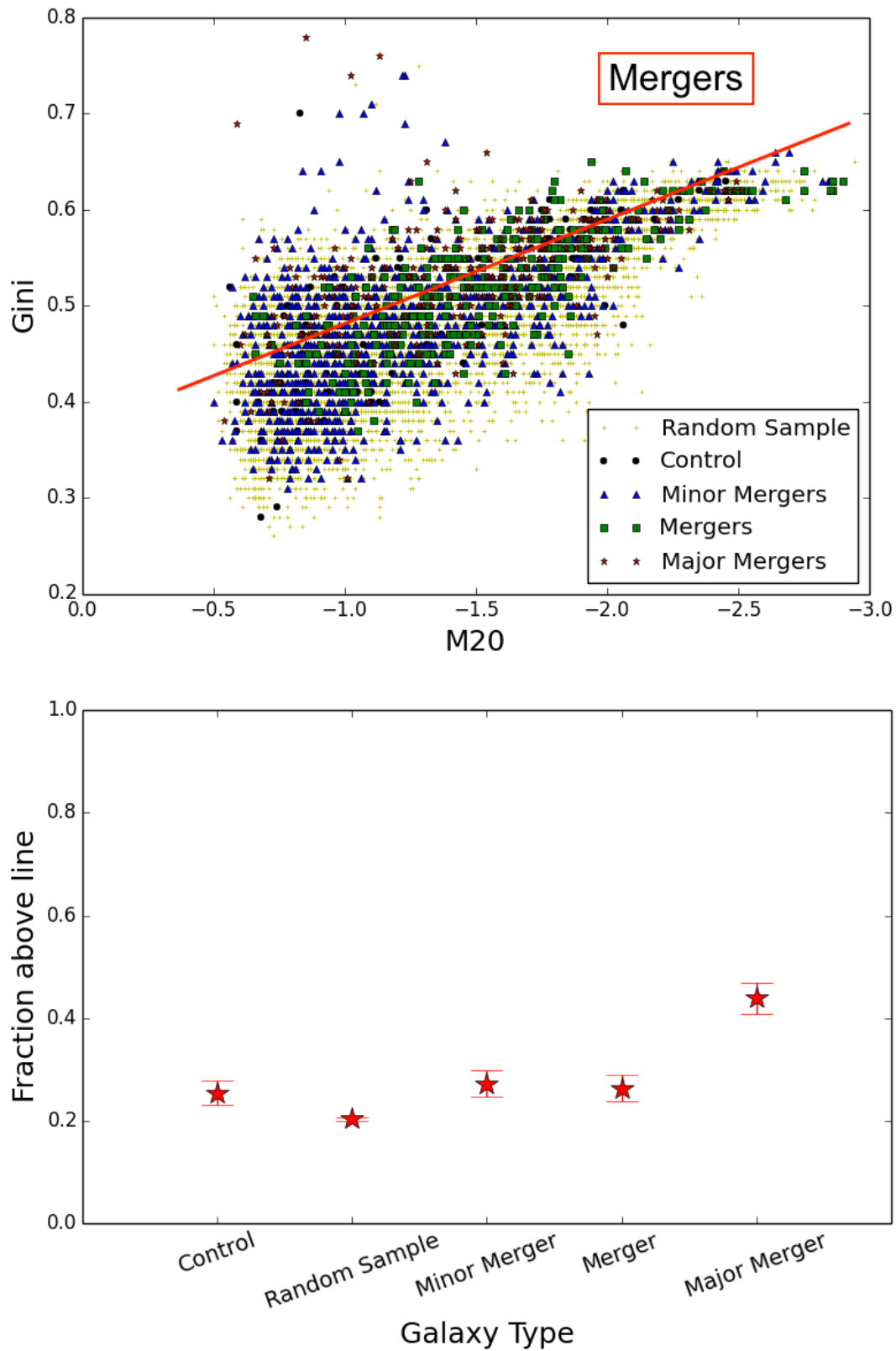
Figure 3.3: Left: the full distribution of Gini and M_{20} from all four viewing angles, split into merger categories (section 3.1). Points above the red line are commonly taken to be mergers (e.g. Lotz et al., 2004; Snyder et al., 2015). Right: the fraction of galaxies in each merger category found above the line. Errors are 1σ binomial errors (Cameron, 2011).
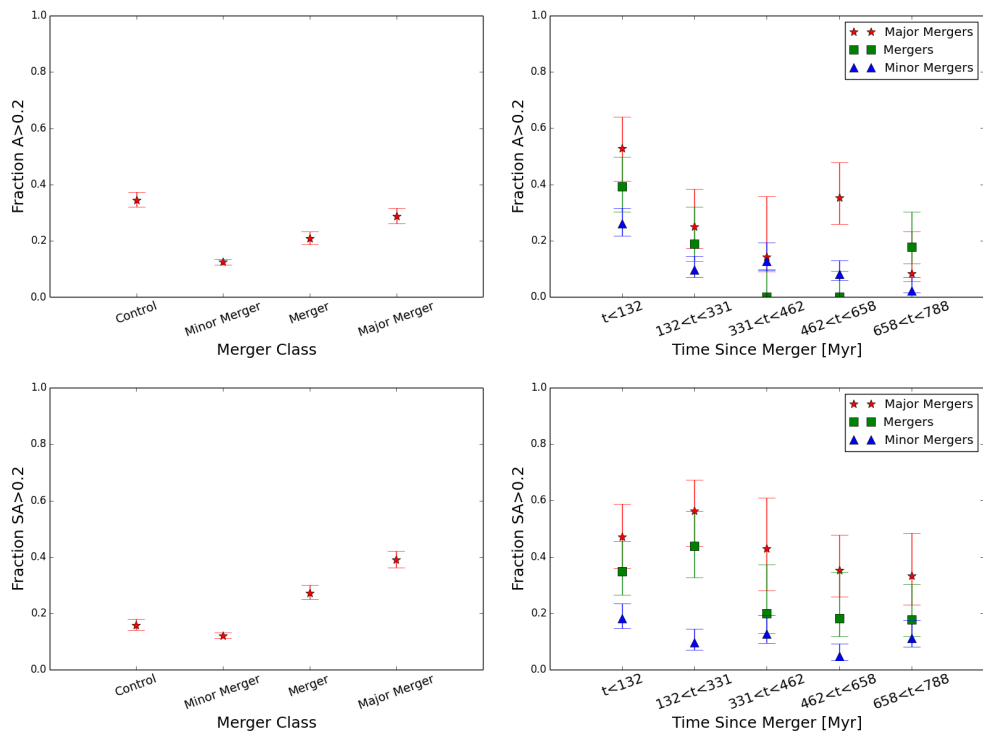
Figure 3.4: Upper and lower plots show the fraction in each category with asymmetry > 0.2 and shape asymmetry > 0.2 respectively. Left-hand figures are split by mass ratio as defined in section 3.1, right-hand figures show the fractions as a function of time since the merger for the three merger classes (red stars are major mergers, green squares mergers and blue triangles minor mergers). Errors are 1σ binomial errors (Cameron, 2011).
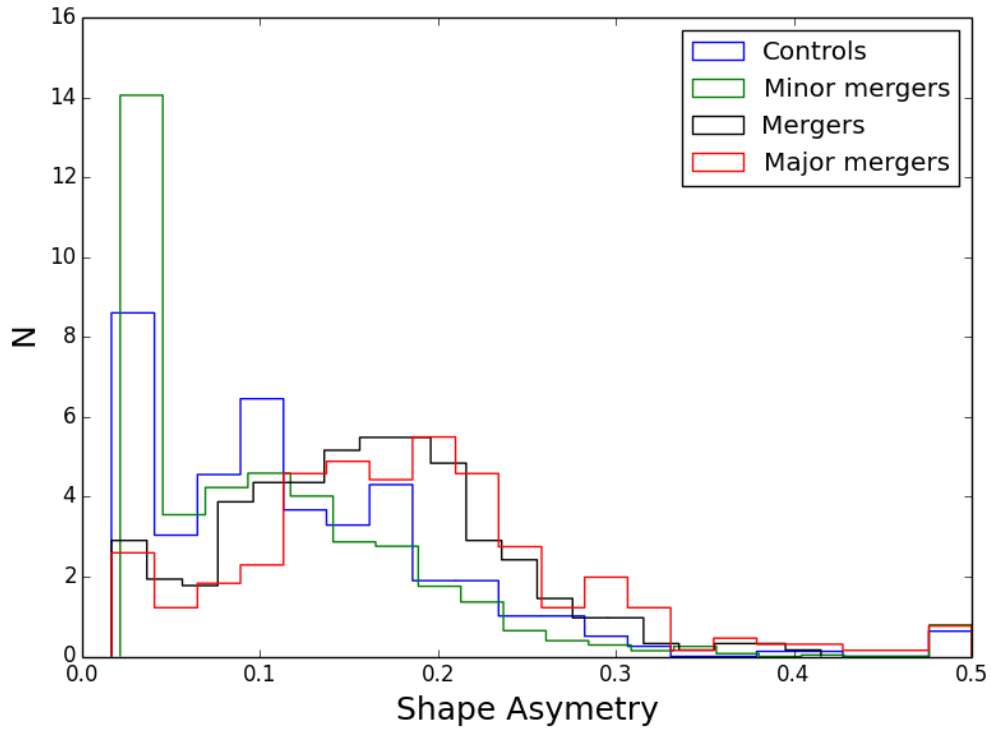
Figure 3.5: Normalised histograms of the shape asymmetry for each category of galaxy.
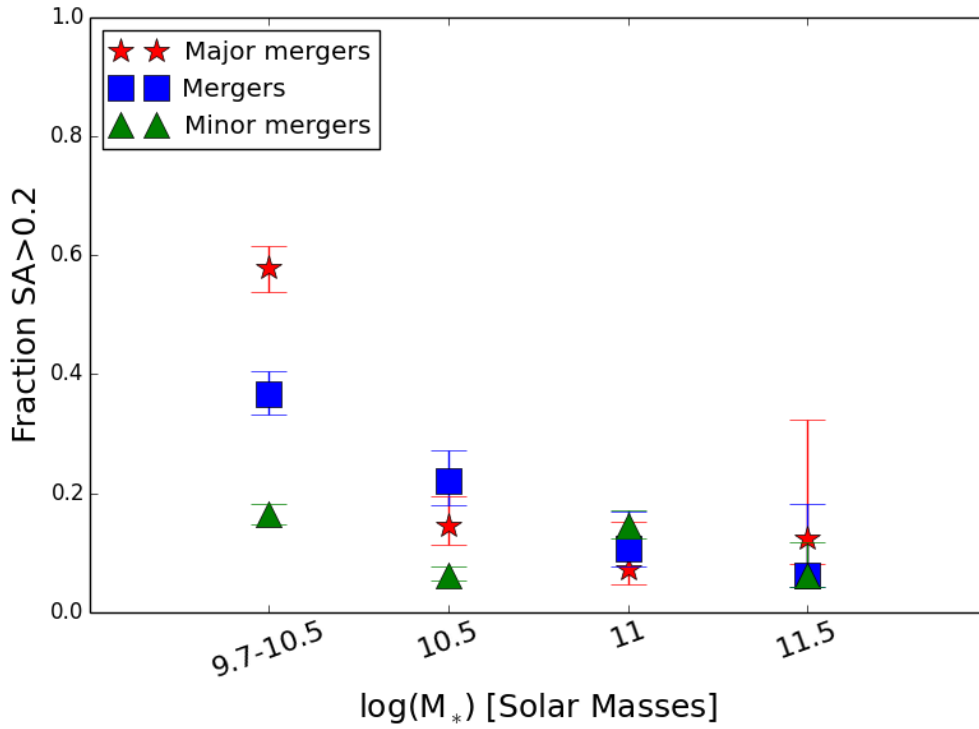
Figure 3.6: Fraction of galaxies with SA>0.2 with increasing stellar mass. Red stars are major mergers, blue squares mergers and green triangles minor mergers.

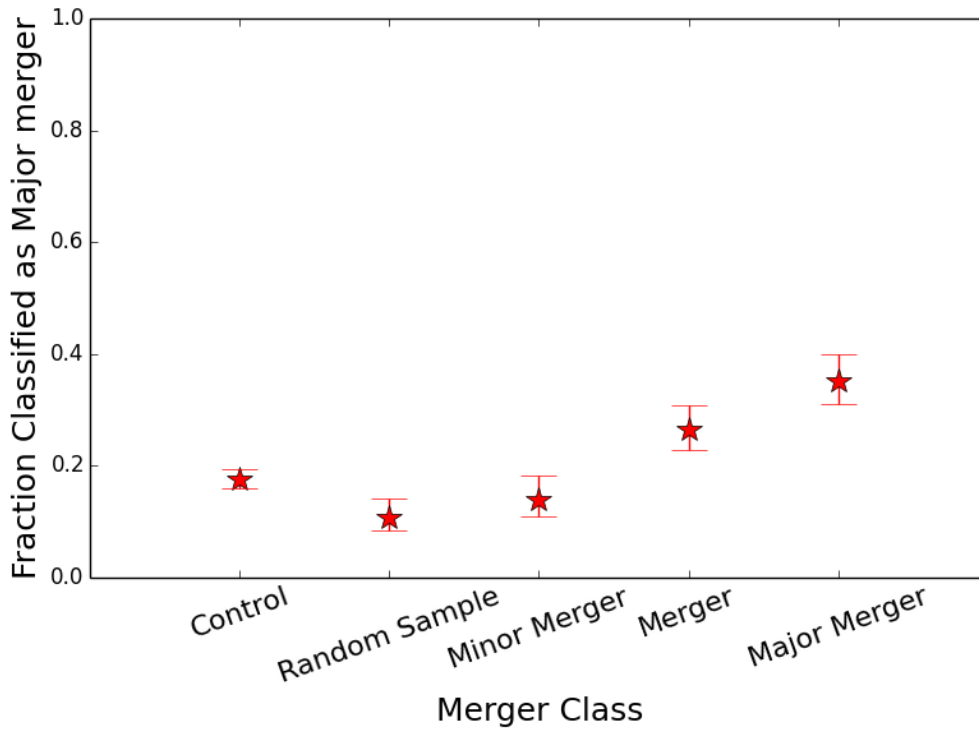**Figure 3.7:** Fractions of each category of galaxy classified as major mergers.

merger. Instead of major or minor merger categories one could use the perhaps less subjective descriptors of what is seen in an image, such as tidal tails, as indicative of some merger class. In reality it appears that major mergers, instead of having some tell-tail feature, simply have higher identified fractions of all individual features, excluding patchy light profiles. There is excellent agreement between visual classification and shape asymmetry (figure 3.10), with ~80% of galaxies with shape asymmetries > 0.3 visually classified as major mergers, implying that a combination of the two is a much better signifier of a merger than either alone. Figure 3.11 shows the fraction of galaxies classified as major mergers as a function of mass: all major mergers are correctly identified as such in the highest mass bin, otherwise with no obvious dependence on mass for classification.

3.3.3 Observational Signatures - Automated Measures

Here and in section 3.3.4 we attempt to quantify the extent to which galaxy mergers increase the likelihood of AGN activity in a given host. First we create fake samples of AGN hosts and control galaxies, as in most observational studies. The control samples are made up with some proportion of mergers to match the local universe merger fractions, i.e. ~2% major mergers, ~2% mergers and ~7% minor mergers, as estimated from the fractions involved in a merger in the random galaxy sample (section 3.1) and in agreement with Rodriguez-Gomez et al. (2015a). If mergers cause fuelling of AGN then, depending on how efficiently they do this, the
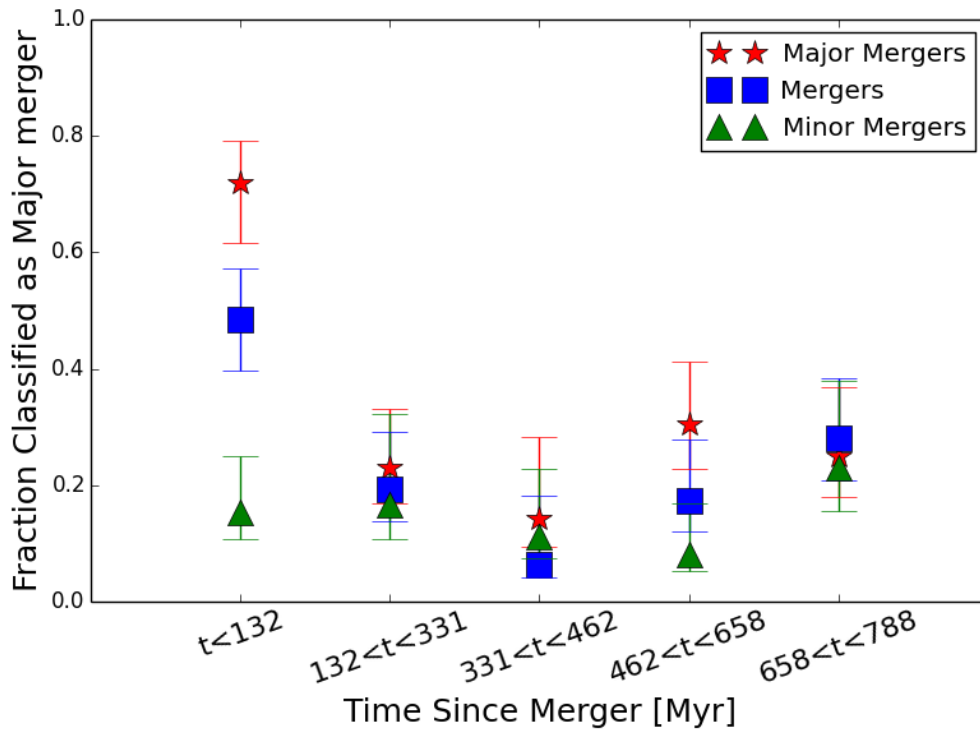
Figure 3.8: Fraction of major mergers, mergers and minor mergers classified as major mergers some time after the onset of the merger. Red stars are major mergers, blue squares mergers and green triangles minor mergers.
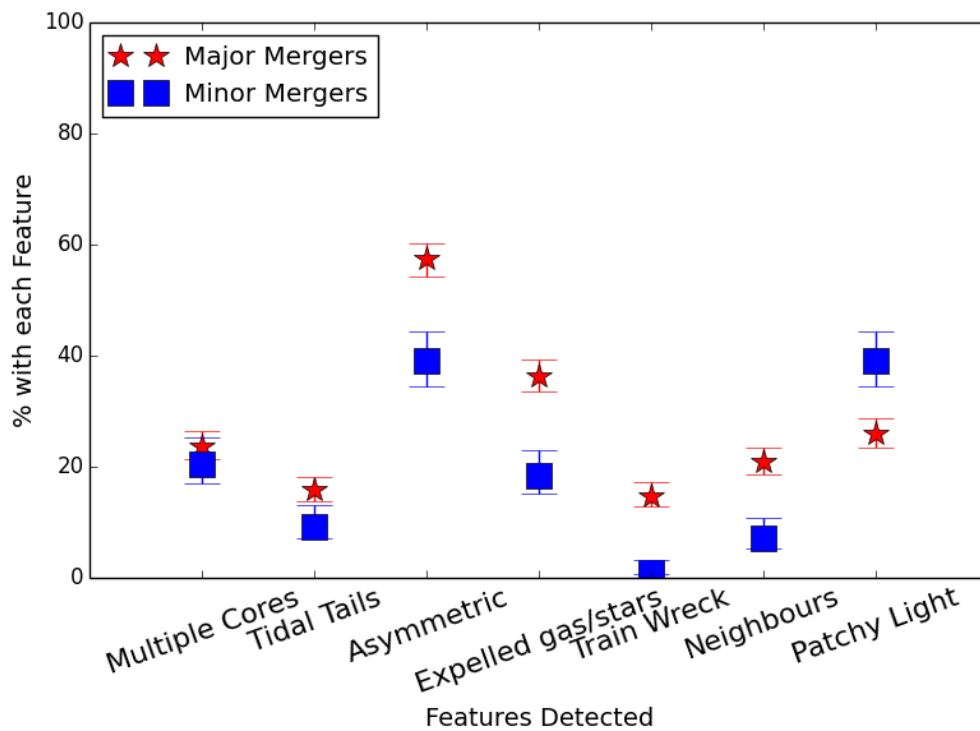

Figure 3.9: Percentage of major and minor mergers with different features identified by classifiers. Red stars are major mergers and blue squares minor mergers.
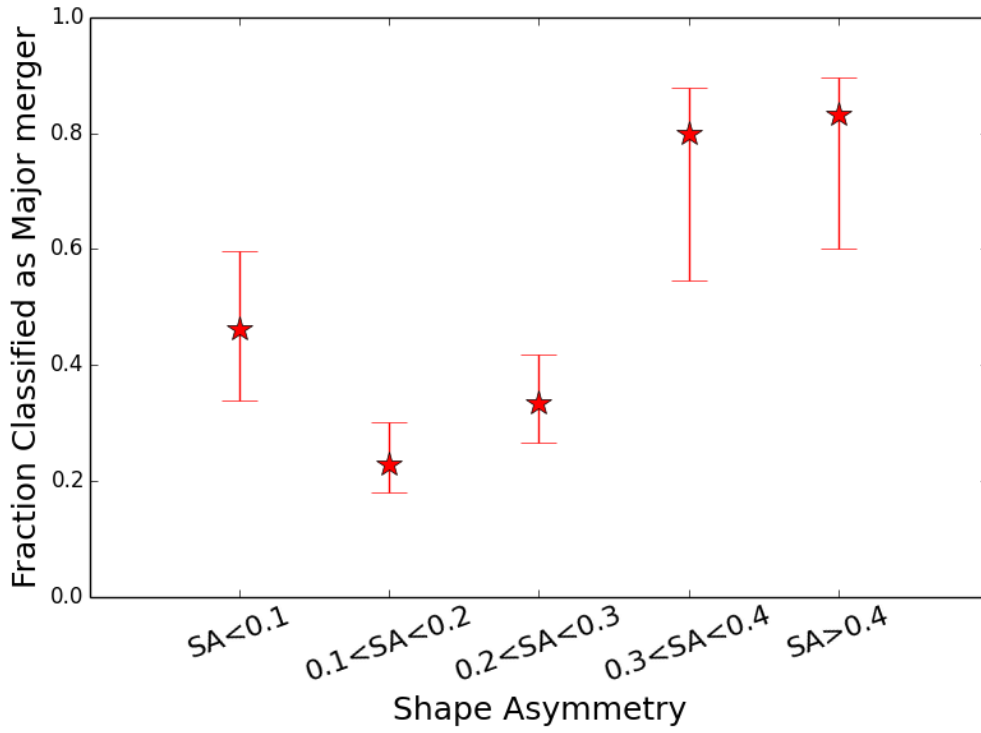
Figure 3.10: Fraction of major mergers classified as such in bins of shape asymmetry.
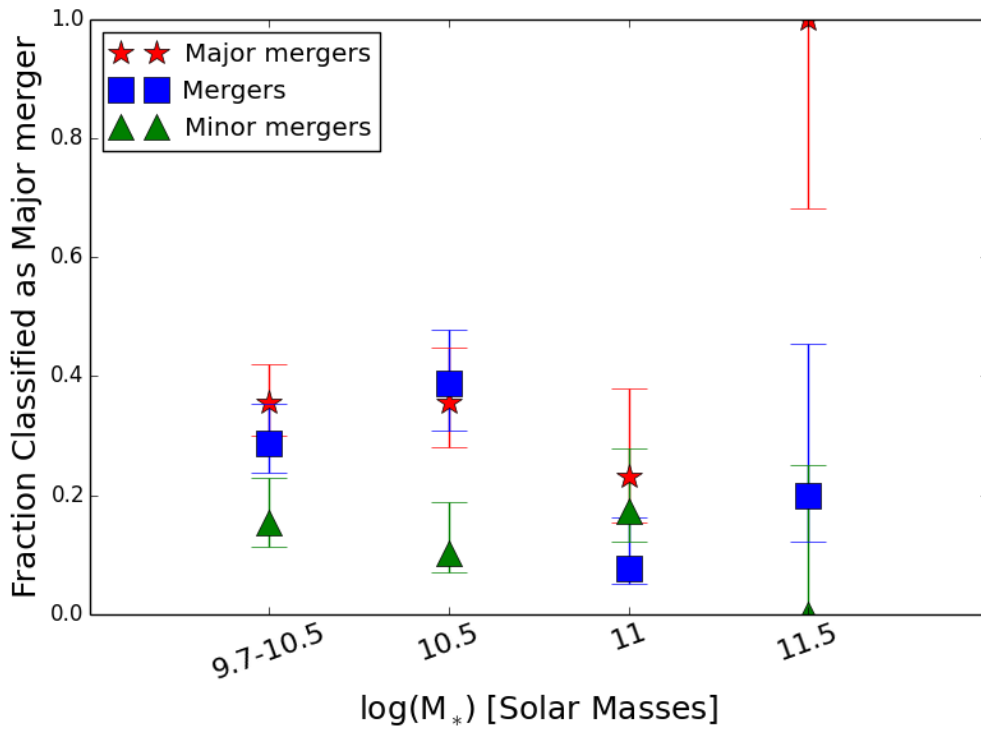

Figure 3.11: Fraction of galaxies classified as major mergers in bins of stellar mass. Red stars are major mergers, blue squares mergers and green triangles minor mergers.

AGN sample will contain some higher fraction of mergers than the control sample. To test how efficiently mergers may enhance accretion onto the AGN, we simply multiply the fraction of major mergers and mergers (i.e. mergers with mass ratios < 10:1) in the AGN sample by some enhancement factor (as seen in figures 3.12 - 3.15) and measure the morphological differences between the AGN and controls, as would be done in observational studies. Control galaxies are drawn with a fraction of mergers corresponding to the fraction in the local Universe (around 2% major mergers; Rodriguez-Gomez et al., 2015a) and the AGN sample with an enhancement factor F more mergers. By comparing the detected difference between the AGN and control samples at different enhancements to similar studies in the real universe we can constrain how efficiently mergers enhance accretion onto the central SMBH.

Samples of ∼50 AGN are drawn, with 3× as many mass-matched control galaxies (reasonably typical of literature studies, e.g. Kocevski et al., 2012; Hewlett et al., 2017). Note that the "AGN" in our samples bear no relation to the simulated AGN, as we wish to select an unbiased sample of galaxies to test a range of possible triggering scenarios. Random samples of AGN and control galaxies are drawn for some level of merger enhancement; the fraction of each group satisfying some criterion (such as having shape asymmetry > 0.2, or lying above the merger line in the Gini-M_{20} plot) is calculated; this process is repeated 50 times for each level of enhancement and the error from random sampling is encapsulated in the standard deviation of the distribution. Figure 3.12 shows the difference between AGN and control samples in the fractions of galaxies with shape asymmetry greater than 0.2, averaged over 50 random selections for a sample size of 50 AGN hosts and 150 control galaxies. Figure 3.12 shows how even with a 10× enhancement to the AGN duty cycle (i.e. with 40% of the AGN population comprised of mergers with mass ratio <10:1), AGN and control samples are only marginally distinguishable by their shape asymmetries, with ∼ 1σ confidence.

In chapter 2 it was found that there was an excess of AGN living in merging systems at z∼2. We hypothesised that a higher background merger rate with a constant triggering efficiency could yield such a result. Figure 3.13 tests this hypothesis by setting the background major merger rate to 10%, comparable to the z∼2 merger rate (Conselice et al., 2003; Rodriguez-Gomez et al., 2015a), and comparing the results to the local major merger fraction (2%). This does indeed yield a far greater difference between AGN and controls for modest enhancement factors, suggesting this could provide an explanation for the observed merger excess in chapter 2. However, a quantitative comparison to chapter 2 data is not possible here, since samples differ in image quality due to differences in instrumentation and redshift, and the effects of a varying star formation rate have not been considered (impacting the observability of merger features, see 3.3.4).
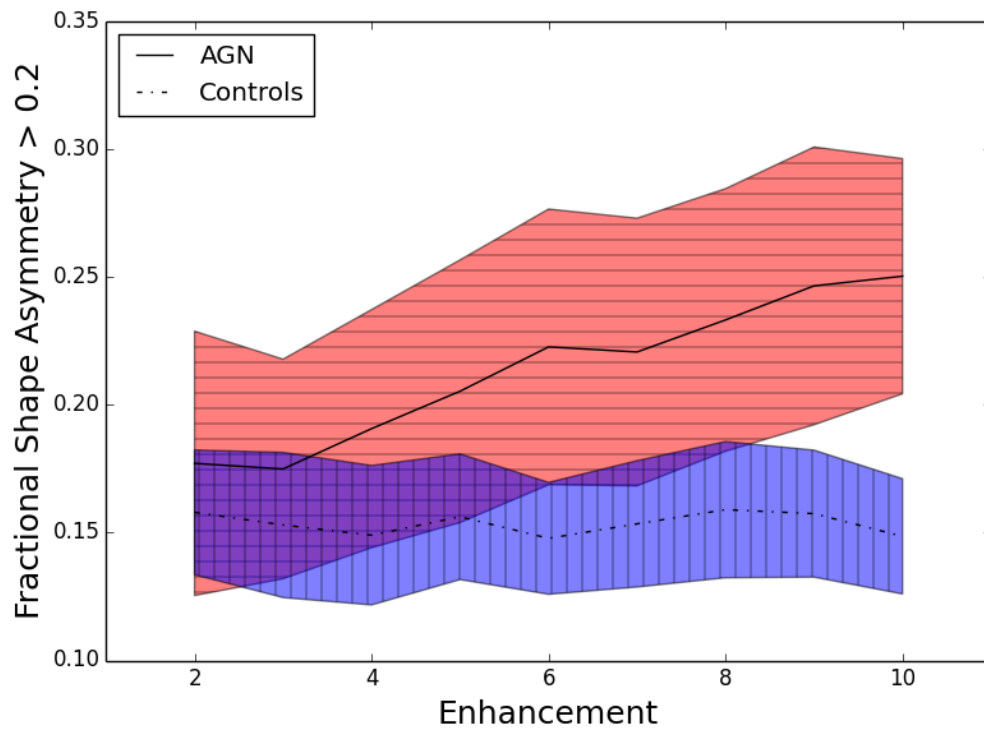
Figure 3.12: Fraction of galaxies with SA>0.2 for increasing values of the enhancement factor. Contemporary values for merger fraction were used (~2% major mergers, ~2% mergers and ~7% minor mergers). The fraction of galaxies with SA.0.2 are averaged over 50 randomly drawn galaxy samples for each level of enhancement and coloured bands represent the 1σ errors from the standard deviation of the distribution.
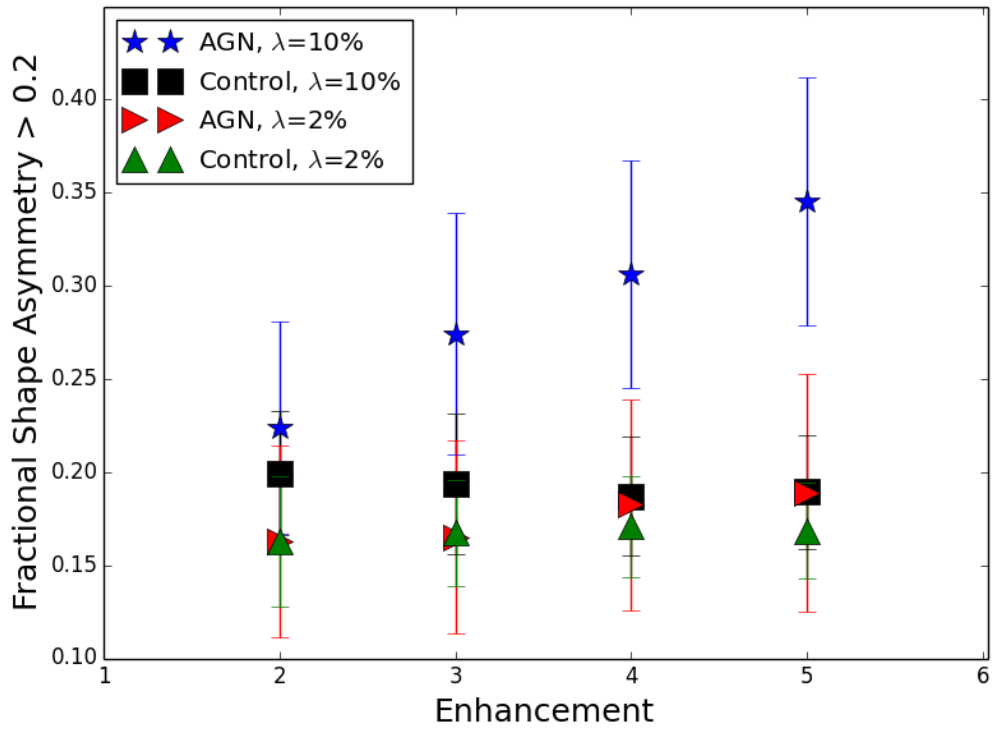
Figure 3.13: Fraction of galaxies with SA>0.2 for increasing values of the enhancement factor with a higher background rate of mergers, as is observed at high redshift. λ=~10% are major mergers, ~10% mergers and ~0% minor mergers (minor mergers are omitted since they are unobservable at high redshift, and because with these merger rates an enhancement factor of 5 leads to 100% of the AGN sample as mergers), in comparison to λ=~2% mergers and major mergers (as in the local Universe). SA fractions are averaged over 50 samples for each level of enhancement and 1σ errors are the standard deviation of the distribution.

3.3.4 Observational Signatures - Visual Classification

Here we present the expected detectable difference between control galaxies and AGN hosts for different levels of enhancement to the nuclear activity from mergers, as found by visual classification. As in the previous section, galaxies are drawn randomly but with the desired fractions of mergers to construct control and AGN samples. The fractions of controls and AGN that are identified visually as major mergers are then compared in the upper panel of figure 3.14 (the lower panel is discussed below). The fraction classified as major mergers is used throughout this section, instead of including those classified as minor mergers, but the results are qualitatively the same if minor merger classifications are included. It is found that even a relatively small enhancement to nuclear activity from mergers should be detectable, though due to the impact of false positives and false negatives on merger detection (false negatives are real mergers which go unidentified, false positives are non-mergers incorrectly identified as mergers), the relationship between the intrinsic enhancement and the detected enhancement is skewed. An intrinsic factor of 10 enhancement yields a factor of 4 detected enhancement in the merger fraction for this simplest model (in which all mergers in all galaxies trigger AGN with uniform probability). The difference between intrinsic and detected merger fractions will likely depend primarily on the depth of the particular survey used and the details of sample selection, in addition to the question asked of classifiers and their personal classification biases (not everyone agrees on what constitutes a merger: for instance, in this study there was ∼60% agreement in classification). In this study we will compare detected and intrinsic merger fractions to observational studies which use SDSS imaging, but in principle the approach taken here is more general, one merely needs to make simulation images resemble images from the particular survey with which they are to be compared (this is simple using SUNPY (Mumford et al., 2015)). By averaging over the classifications of multiple classifiers we mitigate the effects of individual classification bias.

It is clear how important sample selection of AGN and control galaxies is from inspection of figure 3.15, the same as figure 3.14 but split by stellar mass or specific star formation rate (sSFR) (upper and lower panels respectively). Often AGN hosts and control galaxies are matched in stellar mass in observational studies, but the more significant factor appears to be the sSFR, anti-correlating with stellar mass. Galaxies with low mass tend to have proportionately higher sSFRs; faint merger features such as tidal tails are constructed of stars and so the higher the sSFR, the less faint the merger features and the more easily galaxy mergers are identified. In fact, the probability of being classified as a merger correlates strongly with SFR for real mergers and non-mergers alike, as shown in figure 3.16. This causes low mass galaxies to be preferentially identified as mergers compared to high mass galaxies, presumably due to the anti-correlation between stellar mass and sSFR. AGN and control galaxies should

be matched in sSFR where possible, not just mass, since differences in sSFR seem to be the most severe in biasing detection of mergers.

It is unlikely that mergers have a uniform probability of triggering an AGN throughout the duration of the merger. Wild et al. (2010) found a ∼250Myr delay between a merger and the onset of nuclear activity, while theoretical work also suggests this as likely (e.g. Di Matteo et al., 2005; Hopkins & Hernquist, 2006). To recreate this here we selected mergers in the AGN sample with a gaussian probability distribution centred 250Myr after the onset of the merger, with a range of full width at half maxima. In this study, mergers are detected equally well at all times since the merger after the first snapshot (in which the detection probability is enhanced), so this does not have a significant influence on the results. To set a bound on the influence such an effect could have on detection (i.e. to maximise detection of mergers) we instead set the centre of the probability distribution to the onset of the merger, at t=0Myr. The impact of this is shown in the lower panel of figure 3.14, significantly increasing the detected merger fraction, boosting the merger excess in the AGN sample by a factor of ∼2 relative to the upper panel.

3.4 Discussion

In the previous section we described the comparison between mock samples of potential AGN hosts and control galaxies from Illustris with different assumed merger-triggering enhancement to the AGN duty cycle. SDSS-like images have been constructed and analysed both by visual classification and structural parameters (Gini, M20, asymmetry, shape asymmetry). It was found that the abundance of false positive and false negative classifications of mergers adds considerable noise when determining a merger fraction in some galaxy sample and leads to significant underestimation of that merger fraction. The rate of false positives and negatives correlate strongly with other galaxy properties such as the SFR, the duration of time elapsed since a merger, the mass ratio of that merger and the galaxy mass. In this section we discuss these results with regards to the wider literature, in particular to interpret the range of possible merger-enhancements to the AGN duty cycle.

First we discuss the efficacy of different morphology measures in identifying mergers. In section 3.3.1 Gini-M_{20} was shown to be only partially effective at identifying major mergers in Illustris, at a level around 10-20% above controls (for all and recent major mergers respectively), and completely unable to identify mergers with mass ratios greater than 4:1. This may be partially influenced by the fact that ∼10% of galaxies with $M_* \sim 10^{10.5 \text{ - } 11} M_\odot$ have unphysical ring-like structures (Snyder et al., 2015), which may boost values of Gini and M20 due to the higher surface brightness far from the galaxy centre, but this is clearly a subdominant factor. A reasonable interpretation is that too many non-merger processes are able to influence
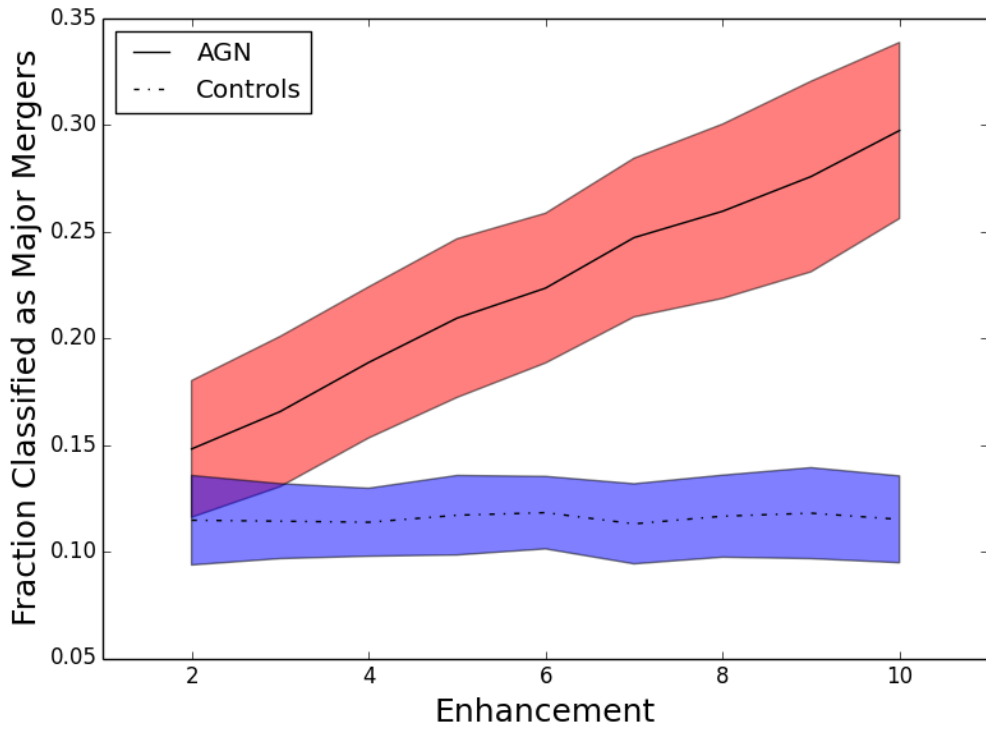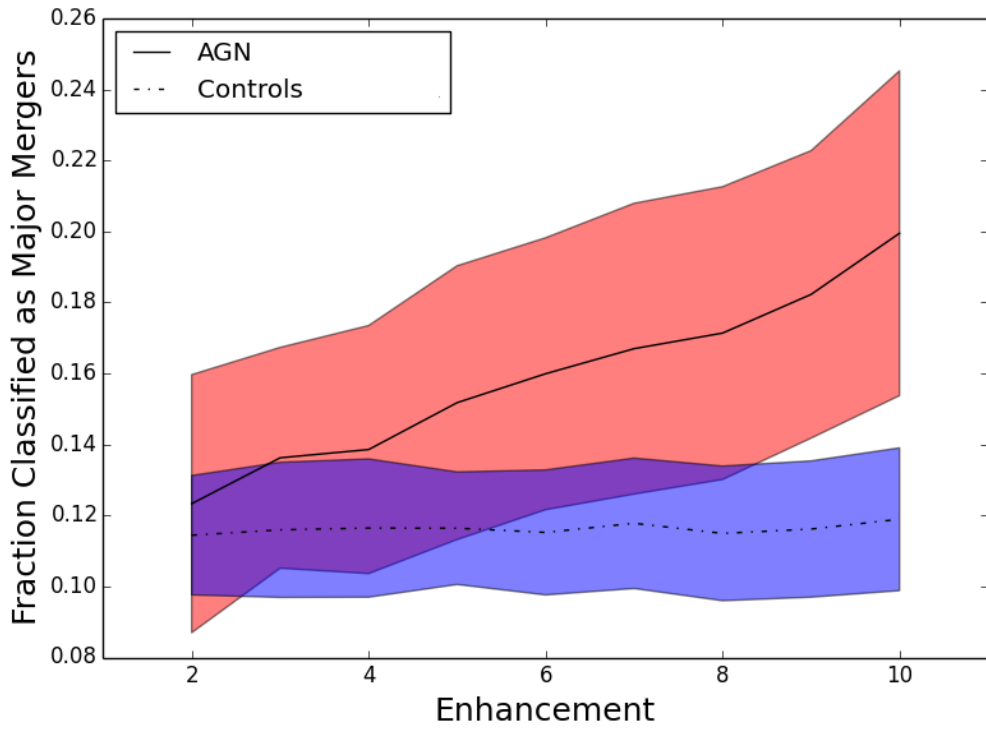
Figure 3.14: Fraction of galaxies visually classified as major mergers for increasing values of the enhancement factor. The upper plot is for a uniform enhancement to the AGN duty cycle over an 800Myr period after the merger. Coloured bands represent the 1σ errors; red for AGN, blue for controls. The lower plot is for a gaussian probability of enhancement about the onset of the merger with a standard deviation of 100Myr. Contemporary values for merger fraction were used ($\sim$2% major mergers, $\sim$2% mergers and $\sim$7% minor mergers). Classification fractions are averaged over 50 samples for each level of enhancement and 1σ errors are the standard deviation of the distribution.

Figure 3.15: Fraction of galaxies visually classified as major mergers for increasing values of the enhancement factor, split by stellar mass (upper) and specific star formation rate (lower). Coloured bands represent the 1σ errors; blue for high mass/sSFR AGN, green for low mass/sSFR AGN, red for low mass/sSFR controls and yellow for high mass/sSFR controls. Contemporary values for merger fraction were used ($\sim$2% major mergers, $\sim$2% mergers and $\sim$7% minor mergers). Classification fractions are averaged over 50 samples for each level of enhancement and 1σ errors are the standard deviation of the distribution.
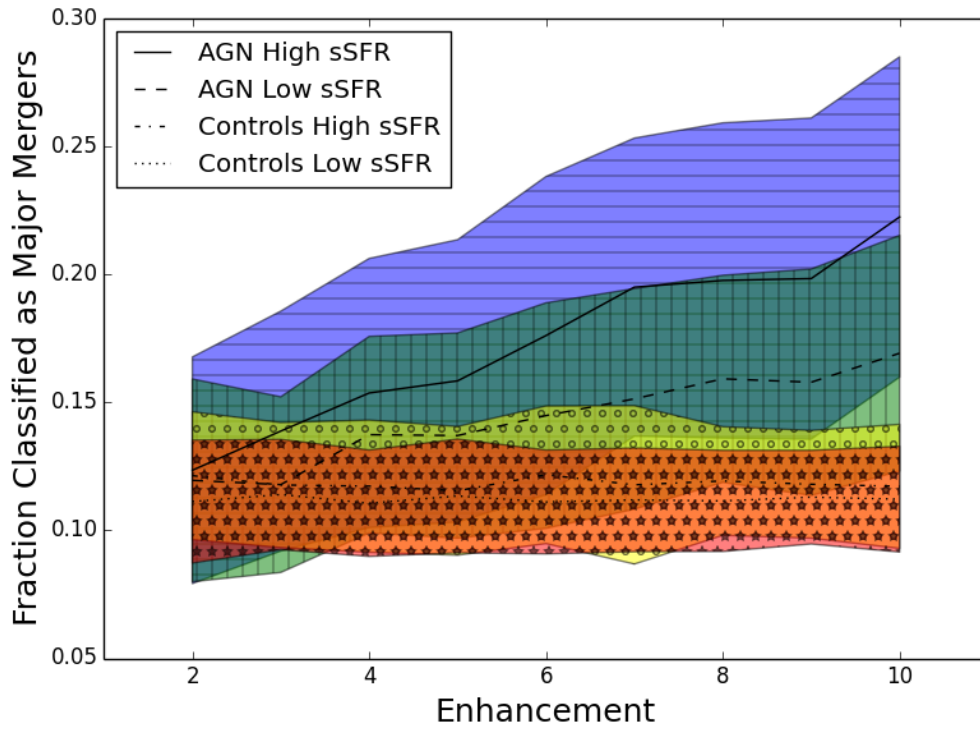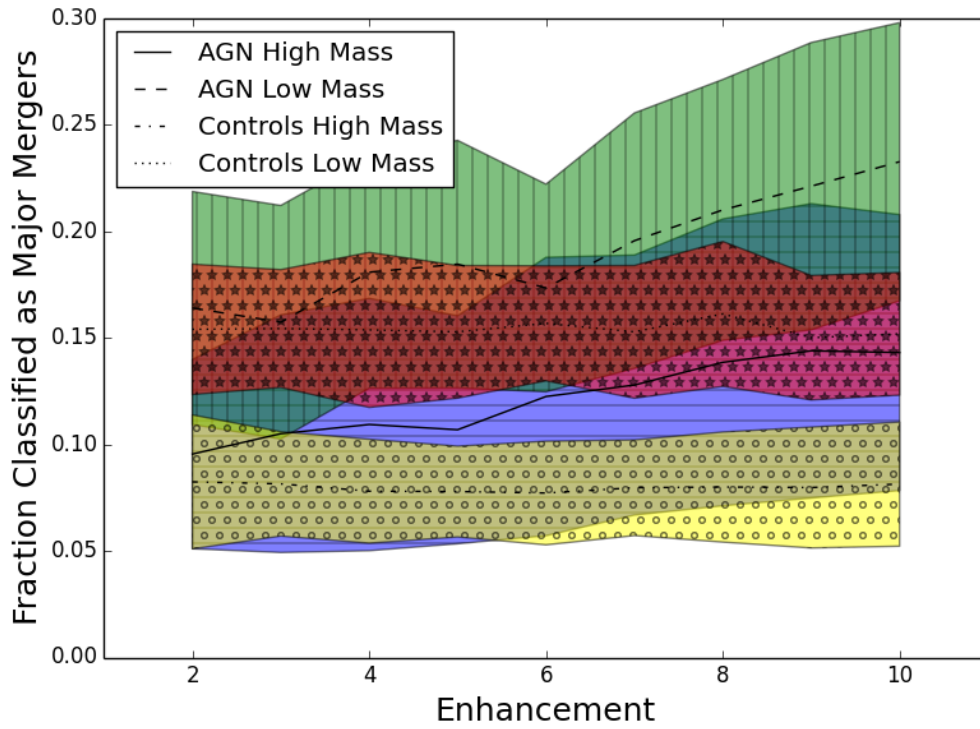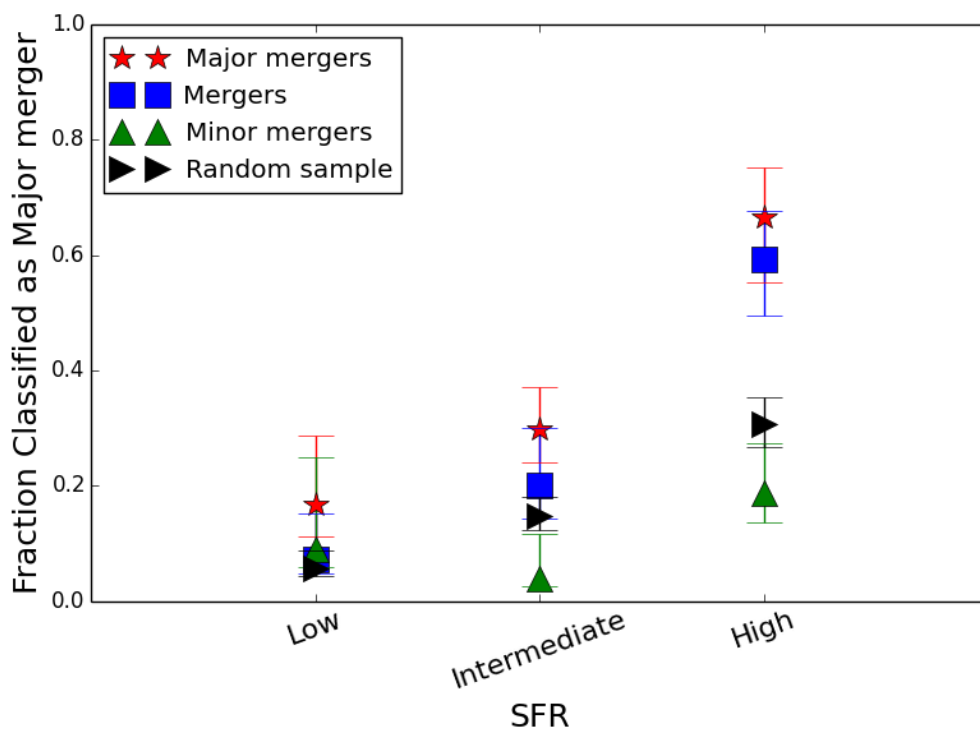
Figure 3.16: Fraction of galaxies visually classified as major mergers in bins of increasing SFR. The mean SFR in each bin is 0.9, 3.3 and 8.7 solar masses per year for low, intermediate and high SFRs respectively.

the values of Gini/M20, such that its usefulness in identifying mergers is limited. For example, the difference between galaxies with low or high SFRs in Gini-M20 space was found to be at least as pronounced as the difference between mergers and non-mergers.

Shape asymmetry (Pawlik et al., 2016) was found to be the most effective parameter for identifying mergers, successfully picking out major mergers for at least 800Myr. Visually, major mergers are the most easily detected, with a detection probability steeply declining with time since the merger. A combination of visual classification and shape asymmetry identifies ∼80% of major mergers within 800Myr. By visual classification there is no strong trend between galaxy mass and likelihood of identifying a merger, except in the highest mass bin where 100% of major mergers are identified. Conversely, mergers of low mass galaxies are considerably more easily distinguished by shape asymmetry, which could be due to the difference in typical initial morphologies between early- and late-type galaxies. Visual classification is not an objective technique and there will be some variance study to study. This depends on the questions asked of classifiers, generally with simpler questions leading to more robust answers (e.g. Darg et al., 2010b), the experience or expertise of the classifiers and the number of classifiers per galaxy. By asking a relatively simple question (does this image contain a major merger, minor merger or non-merger?) of 5 expert classifiers we hope to have minimised the variance of responses and to have defined a classification scheme comparable to other observational studies. Comparison to studies making use of galaxy zoo (with a large number of inexpert classifiers) or studies in which galaxy pairs are counted to infer a merger rate (where the question asked of classifiers is significantly different) will be imperfect. We attempted to use a large group of non-specialist visual classifiers (consisting of ∼20 students) to classify the galaxies, which would have more closely emulated the galaxy zoo process, however the statistics from these classifications were insufficient for further analysis.

Dickinson et al. (2018) compare the classification of Illustris galaxies with SDSS galaxies using galaxy zoo and find that significantly more substructure is identified in the simulated galaxies, particularly for lower mass galaxies, likely due to the inadequate resolution of the simulation. This could bias the distributions of structural parameters, particularly Gini, M20 and asymmetry, if the surface brightness of substructures is significant. Similarly, if human classifiers attribute substructure to galaxy interactions this will increase the detected merger fractions in this work, while correlations between substructure and stellar mass could contribute to the observed difference in the number of detections of mergers in low- and high-mass galaxies (see figure 3.15). However, the substructures identified in Dickinson et al. (2018) pertain to the smoothness of a galaxy, while the features which best predict classification as a major merger here relatively unambiguously pertain to morphological disruption (e.g. asymmetry). The only visually classified feature pertaining to smoothness in this work, "patchy light profiles", was found to anti-correlate with classification as a major merger (see figure

3.9), suggesting the additional substructure does not significantly influence the likelihood of being visually classified as a merger. In addition, the quantity of interest here is the observed ratios of merger fractions in different galaxy samples; since both samples would have similarly enhanced detected merger fractions the ratio should remain unaffected, and comparisons to merger excesses in the literature robust.

Next we discuss the intrinsic merger excess implied by some observed excess as a function of galaxy properties, and compare to observations in the literature. Mock images of potential AGN host galaxies have been assembled, samples constructed with some fraction of mergers corresponding to an inferred enhancement to the AGN duty cycle from galaxy mergers, and compared to random galaxy samples. Although an AGN may "flicker" throughout its lifetime, the relevant quantity is the fraction of some period of time for which the galaxy is active: this is encapsulated in the duty cycle. Control galaxies are taken quasi-randomly from the Illustris simulation and so contain the normal fraction of mergers as in the simulation. Comparing these two samples it is found that even when the merger fraction (mass ratios <10:1) is increased in the AGN sample by a factor of 10, and only recent mergers (most easily detected) are assumed to enhance the AGN duty cycle, this results in a factor of ∼5 detected merger excess by visual classification. This represents the upper bound on what could be detected and is only half the intrinsic excess (figure 3.14). If mergers are assumed to trigger AGN equally efficiently at all times since the merger the situation worsens, with a detected merger excess around a third of the true excess, effectively obscuring the role of mergers in AGN triggering. At higher redshifts the intrinsic merger fraction is higher, and so any enhancement to nuclear activity as a consequence of mergers would rapidly increase the detectable morphology difference between AGN hosts and control galaxies, possibly helping to explain the results of Hewlett et al. (2017) where AGN are found in mergers above the control level at high redshift only.

Figure 3.15 demonstrates the effect a mass or SFR mismatch between AGN and control samples would have on detected merger fractions. Mergers with high mass or low SFR are hard to distinguish from non-mergers with low mass or high SFR respectively. Similar trends are seen with increasing gas fraction and increasing SFR, presumably due to correlations between SFR and gas fraction (and stellar mass etc.). The stronger trend is with SFR however, and this trend persists even if gas fraction or gas mass is controlled for, so we focus on that in the analysis here. If AGN hosts are improperly matched (in mass or SFR) to control galaxies real trends could be hidden in the data. Worse still, the SFR is expected to be enhanced early in a merger (e.g. Di Matteo et al., 2005) before declining rapidly. If an AGN is triggered some time after the onset of the merger, as suggested by both theoretical work and observational studies (e.g. Wild et al., 2010; Carpineti et al., 2012; Kaviraj et al., 2015; Shabala et al., 2017; Di Matteo et al., 2005; Hopkins & Hernquist, 2006; Hopkins, 2012), then while the merger is most easily detected the AGN is likely to be inactive. Mergers are most easily

detected soon after onset both because the merger features have not had time to fade and because the SFR is enhanced, illuminating those merger features. The SFR then rapidly declines and merger features fade as the luminous massive young stars die; the AGN is triggered some time later but the merger associated with it is likely undetectable. As can be seen from figure 3.16 a factor of ∼10 increase in SFR corresponds to a factor ∼5 increase in identification as a major merger. SFR can easily rise by 1-2 orders of magnitude after a merger, rapidly declining ($t_{decline}$<100Myr; Di Matteo et al., 2005; Di Matteo et al., 2007) before the gas has had time to funnel to the SMBH, so the number of non-AGN mergers (assuming a delay between merger and nuclear activity) could itself be boosted by some factor, additionally hiding any AGN excess that may exist. The longer the time delay between merger-induced starburst and nuclear activity the harder it is to observe mergers and the larger the real merger-excess compared to that implied by data. This confluence of events would extremely effectively hide a merger-AGN connection, driving down the observed merger excess in AGN host samples, and is schematically illustrated in figure 3.17.

Though merger-AGN connections may be hidden by time delays between coalescence and black hole growth, there is some evidence that there exist global relations between star formation and nuclear activity. AGN (excluding the highest mass SMBHs in the densest environments, which are often radio loud (e.g. Best et al., 2005) and may be fuelled by, for example, cooling hot halo gas) and star formation are both fuelled by a supply of cold gas, and so it is perhaps unsurprising that global relations between star formation and nuclear activity have been identified in e.g. Mullaney et al. (2012); Sabater et al. (2013); Azadi et al. (2014); Sabater et al. (2015); Ellison et al. (2016). However, AGN vary on much shorter timescales than the SFR and so the correlations are small, debatable, and do not apply to all AGN (e.g. the probability of hosting an AGN broadly rises with galaxy mass, which anti-correlates with SFR Kauffmann et al., 2003; Gavazzi et al., 2011). Since increased SFR enhances the probability of identifying a merger, even if there is no merger to identify (see figure 3.16), we consider the possibility that this may cause a fallacious detection of a merger-AGN connection, which is really just the aforementioned SFR-AGN connection. Sabater et al. (2015) find that the effect of mergers on nuclear activity, once stellar mass and nuclear SFRs are controlled for, is negligible. In this study, they do not use the observed morphology to identify mergers but measures of the galaxy densities and tidal forces due to neighbouring galaxies to determine the prevalence of galaxy interactions, and as such their results are robust against the considerations of this chapter. Azadi et al. (2014) find that star forming galaxies are factor of 2-3 more likely to host an AGN relative to quiescent galaxies. The star forming galaxies have mean SFRs ∼ 10× higher than the quiescent galaxies; such a factor would increase the probability of detecting a merger by a factor of ∼5. Despite this, the SFRs of differently selected AGN vary greatly (e.g. SFRs are shifted to lower values for low excitation radio galaxies, while IR

selected AGN have higher SFRs in Ellison et al., 2016, for example) and so in some samples this could bias against detecting mergers. Overall, given the large range of physical environments AGN inhabit (Mullaney et al., 2012), it is unlikely that a global SFR-AGN connection will significantly influence the detection of a merger-AGN connection. However, these results have more far-reaching consequences than for studies of AGN alone. For instance, in Darg et al. (2010b,a) they find that spiral galaxies are twice as likely to be identified with mergers as ellipticals in galaxy zoo, although major merger rates increase mildly with increasing stellar mass (e.g. Rodriguez-Gomez et al., 2015a); this can be understood as due to different average SFRs of spirals and ellipticals causing easier identification of merging spirals.

The star formation rate density is a factor of ∼2 too high at low redshift in Illustris (Genel et al., 2014; Vogelsberger et al., 2014a; Snyder et al., 2015), and since higher SFRs increase the likelihood of identifying a merger (even for non-interacting galaxies) one might worry that this could influence the analysis here. However, since the difference between SFR in Illustris and in the Universe differs by a factor of $\lesssim 2$ and since both control and AGN samples will be similarly enhanced, the effect is qualitatively negligible.

We will now investigate the merger excesses in AGN samples relative to controls (or lack thereof) reported in the literature in observational studies and reinterpret results based on the analysis presented here. Different studies have a range of sensitivities and attention to detail; here, we limit detailed analysis to low redshift studies with SDSS imaging to ensure a like-for-like comparison with this work is legitimate. Koss et al. (2011) make use of galaxy zoo classifications, comparing merger rates of 185 low redshift (z<0.05) hard X-ray selected AGN to matched controls. They find a factor of ∼4 enhancement of AGN associated with galaxy mergers relative to controls. To find the minimal true merger excess implied by this measurement we can assume only recent major mergers trigger AGN, as in the lower panel of figure 3.14, such that mergers are as detectable as possible. From inspection of figure 3.14 it can be seen that the line representing AGN only reaches an observed factor of 4 enhancement over controls when the true merger excess is a factor of ∼15 larger, with an associated merger fraction making up around 30% of the AGN population (assuming a background merger rate of ∼2%). I.e. approximately a third of the AGN in Koss et al. (2011) are likely to be merger triggered, as opposed to the <1/10th reported. If mergers are assumed to trigger AGN equally efficiently for ∼1Gyr since the merger, the implied merger enhancement to the AGN duty cycle in Koss et al. (2011) is a factor of ∼32 (deduced from extrapolation of the upper panel of figure 3.14), implying mergers account for around two thirds of the AGN triggering in this sample. Assuming some delay between merger and nuclear activity boosts the implied merger fraction, hence merger-enhancement to the duty cycle. Assuming a delay of 250Myr (Wild et al., 2010) gives approximately the same result as assuming all mergers are equally likely to trigger AGN; assuming longer delays does not significantly alter this result, so we take the former as the

upper limit on merger-excess.

Cotini et al. (2013) use visual classification and a combination of structural parameters to determine the merger fractions in a sample of ∼60 hard X-ray selected AGN and ∼250 matched controls, finding an approximately 5× merger excess in the AGN sample (∼20% vs. ∼4%). As above, under the assumption that only recent mergers trigger AGN it is possible to set a lower limit on the likely intrinsic merger excess, equal to around a factor of 21 with an associated merger fraction of ∼42%. If mergers can trigger AGN within ∼800Myr then the approximate upper limit on the merger-excess is a factor of ∼43, with mergers comprising around 86% of the sample. Similarly, Kaviraj et al. (2015) visually classify low redshift (z<0.3) early type radio AGN hosts and compare them to matched controls, finding a factor of 3.3 or 4.7 AGN-merger-excess in the field and in clusters respectively. This implies a true merger excess in the range 11-24 in the field and 20-40 in clusters, corresponding to merger fractions in the AGN sample between 22% and 80%. These results, summarised in table 3.1, taken together imply that around half of all the AGN in the local Universe are merger triggered.

Ellison et al. (2013) and Satyapal et al. (2014) invert the question: instead of finding AGN and determining the fraction that show signs of mergers they find post-merger galaxies and determine the fraction of AGN within them, relative to close galaxy pairs and isolated galaxies. Mergers in these studies are identified from galaxy zoo (Darg et al., 2010b) with an additional stage of visual classification by Ellison et al. (2013). This altered methodology makes direct comparison to merger enhancement factors in this chapter more tricky, since by design the merger fraction is 100%. However, inspection of figure 3.7 suggests that true major mergers are only identified ≲40% of the time, while ∼15% of non-mergers are mis-identified as mergers. Although the merger selection in Ellison et al. (2013) is more rigorous than conducted here, with two layers of merger selection, some non-mergers may have infiltrated the post-merger sample. Ellison et al. (2013) selected AGN from BPT emission line diagnostics and find a factor of ∼4 enhancement of the local (z<0.1) AGN fraction in the post-merger sample relative to the AGN fraction in non-mergers. For the same sample, Satyapal et al. (2014) find a factor of 10-20 enhancement, selecting AGN using WISE IR photometry (as in Stern et al., 2012) (a factor of 10 for W1-W2≥0.5; a factor of 20 for the more robust cut for AGN detection W1-W2≥0.8), implying high levels of obscuration in the post-merger sample. If some of the mergers in the post-merger sample are imposters (non-mergers) the true enhancements will be greater still. This suggests the enhancement to the AGN duty cycle due to mergers is ≳20, i.e. ≳40% of AGN are merger triggered (assuming again a 2% major merger fraction), consistent with the enhancements we infer from other studies in table 3.1.

Though direct comparison to the work here is not possible, we can consider the results of studies with different imaging depths and techniques. Using very high depth imaging (sensi-

Study	Detected excess	Implied excess lower limit	Implied excess upper limit	AGN fraction in mergers
(Koss et al., 2011)	~4	15	32	30-64%
(Cotini et al., 2013)	5	21	43	42-86%
(Kaviraj et al., 2015)	3.3-4.7	11 or 20	24 or 40	22-80%
Combined Mean	4.3	16.8	34.8	51.3±20.9%

Table 3.1: Table of detected merger excesses in AGN samples from the literature and the true merger excesses implied due to the observational biases inferred from analysis of Illustris mock images and outlined in this chapter. The AGN fraction in mergers are calculated assuming a 2% background major merger rate and the error on the mean is the 1σ standard deviation.

tive to surface brightnesses ~26 mag arcsec^{-1}) and careful analysis of the surface brightness of features, Ramos Almeida et al. (2011) investigate the role of interactions in driving powerful radio AGN. They find that 60% of radio galaxies in the local Universe show signs of recent galaxy interactions, as signified by the presence of tidal tails, bridges, fans, shells, amorphous halos, multiple nuclei and irregular features, compared to just ~10% of elliptical galaxies which exhibit shells or ripples in Malin & Carter (1983) (observed with comparable imaging). Shells in particular are often unassociated with galaxy interactions (with around half of the galaxies exhibiting shells in isolated environments, suggesting mergers are an unlikely cause; Malin & Carter, 1983) and so the "control" group of Malin & Carter (1983) likely overestimates the true merger fraction, plausibly bringing the detected merger excess in line with estimates from this work (i.e. $\gtrsim$10). Canalizo & Stockton (2001) present a sample of 9 low redshift AGN between the ultraluminous infrared galaxy and AGN regions of the far-IR colour-colour diagram. From spectral analysis they find a range of starburst ages (t_{burst} = 0-300Myr) and a 100% incidence of interactions, with 8 major mergers. They make use of HST imaging with far higher resolution than the Illustris images here. This is a small sample comprised of a sub-group of AGN and it is not clear how the results apply to the AGN population as a whole, though it represents compelling evidence that mergers can act as an efficient trigger (even if indirectly, via star formation) for nuclear activity.

Up to this point we have mostly considered the influence of major mergers on the central engine. It is plausible that minor mergers, possibly accounting for ~40% of the star formation in the local universe (Kaviraj, 2013), may be important triggers of SMBH growth. Our results suggest that, while deeper imaging may be capable, SDSS-like imaging is unable to reliably detect minor mergers. Minor mergers thus represent a largely un-probed region of parameter space while being significantly more numerous than major mergers. It is plausible that minor mergers could drive much of the SMBH growth in the local universe, but unambiguously identifying minor mergers and distinguishing them from older major mergers remains challenging.
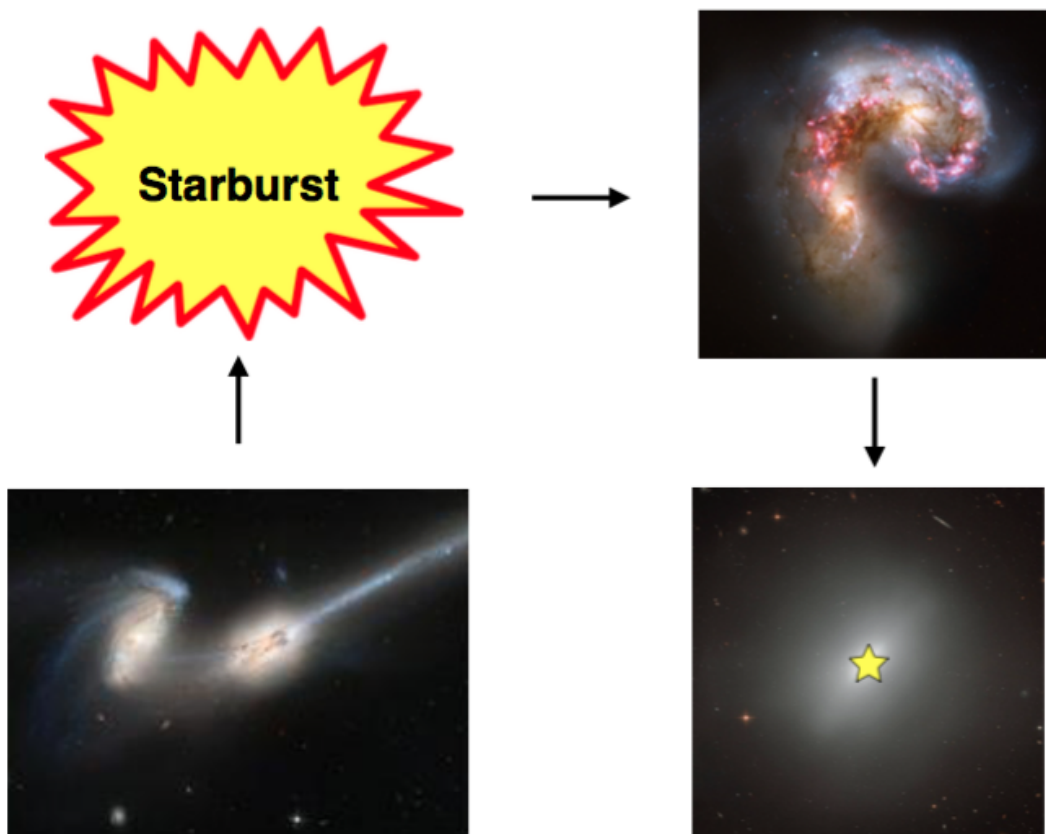
Figure 3.17: A cartoon of the possible time evolution of AGN in galaxy mergers (time proceeding clockwise from bottom left), where an enhanced SFR immediately after the merger, making the merger easily detectable, and a delay before the black hole growth (by which time merger features have faded) conspire to obscure a real underlying merger-AGN connection.

3.5 Conclusions

Mock SDSS-like images of simulated Illustris galaxies were constructed and analysed similarly to real galaxy images, including samples of major mergers, minor mergers and non-interacting galaxies. Structural morphology parameters were measured and a subset of images visually classified by 5 expert classifiers. The efficacy of different structural parameters at identifying mergers was established for $\sim$800Myr after the merger, revealing that shape asymmetry (Pawlik et al., 2016) more robustly identifies mergers than other parameters and is comparably effective to human visual classification (though with a slightly higher false-positive rate). Humans reliably identify young major mergers, but become rapidly less reliable as the merger ages. It is found that the likelihood of identifying a merger is a strong function of the SFR, for both mergers and non-mergers alike, with highly star forming galaxies $\sim 5\times$ more likely to be identified as mergers relative to more quiescent galaxies (mean SFRs of $\sim$9$M_\odot yr^{-1}$ and $\sim$1$M_\odot yr^{-1}$ respectively).

Given the inherent observational biases involved in identifying mergers we reinterpret the merger excesses reported in the literature where comparable imaging is used (i.e. low redshift samples imaged with the SDSS). Mock AGN and control samples are constructed with merger fractions corresponding to the supposed enhancement to the AGN duty cycle from mergers (e.g. if the duty cycle is doubled by mergers, so too is the AGN merger fraction relative to controls). The true merger fraction is then compared to the measured AGN fraction to determine the bias imposed by observational uncertainty, and knowledge of this bias is used to correct the reported merger-excesses in the literature. The observational bias is further dependent on the SFR of the observed galaxy, the time elapsed since the merger, the galaxy mass (related to SFR) and the mass ratio. One may assume a range of triggering models (e.g. AGN are triggered by young major mergers; AGN are triggered by old mergers of massive galaxies etc.) to demonstrate what would be observed in each case, and compare to real observations to constrain what are the likely processes in nature.

In the literature, low redshift AGN hosts are found by Koss et al. (2011); Cotini et al. (2013); Kaviraj et al. (2015) to exhibit an excess of AGN in mergers compared to controls by a factor of 3.3-5. Taking observational biases into account and considering possible delays between mergers and nuclear activity (as reported in literature) leads to the conclusion that the intrinsic merger-excess in these samples is between 11-43$\times$ greater than in their control samples, with the large range arising from the dependence on the specific triggering model assumed. Assuming the true merger rate in the Universe is consistent with the merger rate in the Illustris simulation (Rodriguez-Gomez et al., 2015a) this implies 51$\pm$21% of mergers in the local Universe are merger triggered. This is consistent with other studies where a sample of mergers and non-interacting controls are first selected and then AGN identified (Ellison

et al., 2013; Satyapal et al., 2014). A delay between a merger-induced starburst and nuclear activity could very effectively hide this connection, since as the SFR declines post-merger the merger becomes harder and harder to detect, serving to both mildly enhance the detected merger fraction of inactive control galaxy samples and reduce the detected merger fraction in AGN samples. We are unable to quantify the possible influence of minor mergers, which may constitute a hidden triggering mechanism, since the minor mergers in our sample were undetected by both visual classification and structural parameters.

4

Varied Triggering Mechanisms of Radio, Optical & IR Selected AGN

AGN can be selected by a range of methods (chapter 1). X-rays tend to be copiously produced near the SMBH; powerful radio jets may be launched from spinning SMBHs (Blandford & Znajek, 1977); the thermal output from the disk can be detected in UV/optical emission; gas clouds rapidly orbiting the BH can reprocess light from the central source into broad emission lines; the hard spectrum can excite gas throughout the galaxy and produce highly ionised species of narrow emission lines (BPT method Baldwin et al., 1981); dust from the host galaxy can process AGN light and reemit it in the IR (Stern et al., 2012). No single method is able to detect all AGN in a given volume and different emission processes may trace physically distinct systems, in addition to each having different associated detection biases. For instance, if radio jets are launched by taking energy from the BH spin then their presence may depend on the history of those AGN, in which some process has spun up the BH (such as a merger with a comparable mass BH); identifying AGN from the broad lines requires an unobscured AGN, biasing against dusty galaxies (and AGN with small torus opening angles); IR emission correlates with dust mass, so biases detection toward dust-rich, generally star forming galaxies (Stern et al., 2012).

An interesting question is how the fuelling processes may differ for these different AGN. In this study samples of radio, optical and IR selected AGN were compiled from the Vista survey,

with mass and redshift matched control galaxies, to attempt to understand the relative role of mergers in triggering each. They occupy a narrow range of redshifts, $0.2 < z < 0.5$, and have stellar masses $10^9 < M\,[M_\odot] < 10^{12}$. Optical AGN are selected form the broad emission lines and so are of type 1, radio and IR emission penetrates the dust and these samples are of type 2. The structural parameters Gini, M_{20}, shape asymmetry, concentration and smoothness are computed. Significant differences are found between the hosts of radio AGN compared to their controls and IR AGN compared to theirs. Radio AGN hosts are found by several metrics to exhibit lesser indicators of recent interactions than their controls, while the opposite is true for the IR AGN. Direct comparison to optical AGN is complicated by the presence of the point source, but using metrics which are relatively unaffected by the nuclear light optical AGN are found to be comparable to their control galaxies. This provides evidence that different classes of AGN have significantly different fuelling mechanisms, with varying contributions to their growth due to major mergers. Possible reasons for differences between AGN classes are considered.

The samples of AGN and their controls are introduced in section 4.1; the methodology in section 4.2; the results are presented in section 4.3 and discussed in section 4.4 before conclusions are presented in section 4.5.

4.1 Samples

The VISTA Kilo-Degree Infrared Galaxy survey (VIKINGs) has a 5σ AB J band point source depth of $\sim$22 mags, with imaging over an area of 1500 degrees[2] (Edge et al., 2013) and significant overlap with other surveys, including with the SDSS fields (used to determine galaxy properties and identify AGN). Near infrared imaging is ideal for studying galaxy morphology as it probes the old stars, representative of the distribution of stellar mass through the galaxy (unlike the young blue stars clumped in star forming regions). J band images were used throughout. Selection of AGN and their respective control galaxies is described below.

4.1.1 Optical Sample

Optical selection of AGN is one of the most commonly used methods for identification, since type 1 AGN tend to outshine their host galaxy, rendering detection relatively simple. The Shen AGN catalogue (Shen et al., 2011) contains $\sim$100,000 AGN from the SDSS. In Shen et al. (2011) (and more generally) SMBH masses are derived from fits to broad MgII lines, the FWHM of which gives an estimate of the orbital velocity, and the continuum luminosity, which provides a proxy for the distance to the broad line region (e.g. Peterson, 2006). As described in chapter 1, this allows the BH mass to be calculated. The detection of a broad line region demonstrates that these are type 1 AGN. AGN from the Shen catalogue were searched for in

the Vista fields in the desired redshift range, finding 295 matches.

Stellar masses are typically estimated from spectra (stellar absorption lines Kauffmann et al., 2003) or photometry, but these indicators are highly biased by emission from the AGN. As such, these AGN hosts do not have stellar mass estimates. However, as in chapter 2, the BH mass estimates combined with the M-σ relation of McConnell & Ma (2013) (equation 2.1) can be used to estimate host galaxy masses. A distribution of possible masses was produced for each AGN host, values chosen from this distribution at random and matched to galaxies in the MPA/JHU catalogue for SDSS data release 7. The control galaxies have stellar masses in the range $10^9 < M\,[M_\odot] < 10^{12}$. 798 control galaxies were identified in the Vista fields.

4.1.2 IR Sample

Uncovering type 2 AGN is vital for both understanding the triggering mechanisms responsible for black hole growth and understanding AGN populations as a whole. Obscured AGN may account for $\sim$75% of the AGN population (Comastri et al., 1995; Treister et al., 2004; Ballantyne et al., 2011) but surveys tend to be greatly biased toward type 1 AGN, due to their domination over the galaxy across the electromagnetic spectrum. In the simple picture of AGN depicted in figure 1.2 type 1 and 2 AGN are intrinsically similar, only viewed from different angles. A more complete picture places AGN in a wider galactic environment, where obscuration can have a range of sources. It has been long suggested that mergers may enshroud AGN in dust for a period after coalescence (while galaxies may be most obviously morphologically disturbed Hopkins et al., 2008) and increase line of sight column densities of hydrogen, biasing against detection of AGN in mergers in most surveys (Sanders et al., 1988; Fabian, 1999; Draper & Ballantyne, 2010).

AGN exhibit a near power-law spectrum in the mid-IR, distinct from the blackbody spectrum of stars and therefore relatively easily identified. Mid-IR colours are not so hampered by dust extinction, making them ideal for studying hidden populations of AGN. Stern et al. (2012) use the Wide-field Infrared Survey Explorer (WISE; Wright et al., 2010) bands, W1 and W2, to identify the AGN power-law contribution to IR colours and thus find AGN. WISE is an all-sky survey with a 5σ point source depth of 0.08 and 0.11mJy in W1 and W2 respectively (Wright et al., 2010). A simple cut of W1-W2$\geq$0.8 suffices to identify most AGN at low redshift; for instance in the redshift range under investigation here this criterion identifies $\sim$84% of the (type 1) AGN in the Shen catalogue. The Stern criterion was applied to targets in the SDSS spectroscopically identified as galaxies to find obscured AGN, yielding 578 AGN candidates. Of these, only 20 lay in the VIKINGs fields. These make up the IR selected AGN sample used in this work.

Unlike the optically selected AGN sample, these AGN are of type 2: the spectra are not so
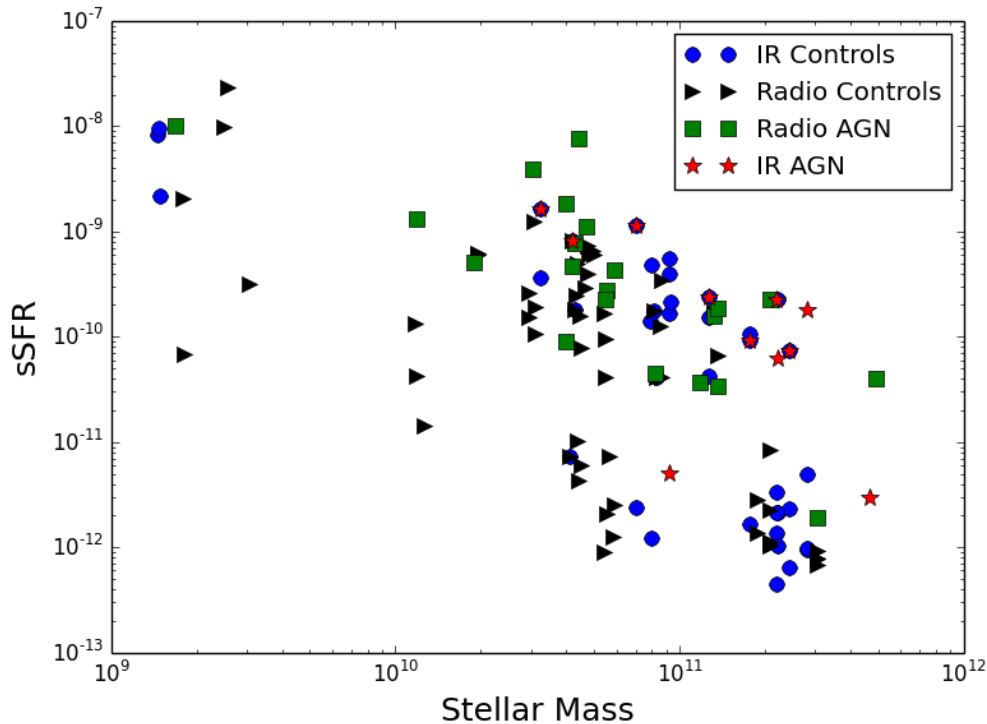
Figure 4.1: Stellar mass [in Solar masses] vs specific star formation rate for IR and radio selected AGN and their controls. Values are taken from MPA/JHU catalogues. Star formation rates are lacking for some galaxies, which is why there are more controls than AGN hosts at low stellar mass.

contaminated by nuclear emission and stellar masses and star formation rates can be calculated (though scattered continuum light can bias measurements, see section 4.4). Stellar masses are calculated from stellar absorption lines (Kauffmann et al., 2003) in 13 cases and star formation rates measured from the 4000Å break in 11 cases. They are matched directly to galaxies in the MPA/JHU catalogue in stellar mass and redshift to the closest possible matches. For the 7 AGN hosts with no mass measurements the mean of the 13 known masses is assumed, leading to a total of 40 control galaxies. A plot of stellar mass vs specific star formation rate for IR selected AGN, radio selected AGN (below) and their controls is shown in figure 4.1. As can be seen, IR and radio AGN apparently have significantly higher star formation rates than their controls. Possible causes of this measurement and the potential effect on results is discussed in section 4.4.

4.1.3   Radio Sample

Finally, radio loud AGN are known to lie in significantly different galaxies and galactic environments than radio quiet AGN (Hickox et al., 2009, e.g.). Only ~10% of AGN are radio loud (Stern et al., 2000) and they tend to live in massive galaxies (harbouring the most massive BHs), often in dense galaxy clusters (Tadhunter, 2016). Such a distinct population may be expected to have different accretion histories, triggering mechanisms or physical characteris-

tics to the rest of the AGN population. Radio frequency photons are emitted predominantly by two processes: Bremsstrahlung and synchrotron radiation, both emitted due to the fact that accelerating charges emit radiation (Larmor, 1897). Bremsstrahlung is emitted from rapidly decelerating charged particles as they interact with other charged particles: e.g. from relativistic electrons encountering 'stationary' protons. Synchrotron radiation is caused by charged particles spiralling around magnetic field lines and is the dominant emission mechanism for radio AGN. Radio emission is insensitive to dust obscuration and so can offer a relatively clean selection of AGN.

The Faint Images of the Radio Sky at Twenty-cm (FIRST) survey covers the SDSS fields with the Very Large Array (Becker et al., 1994) with a point source sensitivity of 1mJy (Gregg et al., 1996). Searching the SDSS fields for objects with $L_{1.4GHz} > 10^{24}WHz^{-1}$ radio loud sources, assumed to be AGN, were identified (as in Tadhunter, 2016, since above this luminosity contamination from star forming galaxies is very low). 1600 AGN candidates were identified from the SDSS targets spectroscopically identified as galaxies (to ensure AGN are of type 2), 67 of which lay in the VIKINGs fields and 24 of which had existing stellar mass measurements. As with the IR sample, control galaxies are mass and redshift matched where possible, and redshift matched to a galaxy with the mean mass of the known AGN host mass distribution otherwise, leading to 70 matched control galaxies. The distribution of known stellar masses and star formation rates is shown in figure 4.1. As with IR AGN, radio AGN tend to have higher star formation rates than their controls, which may somewhat bias results (see section 4.4). 100% of radio-selected AGN are also identified as AGN with WISE photometry, while only 3 IR-selected AGN are identified as radio-loud.

4.2 Methodology

To compare the morphologies of galaxies, in particular to evaluate the likely fractions of mergers in each category, the structural parameters Gini, M_{20}, shape asymmetry, smoothness and concentration are calculated. Of these, only smoothness and concentration were not discussed in chapter 3, so are described below. Asymmetry (Abraham et al., 1996) is omitted in favour of the updated shape asymmetry parameter (Pawlik et al., 2016).

Concentration measures the relative density of flux in the centre to the outskirts of a galaxy (Bershady et al., 2000) by:

$$C = 5 \log \left(\frac{R_{80}}{R_{20}} \right) \tag{4.1}$$

where R_{80} and R_{20} are the radii of circular apertures containing 80% and 20% of the light respectively. The total flux is defined by a circular aperture with radius set by the binary

detection mask (the distance between the outermost pixel associated to the galaxy and the central brightest pixel), as per Pawlik et al. (2016).

The smoothness is a measure of the clumpiness of the light distribution in a galaxy, sensitive to small structures within a galaxy such as star forming regions (Conselice, 2003). The image is smoothed with a Gaussian filter, with a smoothing radius set by R_{20}, and compared to the original image pixel by pixel:

$$S = \frac{\Sigma |I_0 - I_\sigma|}{\Sigma |I_0|} \tag{4.2}$$

where I_0 is the intensity of a pixel in the original image and I_σ is the intensity in the smoothed image. If light were equally distributed across the galaxy the smoothing filter would change nothing and the value of smoothness would be 0; the clumpier the original image the higher the value of smoothness.

4.3 Results

The results of parametric structural parameter measurements are presented here. The key comparisons are between AGN and their matched controls, and between the differently selected groups (IR AGN to controls versus radio AGN to controls for instance). For completeness optical AGN (and controls) are included in the analysis, with the caveat that the presence of the point source severely biases certain measures. Where comparisons are consequently unreliable this will be indicated.

Each structural parameter provides new information about the distribution of light in a galaxy. Concentration (C) is sensitive to the central flux density, shape asymmetry (SA) is sensitive to structure on the outskirts of galaxies, M_{20} provides information about the physical separation of bright regions, smoothness (S) is sensitive to the small structures within galaxies and Gini (G) is sensitive to the average intensity inequality across pixels. Correlations certainly exist between these measures: for instance, if a galaxy has high concentration it will also have high Gini coefficient, since much of the light will be focussed in a few central pixels. However, if there are bright regions far from the galaxy centre the Gini index will be high and concentration low, allowing Gini to identify structures in galaxies which concentration alone would miss.

The product of interest here is not so much the distribution of parameter values as the relations between parameters. A high value of G could indicate the presence of multiple cores, star forming clumps, or a highly concentrated galaxy such as an elliptical; a high value of G and SA simultaneously much more strongly implies the former, since SA is insensitive to intragalactic structure. Such relations allow lines to be drawn above which galaxies can be said to have high probability of interaction.
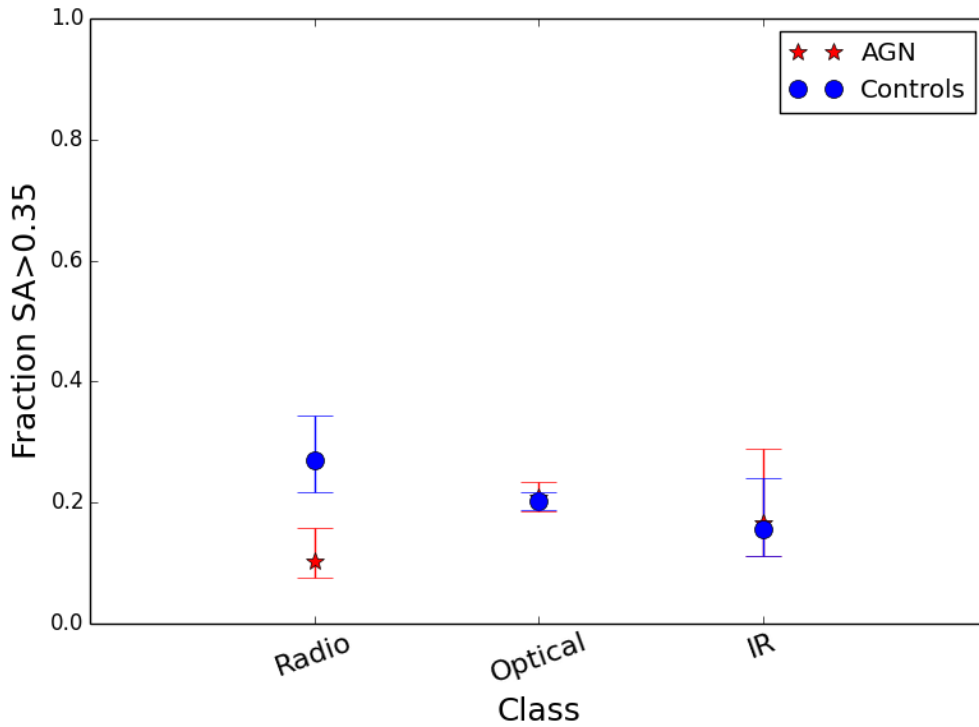
114

Figure 4.2: The fraction of galaxies in each selection class with shape asymmetry values greater than 0.35, the value above which 20% of the galaxies lie, indicative of a recent galaxy interaction. Errors are 1σ binomial errors (Cameron, 2011).

First, as in chapter 3, the fractions of AGN and control galaxies with high values of shape asymmetry in each category can be compared (Pawlik et al., 2016), as in figure 4.2. The histograms of SA distributions are shown in figure 4.3. Since VIKINGs is a deeper survey than the SDSS (Edge et al., 2013) more low surface brightness features are detected than for galaxies in chapter 3, meaning values of shape asymmetry are higher and the previous threshold for interaction, shape asymmetry > 0.2, is inappropriate. The threshold used here, shape asymmetry > 0.35, is chosen such that 80% of all the VIKINGs galaxies (optical, IR and radio AGN and controls) are identified as non-interacting galaxies. Varying the threshold does not qualitatively influence results or any conclusions drawn. It is found that IR and optical AGN have comparable shape asymmetry values to their controls, while radio AGN have fewer highly asymmetric galaxies than their controls, significant at the 2.2σ level.

Conselice (2003) find mergers to have values of asymmetry (A) > 0.38 and S > A. Similarly, galaxies with values:

$$SA > 0.35 \quad \text{and} \quad S > SA \tag{4.3}$$

are taken to be merger candidates here. High values of S imply clumpy morphologies, such as due to star formation, and high values of SA are associated with galaxy mergers due to
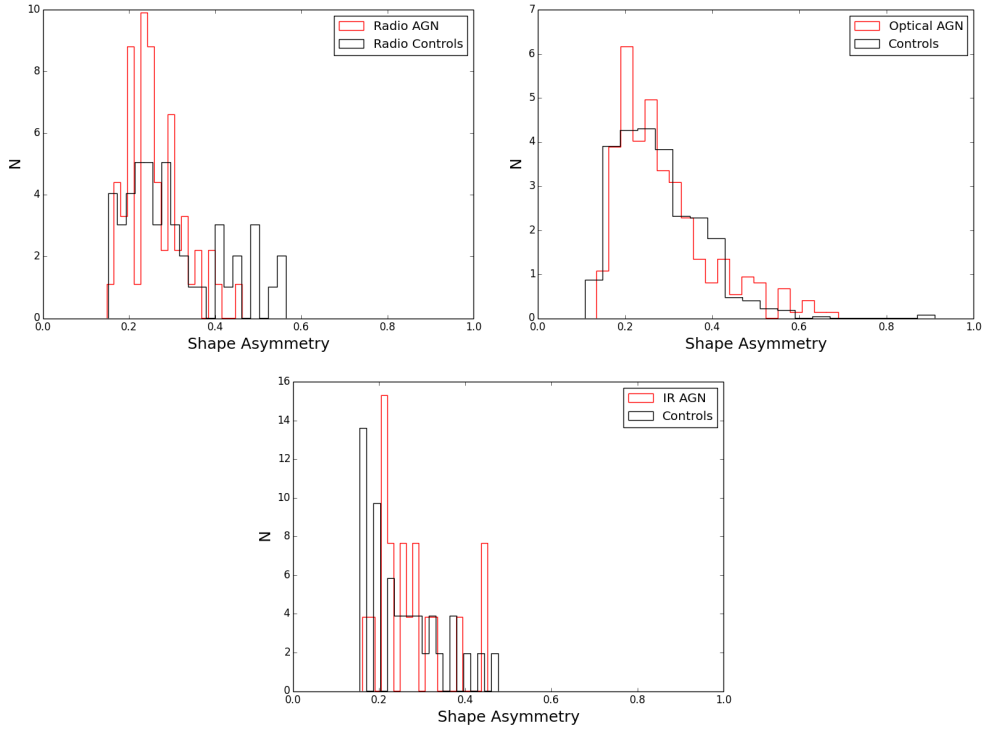
115

Figure 4.3: Shape asymmetry histograms for the radio, optical and IR samples, plus controls.

disruptive gravitational torques. Figure 4.4 shows the fractions of each galaxy group satisfying these criteria. No difference between AGN and their controls is found.

Lotz et al. (2004) set the criterion for interaction used here:

$$G > (-0.4 \times \text{SA}) + 0.66 \tag{4.4}$$

the only difference to the definition in Lotz et al. (2004) being the substitution of asymmetry for SA. Similar to the above case, G tends to have high values in interacting systems due to the presence of multiple cores or star formation biasing the light distribution and SA relates to mergers as described above. Figure 4.5 shows the fractions of AGN and their controls which satisfy this criterion, indicative of mergers. The values of G for optical AGN are artificially shifted to higher values by the point source and so optical AGN are omitted. Stark differences between radio AGN and their controls compared to IR AGN and theirs remain however. Radio AGN are found to be significantly less likely to be associated with mergers than the AGN with $\sim 3.4\sigma$ significance, while IR AGN show the opposite (though not statistically significantly).

A commonly used criterion for interaction is location in G-M_{20} parameter space (Lotz et al., 2004):
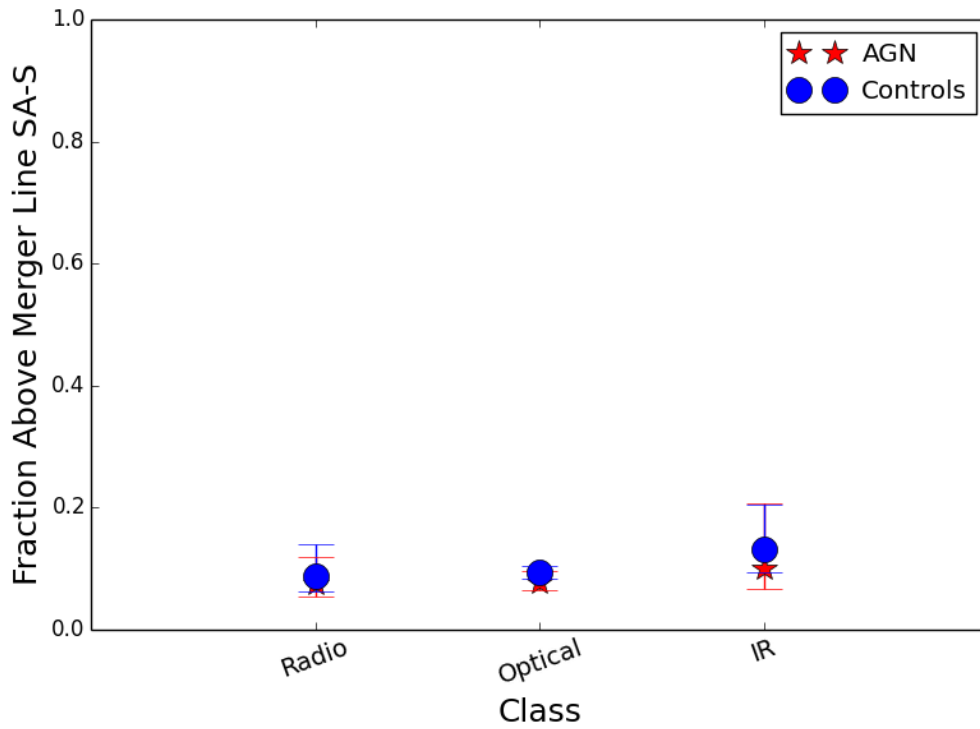
$$G > (-0.115 \times M_{20}) + 0.384 \tag{4.5}$$

116

Figure 4.4: The fraction of galaxies in each selection class which satisfy the criteria of equation 4.3. Errors are 1σ binomial errors (Cameron, 2011).

Figure 4.5: The fraction of galaxies in each selection class which satisfy the criteria of equation 4.4. Optical AGN have biased measurements of Gini and the comparison to their controls is unreliable and as such they are omitted. Errors are 1σ binomial errors (Cameron, 2011).
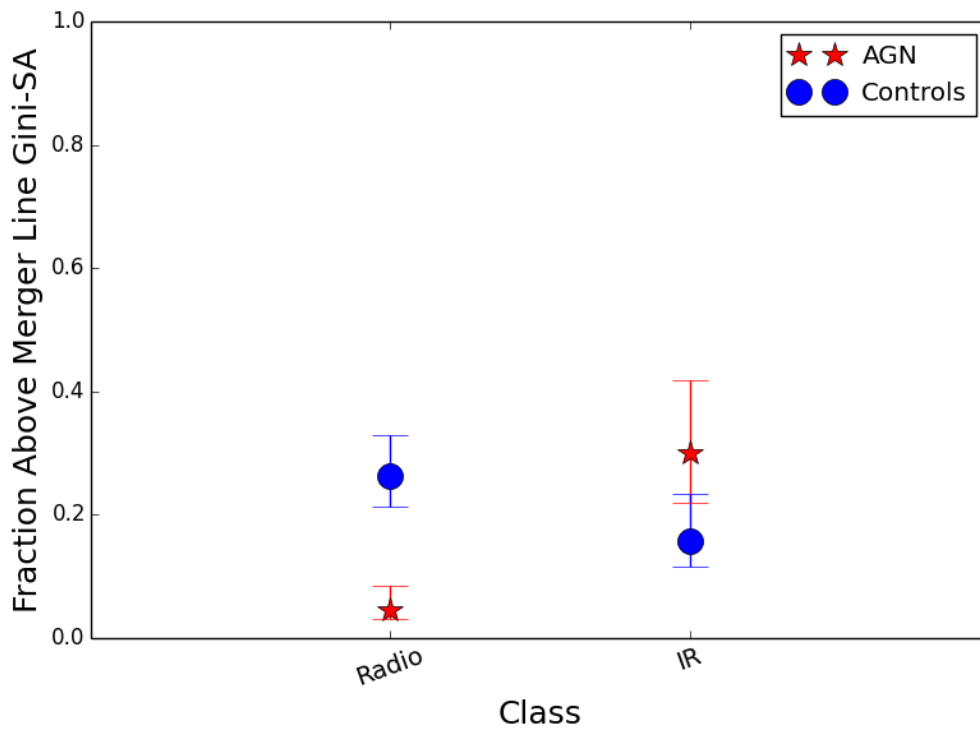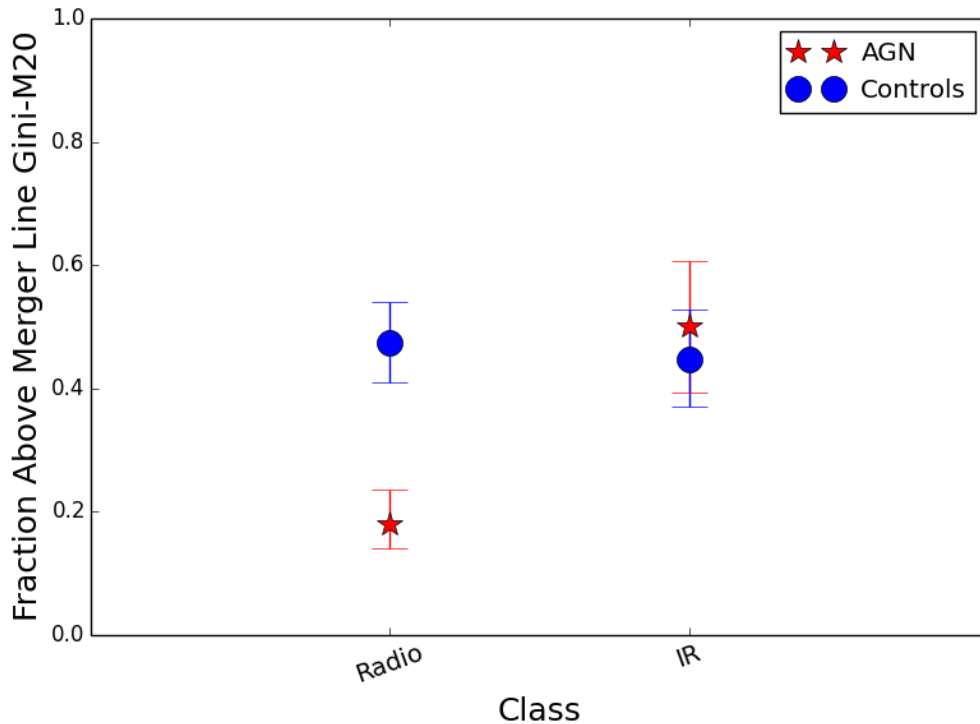
Figure 4.6: The fraction of galaxies in each selection class which satisfy the criteria of equation 4.5. Optical AGN have biased measurements of both Gini and M20 so the comparison to their controls is unreliable and as such they are omitted. Errors are 1σ binomial errors (Cameron, 2011).

as used in chapter 3. High values of G imply unequal distributions of flux across pixels while high values of M_{20} imply that those bright pixels are far from the galaxy centre. Combined, the above criterion can be used to exclude classes of galaxies (such as ellipticals which have high values of G but low values of M_{20}) and suggest the existence of multiple cores, a strong merger indicator. As in figure 4.5 the optically selected sample is unreliable in both G and M_{20} due to the point source, so is omitted. Figure 4.6 again shows a significant ($\sim 3.5\sigma$) difference between radio AGN and their control galaxies, with the controls again more likely to exhibit signs of interaction. The IR AGN are comparable to their controls.

Finally, C-M_{20} can be used to identify interacting systems (Lotz et al., 2004; Holwerda et al., 2014) using the following criterion:

$$C > (-2.5 \times M_{20}) + 1 \tag{4.6}$$

C and M_{20} are correlated since C measures the relative flux in the galaxy centre and M_{20} measures the distance of the brightest pixels to the centre. Galaxies resultantly tend to lie along a line on a C-M_{20} plot. Outliers from this can be inferred to have irregular morphology. In figure 4.7 the optical sample is again unreliable, but a significant ($\sim 3.4\sigma$) excess of IR AGN in merger candidates relative to controls is detected. Radio AGN again show a slight deficit
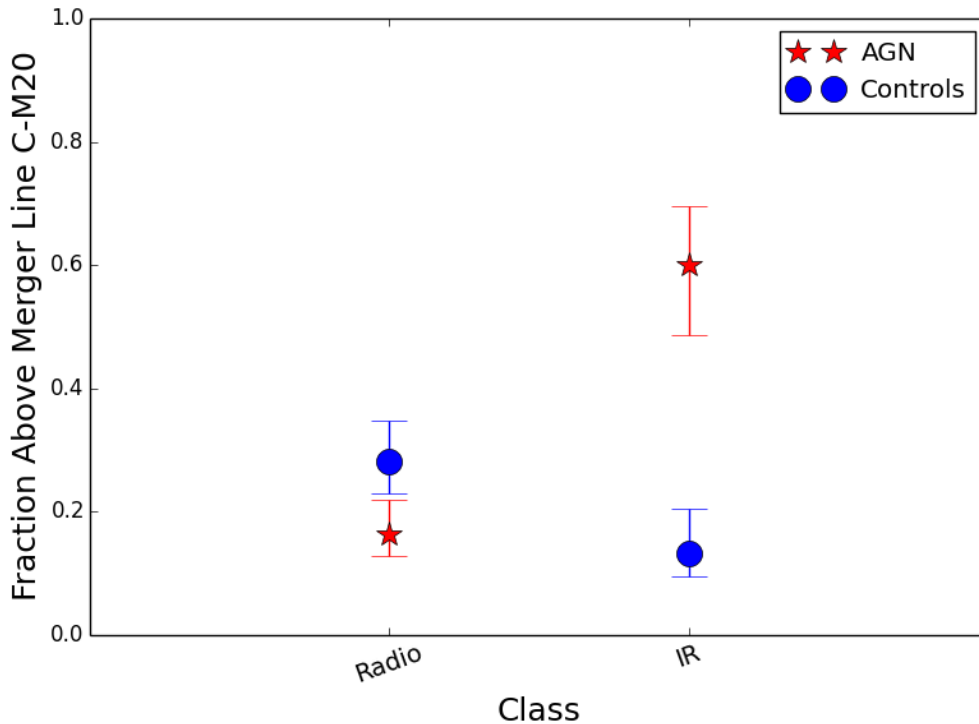
Figure 4.7: The fraction of galaxies in each selection class which satisfy the criteria of equation 4.6. Optical AGN have biased measurements of both Gini and M20 so the comparison to their controls is unreliable and as such they are omitted. Errors are 1σ binomial errors (Cameron, 2011).

relative to controls, but with $<2\sigma$ significance.

4.4 Discussion

In the previous section we compared the fractions of galaxies classified as mergers using structural parameters, separated by AGN selection method. The optically selected AGN tend to have unreliable evaluation due to point source contamination, but radio and IR selected AGN are of type 2, thus have no point source in images, and so the discussion will focus on these two sets of galaxies. We find significant differences between radio-selected AGN and their controls compared to IR-selected AGN, with radio-loud AGN significantly less likely to be observed in mergers than their controls and the opposite for IR AGN (by some measures), broadly consistent with other studies. We will consider possible reasons for these differences, including differences in SFR and stellar mass between different AGN samples and their controls, a contrasting methodology to that used in the literature, and exploring the possibility that an explanation for this difference lies in an evolutionary sequence for AGN fuelling.

First we consider the possibility that a mismatch in SFR may lead to the observed difference between radio-selected AGN and IR-selected AGN relative to their controls. From inspection of figure 4.1 it seems that both radio- and IR-selected AGN have higher sSFRs than their controls,

indeed KS tests of the sSFR distributions for radio and IR AGN compared to their controls give P values 0.004 and 0.024 respectively. However, radio AGN hosts have comparable sSFRs to IR AGN hosts, with a KS test P value of 0.645. From the results of Chapter 3 it is expected that higher SFRs lead to a higher rate of detected mergers; however, while both IR and radio AGN have higher SFRs than their controls radio AGN appear to be less associated with mergers while the opposite is true for IR-selected AGN. Estimation of SFRs in AGN hosts is notoriously unreliable since nuclear emission can dominate the optical emission lines or other spectral features used to estimate the SFR. For these galaxies SFRs are estimated from the strength of the 4000Å break (Brinchmann et al., 2004): metals in stellar atmospheres scatter and absorb blue radiation, leading to a characteristic drop-off in intensity below 4000Å. This dip can be offset by a large population of recently formed massive blue stars and so the strength of the 4000Å break provides information about the rate of star formation. However, scattered light from the central AGN, or shock heating from jets, may provide this blue emission and contaminate estimates of the SFRs. Radio AGN typically have low SFRs (Tadhunter, 2016) and so the relatively high rates registered in catalogues for some of the radio AGN here raised suspicion. Inspection of the spectra indeed suggests the radio AGN SFRs cannot be trusted, with significant AGN components. IR AGN will suffer from some of the same systematic issues, with contamination in some sources, but tend to have intrinsically higher SFRs. This is perhaps reflected in the higher median value of SFR ($\sim31M_\odot yr^{-1}$ vs $\sim19M_\odot yr^{-1}$ for IR and radio AGN respectively) but lower mean ($\sim32M_\odot yr^{-1}$ vs $\sim41M_\odot yr^{-1}$). If the SFRs are intrinsically higher for the IR AGN than the radio AGN this may yield a higher detected merger fraction, as outlined in chapter 3, partially or fully explaining the differences across the samples, but a full analysis of the SFRs of the hosts is beyond the scope of this work.

Next we consider the possibility that the techniques used here are not sophisticated enough to identify relevant merger features. Some literature studies use deep photometry and visual inspection to identify merger features (e.g. Ramos Almeida et al., 2011), compared to the relatively shallow imaging and analysis of structural parameters used here. Radio-loud AGN are predominantly found in elliptical galaxies (e.g. Hickox et al., 2009), which typically have low asymmetry and smooth distributions of stellar light, meaning that many of the measures used here are biased against identifying mergers. In particular, Gini, M20 and smoothness are expected to take on low values while concentration will be relatively high. Contrarily, IR-selected AGN are more commonly associated with late type galaxies (Hickox et al., 2009; Stern et al., 2012) and so, without any contribution from mergers, one might expect to find a higher incidence of morphologies consistent with tidal disruption in the IR-selected sample compared to the radio sample, as seen in figures 4.5, 4.6 and 4.7. The control galaxies are merely mass- and redshift-matched, and so will have comparable morphologies across categories due to the similar stellar masses of these AGN host samples. While this may be a contributing

factor, the use of multiple structural parameters in concert should ensure that the (significant) differences between samples are not just due to different typical host galaxies but are due to a real difference in the prevalence of tidal features between AGN classes.

Finally we consider the possibility that the differences in AGN morphological properties are due to intrinsic differences in the fuelling mechanisms for different AGN. IR- and radio-selected AGN are both of type 2, meaning the central engine is obscured from view by intervening gas and dust. An orientation-based AGN unification model would explain this as due to a torus of dust enclosing the SMBH (Barthel, 1989), but there is no fundamental reason for gas and dust dispersed through the galaxy not to provide similar obscuration. It has been hypothesised that increased line-of-sight column densities of gas and dust following a merger can hide nuclear activity; distinguishing this effect from chance alignment of a torus is challenging, but one possible consequence would be for type 2 AGN to be disproportionately associated with galaxy mergers relative to type 1 AGN (e.g Kocevski et al., 2015). IR-selected AGN are particularly effective for this analysis, since the nuclear emission can be reprocessed by dust and so the presence of an AGN inferred from unimpeded long-wavelength radiation. The literature is mixed regarding the merger rates for obscured vs unobscured AGN. Schawinski et al. (2012) found a low fraction of mergers in dusty obscured galaxies at z~2, consistent with the comparable unobscured AGN sample of Schawinski et al. (2011). Kocevski et al. (2015) found increased evidence of galaxy interactions as a function of obscuration, suggesting AGN tend to be obscured soon after the merger, thus hiding an intrinsic connection between the merger and black hole growth. In this work, IR-selected type 2 AGN are found by some measures (C-M20 and to a lesser extent Gini-SA, figures 4.7 and 4.5 respectively) to show greater incidence of galaxy interactions relative to their controls. Other measures find IR AGN and their controls to be comparable (figures 4.2, 4.4 and 4.6). Overall we fail to determine conclusively the role of mergers on the IR AGN population, but find hints that mergers play a greater role with up to 3.4σ excess relative to controls.

While relatively high dust masses are typical of IR-selected AGN (Stern et al., 2012), leading to a slight preference to detect AGN in interacting or star forming galaxies, the mechanism by which radio emission is produced by the AGN is largely unrelated to galaxy properties and is instead related to intrinsic properties of the SMBH (e.g. its spin Blandford & Znajek, 1977). Therefore, while IR-selected AGN have a slight bias to be obscured by dust and gas throughout the galaxy, radio-selected AGN have no such bias and a simple orientation-based unification model is likely sufficient to explain the obscuration of radio AGN (though some may still be obscured by intervening gas and dust). All of the radio-selected AGN here are also detected using WISE photometry. In the literature radio AGN are categorised into numerous classes, dependent primarily on orientation and the extent and morphology of radio emission. Different classes of radio galaxies have different inferred triggering mechanisms

and radio loudness. One of the most common categorisations is the Fanaroff-Riley classification. Fanaroff-Riley class 1 objects have lower radio luminosities (relative to galaxy luminosity) and compact radio jets, localised near the central SMBH; Fanaroff-Riley class 2 objects have extended radio lobes, sometimes extending megaparsecs, and higher radio luminosities (Fanaroff & Riley, 1974). The former are typically found to have elliptical morphologies inconsistent with merger-triggering as a dominant fuelling mechanism, live in dense environments and have low SFRs (Tadhunter, 2016); the latter show greater evidence of galaxy interactions, live in groups and have slightly higher typical SFRs (Tadhunter, 2016). A comparison of radio loudness at 1.4GHz to r-band magnitude demonstrates that the majority of radio AGN in this study have Fanaroff-Riley class 1, and so the expectation from previous studies is for these to have undisturbed elliptical morphologies. The radio AGN here show less evidence of recent mergers than the control sample, suggesting they are typical of Fanaroff-Riley class 1 radio galaxies. To power a megaparsec-scale jet (class 2) requires large quantities of gas to be continually funnelled to the AGN over a long period of time and therefore, since these are typically gas-poor galaxies, may require a merger or some other cataclysmic event. The radio AGN here likely have smaller jets and correspondingly lower radio luminosities (class 1) and therefore it is likely they are fuelled more slowly from the insubstantial gas reservoirs within the galaxy. At low accretion rates, with low gas densities, energy is likely inefficiently radiated away. Disks may puff up into a quasi-spherical structure and this hot-gas accretion may be the fuelling mechanism for the radio AGN investigated here.

Taken together with the literature this broadly supports an evolutionary scenario for AGN: major mergers tend to trigger significant star formation (Toomre & Toomre, 1972), yielding bright tidal features; once merger features have somewhat faded the black hole starts to grow rapidly (Wild et al., 2010) but is enshrouded in gas and dust (e.g. Hopkins et al., 2008; Kocevski et al., 2015), occluding parts of the spectrum but remaining visible to mid-IR colours (Stern et al., 2012) and ultra-hard X-rays (Kocevski et al., 2015). Merger features continue to fade while the star formation rate decays; the AGN reveals itself in optical and X-ray colours due to AGN feedback (e.g Di Matteo et al., 2005; Hopkins et al., 2008), and/or the galaxy coming into dynamical equilibrium, leading to little-to-no observed merger-AGN connection in X-ray and optical studies (e.g. Villforth et al., 2017). SMBHs from the two galaxies settle into the new galaxy centre and may eventually merge, yielding a single rapidly spinning BH at the heart of a massive gas-poor elliptical galaxy. Further accreted gas is launched efficiently from the BH in jets, extracting energy from the BH spin (Blandford & Znajek, 1977) and causing the BH to grow slowly; the highest radio luminosities require large quantities of gas, meaning Fanaroff-Riley class 2 objects are preferentially associated with mergers (Tadhunter, 2016).

4.5 Conclusions

We have collated samples of AGN using IR-selection, optical-selection and radio-selection to compare the triggering mechanisms for different classes of AGN. Using combinations of structural parameters we assess the probable fractions of mergers in each class in comparison to mass- and redshift-matched control galaxies. Since optical AGN are identified by broad optical lines many of the structural parameters are biased by nuclear emission and so we focus on IR and radio AGN. It is found that radio AGN hosts exhibit less signs of interaction than their controls while the opposite is true (though with less confidence) for IR AGN. We cannot rule out the possibility that this is due to a mismatch in SFR since AGN contamination makes SFR measurements unreliable. If radio controls have higher SFRs than the radio AGN they will likely be observed with higher merger fractions, while if IR AGN have higher SFRs than their controls this could explain the slight observed merger-excess (as shown in chapter 3). It is also possible that structural parameters are not able to detect faint or old merger features and thus that the methodology applied here is not as sensitive as required, particularly for radio AGN. In spite of these cautions, the data are broadly consistent with the literature and suggest that IR-selected and obscured AGN are more commonly associated with mergers than optical or X-ray selected AGN, while Fanaroff-Riley class 1 radio AGN are fed stochastically by intragalactic processes or mergers too minor to robustly detect here. This can be understood as part of a sequence, where IR AGN are preferentially associated with mergers which enshroud the black hole in gas and dust, and occupy a diverse range of environments and galaxies. Lower luminosity radio AGN tend to occupy massive gas-poor elliptical galaxies where low-moderate rates of accretion drive weak radio jets and are primarily fuelled stochastically, while more rapid fuelling by major mergers leads to powerful jets.

5

Conclusions

In this thesis the connections between nuclear activity and galaxy mergers have been investigated. The conditions of the Universe have changed dramatically through cosmic time with gas fractions, SFRs and galaxy number densities declining rapidly since z~3. The relative role of mergers in driving BH growth may similarly be expected to change through time. Additionally, one might expect the triggering mechanism to vary with AGN luminosity: perhaps the most rapidly accreting BHs require the most cataclysmic events (i.e. mergers) to supply the BH with ample fuel. In chapter 2 we investigated the role of mergers in triggering AGN through cosmic time and as a function of AGN luminosity. We assembled a sample of 106 HST imaged luminous X-ray selected AGN between redshifts of 0.5 and 2.2, with matched control galaxies. Control galaxies were made indistinguishable from AGN hosts by adding stars, magnitude matched to AGN, to their centres. Using GALFIT, symmetric models of the galaxy light profiles were created and subtracted from images to reveal previously hidden low surface brightness features which may indicate recent mergers. A new method for quantifying morphological disturbance was created, summing the residuals (post model-subtraction) to obtain a numerical measure of disruption. This works because models are smooth and symmetric, while galaxies tend not to be after interactions with other galaxies. Visual classification of galaxies was also conducted and confirms the efficacy of the residual flux method in identifying merger candidates. It was found that, overall, the fractions of AGN in mergers was comparable to the fractions of merging control galaxies. No trend is found when the sample is split into lumi-

nosity bins, with AGN of all luminosities similarly likely to be identified as mergers as their control galaxies. However, when split into bins of redshift a mild trend emerges, with higher redshift AGN more likely to be found in mergers than their control galaxies. At z∼2 there is a factor of 4 enhancement of AGN in galaxies visually classified as mergers compared to control galaxies, with 99% confidence. Despite this trend, the fraction of galaxies exhibiting signs of major mergers is small (≤20%), suggesting that most luminous X-ray selected AGN are not triggered by major mergers.

In chapter 3 we attempt to reinterpret and contextualise studies of the AGN-merger connection in the literature through a comparative analysis of the Illustris simulation and observational data. Mock images of mergers and non-mergers from the Illustris simulation were obtained and made to resemble SDSS images. With real astronomical data it is unknown whether a galaxy has truly undergone a merger and so interpretation of results is problematic. Using highly realistic Illustris galaxy images all relevant parameters - the time elapsed since a merger occurred, the mass ratio of that merger, the SFR, gas mass, stellar mass etc. - are known in advance. This allows determination of the efficacy of different structural parameters and of visual classification in identifying mergers with a range of parameters. It is found that shape asymmetry is the most effective structural parameter, correctly identifying ≳30% of major mergers for ≳800Myr since the onset of the merger. Visual classification is highly effective at detecting the most recent major mergers (identifying ∼70% of major mergers within 132Myr) but becomes rapidly less reliable as the merger ages. The probability of being classified as a major merger is a strong function of the SFR, with the most rapidly star forming galaxies ∼5 times more likely to be classified as a merger than the least rapidly star forming galaxies, raising concerns that a mismatch in SFR between AGN hosts and control galaxies in the literature could yield strong but fallacious correlations between mergers and nuclear activity. Otherwise, a delay between a merger-induced starburst and nuclear activity could effectively hide real connections between mergers and AGN.

Samples of "AGN" and control galaxies were constructed with a factor, F, more mergers in the AGN sample, where F is the hypothesised enhancement to the duty cycle due to mergers. Comparing the detected merger enhancement to the intrinsic (F) reveals the factor by which studies in the literature may have underestimated the true merger enhancement. Comparable literature studies (those using SDSS imaging of low-redshift AGN and controls) find 3.3-5× more mergers in the AGN sample than the controls. Since many mergers go undetected and some non-mergers are mis-identified as mergers, the implied intrinsic merger excesses are 11-43× greater in the AGN samples than control samples. The large range arises from a dependence on the particular triggering model assumed (e.g. assuming only recent major mergers trigger AGN vs. assuming mergers trigger AGN with some delay). This implies 51±21% of AGN in the local Universe are merger triggered, consistent with studies where mergers are

selected first and their AGN fractions compared to the AGN fractions in isolated galaxies. In addition, though a direct comparison to the high redshift work of chapter 2 was not attempted, it is suggested that an increasing underlying merger fraction with increasing redshift could be sufficient to explain the apparent redshift evolution of merger triggering reported in chapter 2.

In chapter 4 we further investigate how observational biases may influence quantification of the role of mergers in AGN triggering, and explore the triggering of physically distinct AGN. AGN selected using optical, IR and radio emission are compared to matched controls and each other within a narrow redshift range ($0.2 < z < 0.5$). Different selection methods may trace inherently different environments, so mergers may play different roles for each of these samples, and for X-ray selected AGN as in chapter 2. Mergers may tend to obscure the X-rays and optical emission from AGN by increasing line-of-sight column densities of gas and dust, biasing against finding mergers in such samples. IR and radio emission will tend to penetrate the gas and dust, while radio AGN tend to be found in massive gas poor ellipticals. Combinations of structural parameters were used to identify merger candidates amongst the three samples and their controls. Optical AGN, unlike radio or IR, are of type 1 and so several of the structural parameters are heavily biased for this sample. It is found that radio AGN hosts show lesser signs of recent interactions than their controls, while the opposite is true for IR AGN. Optical AGN are found to have equivalent merger fractions to their controls by measures of morphology unbiased by the point source. This lends support to the hypothesis that mergers drive nuclear activity but obscure the AGN behind thick columns of gas and dust, such that IR-selected AGN show the greatest signs of recent mergers (as IR penetrates the gas and dust). Once the optical and X-ray emission is visible (possibly due to feedback) the merger features have largely faded and correlations between AGN and mergers are found to be weak or non-existent. FRI radio AGN tend to occupy massive gas poor ellipticals and are fed primarily by cooling flows, not mergers; the most extreme radio AGN (FRII) require larger pools of gas to sustain such powerful jets and may be largely fuelled by galaxy mergers.

Taken together, these results suggest galaxy mergers play a significant role in driving SMBH growth. Observational biases can easily hide intrinsic connections; AGN must be robustly detected using multiple wavelengths to reveal the whole population; and the effects of galactic parameters on observations, such as SFR, must be carefully considered and controlled for if results are to be reliable.

Acknowledgments

I would like to thank my supervisors, Carolin Villforth and Vivienne Wild, for endless support, patience, guidance and discussions. My dad, for refusing to answer a question with anything but another question, which taught me to think independently, and in fact for questioning pretty much anything which seemed self-evident but often transpired not to be. My mum, for often warning me that "there's a lot of maths in physics", for going cross-eyed at explanations and then asking surprisingly insightful questions, and for summarising concepts in poetry more clearly than seems reasonable. My sister who, with the patience of a teacher, taught me the invaluable skill of explaining concepts in relatable terms, while I babbled for probably years. My partner, for constantly helping with $\LaTeX$[1] and coding issues, and for enlightening discussion. To friends, family, teachers and the boundary conditions of the Universe for making any of this possible. Also to my funding body, the STFC, who supported this work.

Bibliography

Abbott B. P., et al., 2016, Physical Review Letters, 116, 1

Abraham R. G., Tanvir N. R., Santiago B. X., Ellis R. S., Glazebrook K., van den Bergh S., 1996, MNRAS, 279, L47

Abraham R. G., van den Bergh S., Nair P., 2003, ApJ, 588, 218

Aird J., et al., 2010, MNRAS, 401, 2531

Alexander D. M., et al., 2008, The Astronomical Journal, 135, 1968

Allevato V., et al., 2011, ApJ, 736, 99

Antonucci R. R. J., Miller J. S., 1985, ApJ, 297, 621

Arshakian T. G., 2005, A&A, 436, 817

Azadi M., et al., 2014, ApJ, 806, 187

Baldry I. K., Balogh M. L., Bower R., Glazebrook K., Nichol R. C., 2004, ApJ, 600, 681

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469

Baldwin a., Phillips M. M., Terlevich R., 1981, Publications of the Astronomical Society of the Pacific, 93, 817

Ballantyne D. R., Draper a. R., Madsen K. K., Rigby J. R., Treister E., 2011, ApJ, 736, 15

Barnes J. E., 1988, The Astrophysical Journal, 331, 699

Barnes Joshua E. Hernquist L. E., 1991, ApJ, 370, L65

Barthel P. D., 1989, The Astrophysical Journal, 336, 606

Becker R. H., White R. L., Helfand D. J., 1994, Astronomical Data Analysis Software and Systems III, 61, 165

Beckwith S. V. W., et al., 2006, The Astronomical Journal, 132, 1729

Bekenstein J., 1973, Physical Review D, 7, 2333

Bennert N., Canalizo G., Jungwiert B., Stockton A., Schweizer F., Peng C. Y., Lacy M., 2008, ApJ, 677, 846

Bentz M. C., Peterson B. M., Netzer H., Pogge R. W., Vestergaard M., 2006, Astrophysical Journal, 697, 160

Bershady M. a., Jangren A., Conselice C. J., 2000, ApJ, 119, 2645

Best P. N., Kauffmann G., Heckman T. M., Brinchmann J., Charlot S., Ivezić Ž., White S. D. M., 2005, MNRAS, 362, 25

Bharadwaj S., Bhavsar S. P., Sheth J. V., 2004, ApJ, 606, 25

Biermann P. L., Zier C., 2008, A&A, 2, 20

Biviano a., Katgert P., Mazure A., Moles M., Den Hartog R., Perea J., Focardi P., 1997, A&A, 321, 84

Blandford R. D., McKee C. F., 1982, Astrophysics Journal, 255, 419

Blandford R. D., Znajek R. L., 1977, Mon. Not. R. astr. Soc, 179, 433

Blanton M. R., et al., 2017, ApJ, 154, 34

Bojowald M., 2002, Classical and Quantum Gravity, 19, 2717

Bonning E. W., Cheng L., Shields G. A., Salviander S., Gebhardt K., 2007, ApJ, 659, 211

Bournaud F., Dekel A., Teyssier R., Cacciato M., Daddi E., Juneau S., Shankar F., 2011, ApJ, 741, 33

Bouwens R. J., et al., 2014, Astrophysical Journal, 795, 126

Brinchmann J., Charlot S., Heckman T. M., Kauffmann G., Tremonti C., White S. D. M., 2004, MNRAS, 351, 1151

Bruce V. a., Dunlop J. S., Mortlock A., Kocevski D. D., McGrath E. J., Rosario D. J., 2015, MNRAS, 14, 1

Buitrago F., Trujillo I., Conselice C. J., Häußler B., 2013, MNRAS, 428, 1460

Bundy K., Fukugita M., Ellis R. S., Targett T. a., Belli S., Kodama T., 2009, ApJ, 679, 14

Cameron E., 2011, Publications of the Astronomical Society of Australia, 28, 128

Canalizo G., Stockton A., 2001, ApJ, 555, 719

Capak P., et al., 2007, The Astrophysical Journal Supplement Series, 172, 99

Cappelluti N., et al., 2009, Astronomy and Astrophysics, 497, 635

Carpineti A., Kaviraj S., Darg D., Lintott C., Schawinski K., Shabala S., 2012, MNRAS, 420, 2139

Carroll S. M., 2001, Living Reviews in Relativity, 4, 56

Cattaneo a., Combes F., Colombi S., Bertin E., Melchior a. L., 2005, MNRAS, 359, 1237

Cedric L., Shaun C., 1994, MNRAS, 271, 676

Chandrasekhar S., 1931, ApJ, 74, 81

Christian C., Lintott C., Smith A., Fortson L., Bamford S., 2012, Citizen Science: Contributions to Astronomy Research. Vol. 1, Organizations, People and Strategies in Astronomy

Cisternas M., et al., 2011, ApJ, 726, 57

Clowe D., Gonzalez A., Markevitch M., 2004, ApJ, 604, 596

Comastri a., Setti G., Zamorani G., Hasinger G., 1995, A&A, 196, 1

Conselice C. J., 2003, The Astrophysical Journal Supplement Series, 147, 1

Conselice C. J., Bershady M. a., Dickinson M., Papovich C., 2003, ApJ, 126, 1183

Cotini S., Ripamonti E., Caccianiga A., Colpi M., Ceca D. R., Mapelli M., Severgnini P., Segreto A., 2013, MNRAS, 431, 2661

Cowie L. L., Hu E. M., Songaila A., 1995, The Astronomical Journal, 110, 1576

Cox T. J., Jonsson P., Somerville R. S., Primack J. R., Dekel A., 2008, MNRAS, 384, 386

Croton D. J., et al., 2006, Monthly Notices of the Royal Astronomical Society, 365, 11

Darg D. W., et al., 2010a, MNRAS, 401, 1043

Darg D. W., et al., 2010b, MNRAS, 401, 1552

De Bernardis P., et al., 2000, Nature, 404, 955

De Loore C., De Greve J. P., Lamers H. J. G. L. M., 1977, A&A, 61, 251

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Di Matteo P., Combes F., Melchior A.-L., Semelin B., 2007, A&A, 468, 61

Dickinson H., et al., 2018, ApJ, 853, 194

Draper a. R., Ballantyne D. R., 2010, ApJ Lett., 715, 99

Draper a. R., Ballantyne D. R., 2012, ApJ, 753, 37

Dubois Y., et al., 2014, MNRAS, 444, 1453

Dunlop J. S., McLure R. J., Kukula M. J., Baum S. a., O'Dea C. P., Hughes D. H., 2003, MNRAS, 340, 1095

Eales S., et al., 2018, MNRAS, 473, 3507

Easson D. A., Brandenberger R. H., 2001, JHEP, 2001, 1

Easson D. a., Frampton P. H., Smoot G. F., 2011, Phys. Lett. D, 696, 273

Edge a., Sutherland W., Kuijken K., Driver S., McMahon R., Eales S., Emerson J. P., 2013, The Messenger, 154, 32

Einstein A., 1905, Annalen der Physik, 17, 891

Einstein A., 1915, Preussische Akademie der Wissenschaften, Sitzungsberichte, 2, 844

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Ellison S. L., Mendel J. T., Patton D. R., Scudder J. M., 2013, Monthly Notices of the Royal Astronomical Society, 435, 3627

Ellison S. L., Teimoorinia H., Rosario D. J., Trevor Mendel J., 2016, MNRAS, 458, 34

Ellison S. L., Sánchez S. F., Ibarra-Medel H., Antonio B., Mendel J. T., Barrera-Ballesteros J., 2018, MNRAS, 474, 2039

Elvis M., Marengo M., Karovska M., 2002, ApJ, 567, 107

Fabian a. C., 1999, MNRAS, 308, 39

Fabian a. C., 2012, Annual Review of Astronomy and Astrophysics, 50, 455

Faith E. A., 1909, Lick Observatory bulletin, 149, 71

Fakhouri O., Ma C. P., Boylan-Kolchin M., 2010, MNRAS, 406, 11

Fanaroff B. L., Riley J. M., 1974, MNRAS, 167, 31

Fanidakis N., et al., 2012, MNRAS, 419, 2797

Ferrarese L., 2002, ApJ, 578, 90

Ferrarese L., Merritt D., 2000, ApJ, 539, 9

Finn R. A., Impey C. D., 2001, ApJ, 557, 578

Frenk C. S., White S. D. M., 2012, Annalen der Physik, 524, 507

Furlanetto S. R., Briggs F. H., 2004, New Astronomy Reviews, 48, 1039

Gabor J. M., et al., 2009, ApJ, 691, 705

Garriga J., Vilenkin A., Zhang J., 2016, Journal of Cosmology and Astroparticle Physics, 2016, 1

Gavazzi G., Savorgnan G., Fumagalli M., 2011, A&A, 534, 31

Gebhardt K., et al., 2000, The Astrophysical Journal Letters, 543, 5

Genel S., et al., 2014, MNRAS, 445, 175

Goulding A. D., et al., 2017, Publications of the Astronomical Society of Japan, 0, 1

Gregg M. D., Becker R. H., White R. L., Mcmahon R. G., Hook I. M., 1996, ApJ, 112, 407

Grogin N. a., et al., 2005, ApJ, 627, L97

Guth A. H., 1981, Physical Review D, 23, 347

Haring N., Rix H.-W., 2004, ApJ, 604, 89

Harrison C. M., 2017, Nature Astronomy, 1, 1

Hasinger G andCappelluti N., Brunner H., Brusa M., Comastri A., Elvis M., Finoguenov A., Fiore F., 2007, ApJ Supplementary Series, 172, 29

Hawking S., 1974, Nature, 248, 30

Heckman T. M., Smith E. P., Baum S. A., van Breugel W. J. M., Miley G. K., Illingworth G. D., Bothun G. D., Balick B., 1986, ApJ, 311, 526

Hewlett T., Villforth C., Wild V., Mendez-Abreu J., Pawlik M., Rowlands K., 2017, MNRAS, 470, 755

Hickox R. C., et al., 2009, ApJ, 696, 891

Hirano S., Hosokawa T., Yoshida N., Kuiper R., 2017, Science, 357, 1375

Ho L. C., 1999, ApJ, 516, 672

Hocking A., Geach J. E., Sun Y., Davey N., 2017, MNRAS, 22, 1

Holwerda B. W., et al., 2014, ApJ, 781, 19

Hopkins P. F., 2012, MNRAS letters, 420, 8

Hopkins P. F., Hernquist L., 2006, ApJ Supplementary Series, 166, 37

Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJ Supplementary Series, 163, 1

Hopkins P. F., Bundy K., Hernquist L., Ellis R. S., 2007, ApJ, 659, 24

Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, The Astrophysical Journal Supplement Series, 175, 356

Horowitz G. T., Marolf D., 1998, JHEP, 9807, 014

Hubble E. P., 1926, ApJ, 64, 321

Hubble E. P., 1929, Proceedings of the National Academy of Sciences of the United States of America, 15, 168

Isserstedt J., Schindler R., 1986, Astronomy and Astrophysics, 167, 1689

Jeans J. H., 1902, Philosophical Transactions of the Royal Society of London, 199, 1

Jiang L., Fan X., Ivezić Ž., Richards G. T., Schneider D. P., Strauss M. a., Kelly B. C., 2007, ApJ, 656, 680

Jonsson P., 2006, MNRAS, 372, 2

Karouzos M., Jarvis M. J., Bonfield D., 2014, MNRAS, 439, 861

Kaspi S., Smith P. S., Netzer H., Maoz D., Jannuzi B. T., Giveon U., 2000, The Astrophysical Journal, 533, 631

Kaspi S., Smith P. S., Netzer H., Maoz D., Jannuzi B. T., Giveon U., 2005, The Astrophysical Journal, 629, 61

Kauffmann G., Haehnelt M., 2000, MNRAS, 311, 576

Kauffmann G., Heckman T. M., 2009, MNRAS, 397, 135

Kauffmann G., et al., 2003, Monthly Notices of the Royal Astronomical Society, 346, 1055

Kaviraj S., 2013, MNRAS, 440, 2944

Kaviraj S., et al., 2013, Monthly Notices of the Royal Astronomical Society, 428, 925

Kaviraj S., Shabala S. S., Deller A. T., Middelberg E., 2015, MNRAS, 452, 774

Kennicutt Robert C. J., 1989, ApJ, 344, 685

Kennicutt Keel der Hulst V., Hummel 1987, Proceedings of a conference at the California Institute of Technology

Khlopov M. Y., Rubin S. G., Sakharov A. S., 2005, Astroparticle physics, 23, 265

King A., 2003, ApJ, 596, 27

Knapen J. H., Shlosman I., Peletier R. F., 2000, ApJ, 529, 93

Kocevski D. D., et al., 2012, ApJ, 148, 10

Kocevski D. D., et al., 2015, ApJ, 814, 104

Koekemoer a., et al., 2007, Astrophysical Journal, Supplement Series, 172, 196

Kormendy J., Richstone D., 1995, Annual Review of Astronomy and Astrophysics, 33, 581

Koss M., Mushotzky R., Veilleux S., Winter L. M., Baumgartner W., Tueller J., Gehrels N., Valencic L., 2011, ApJ, 739, 57

Kroupa P., 2001, MNRAS, 322, 231

Kuiper G. P., 1938, Astrophysal Journal, 88, 472

Larmor J., 1897, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 44, 503

Lauer T. R., Tremaine S., Richstone D., Faber S. M., 2007, ApJ, 670, 249

Laughlin G., Bodenheimer P., Adams F. C., 1997, ApJ, 482, 420

Lawrence a., 1991, Royal Astronomical Society, 252, 586

Lin D. N. C., Pringle J. E., 1987, ApJ, 320, 4

Loeb A., Rasio F. A., 1994, ApJ, 432, 52

Lotz J. M., Primack J., Madau P., 2004, ApJ, 128, 163

Lotz J. M., Jonsson P, Cox T. J., Primack J. R., 2010a, MNRAS, 404, 575

Lotz J. M., Jonsson P, Cox T. J., Primack J. R., 2010b, MNRAS, 404, 590

Lusso E., et al., 2009, A&A, 34, 14

Maccarone T. J., Kundu A., Zepf S. E., Rhode K. L., 2007, Nature, 445, 183

Maldacena J., 1999, International Journal of Theoretical Physics, 38, 1113

Malin D. F., Carter D., 1983, ApJ, 274, 534

Marconi A., Hunt L. K., 2003, ApJ, 589, 21

Massey R., Stoughton C., Leauthaud A., Rhodes J., Koekemoer A., Ellis R., Shaghoulian E., 2010, MNRAS, 401, 371

McConnell N. J., Ma C.-P., 2013, arxiv, 1211, 2816

Mechtley M., et al., 2016, The Astrophysical Journal, 830, 1

Mellier Y., Mathez G., 1987, Astronomy and Astrophysics, 175, 3

Meurs E. J. A., Wilson A. S., 1984, A&A, 136, 206

Mihos J. C., Hernquist L., 1994, ApJ, 437, 47

Mihos J. C., Hernquist L., 1996, ApJ, 464, 641

Mobasher B., Capak P., Scoville N. Z., Dahlen T., Salvato M., 2007, The Astrophysical Journal Supplement Series, 172, 117

Mortlock D. J., et al., 2011, Nature, 474, 616

Mullaney J. R., et al., 2012, MNRAS, 419, 95

Mullaney J. R., Alexander D. M., Fine S., Goulding a. D., Harrison C. M., Hickox R. C., 2013, Monthly Notices of the Royal Astronomical Society, 433, 622

Mumford S. J., et al., 2015, Computational Science and Discovery, 8, 014009

Nelson D., et al., 2015, Astronomy and Computing, 13, 12

Nenkova M., Sirocky M. M., Nikutta R., Ivezic Z., Elitzur M., 2008, ApJ, 685, 160

Nesvadba P. H. N., Lehnert M. D., 2008, A&A, 491, 407

Nesvadba N. P. H., Lehnert M. D., Eisenhauer F., Gilbert a., Tecza M., Abuter R., 2006, The Astrophysical Journal, 650, 693

Netzer H., Trakhtenbrot B., 2007, ApJ, 654, 9

Oesch P. A., et al., 2016, ApJ, 819, 129

Oh S., Oh K., Yi S. K., 2012, The Astrophysical Journal Supplement Series, 198, 4

Omukai K., Nishi R., 1998, ApJ, 508, 141

Pawlik M. M., Wild V., Walcher C. J., Johansson P. H., Villforth C., Rowlands K., Mendez-Abreu J., Hewlett T., 2016, MNRAS, 20, 1

Peebles P. J. E., 1972, ApJ, 178, 371

Peirani S., Crockett R. M., Geen S., Khochfar S., Kaviraj S., Silk J., 2010, MNRAS, 405, 2327

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, ApJ, 124, 29

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2009, ApJ, 139, 41

Penrose R., 2012, AIP Conference Proceedings, 1446, 233

Peterson B., 2006, The Broad-Line Region in Active Galactic Nuclei. Vol. 693, Springer, Berlin, Heidelberg

Peterson B. M., Ferrarese L., Gilbert K. M., Kaspi S., Malkan M. a., 2004, ApJ, 613, 682

Planck-Collaboration 2016, A&A, 594, 1

Pogge 1988, ApJ, 328, 519

Pounds K. A., Reeves J. N., King A. R., Page K. L., Brien P. T. O., Turner M. J. L., 2003, MNRAS, 713, 705

Press W. H., Schechter P., 1974, ApJ, 187, 425

Rafferty D. a., McNamara B. R., Nulsen P. E. J., Wise M. W., 2006, ApJ, 652, 216

Ramos Almeida C., Tadhunter C. N., Inskip K. J., Morganti R., Holt J., Dicken D., 2011, MNRAS, 410, 1550

Ranalli P., Comastri A., Setti G., 2003, Astronomy and Astrophysics, 399, 39

Reynolds C., 2013, Classical and Quantum Gravity, 30, 22

Riess A. G., et al., 1998, The Astronomical Journal, 116, 1009

Risaliti G., Maiolino R., Salvati M., 1999, ApJ, 522, 157

Rix H.-W., Zaritsky D., 1995, ApJ, 447, 82

Robertson B., Hernquist L., Cox T. J., Di Matteo T., Hopkins P. F., Martini P., Springel V., 2006a, ApJ, 641, 90

Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006b, ApJ, 645, 986

Rodriguez-Gomez V., et al., 2015a, Monthly Notices of the Royal Astronomical Society, 64, 49

Rodriguez-Gomez V., et al., 2015b, MNRAS, 64, 49

Rubin V. C., Ford W. Kent J., 1970, ApJ, 159, 379

Sabater J., Best P. N., Argudo-fernández M., 2013, MNRAS, 430, 638

Sabater J., Best P. N., Heckman T. M., 2015, MNRAS, 447, 110

Salpeter E. E., 1955, Astrophysical Journal, 121, 161

Sanders D. B., Soifer B. T., Elias J. H., Madore B. F., Matthews K., Neugebauer G., Scoville N. Z., 1988, ApJ, 325, 74

Satyapal S., Ellison S. L., McAlpine W., Hickox R. C., Patton D. R., Trevor Mendel J., 2014, MNRAS, 441, 1297

Schawinski K., Urry C. M., Virani S., Coppi P., Bamford S. P., 2010, ApJ, 711, 284

Schawinski K., Treister E., Urry C. M., Cardamone C. N., Simmons B., Yi S. K., 2011, ApJ, 727, L31

Schawinski K., Simmons B. D., Urr C. M., Treister E., Glikman E., 2012, MNRAS letters, 425

Schaye J., et al., 2015, MNRAS, 446, 521

Schmidt M., 1963, Nature, 197, 1040

Schödel R., et al., 2002, Nature, 419, 694

Scoville N., Aussel H., Brusa M., Capak P., Carollo C. M., 2007, The Astrophysical Journal Supplement Series, 172, 1

Seyfert C. K., 1943, ApJ, 97, 28

Shabala S. S., Deller A., Kaviraj S., Middelberg E., Turner R. J., Ting Y. S., Allison J. R., Davis T. a., 2017, MNRAS, 464, 4706

Shade D., 1995, ApJ

Shankar F., Bernardi M., Haiman Z., 2009, ApJ, 694, 867

Shankar F., Weinberg D. H., Miralda-Escudé J., 2013, MNRAS, 428, 421

Shen Y., et al., 2011, ApJ Supplementary Series, 194, 45

Shen Y., et al., 2015, ApJ, 805, 96

Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, MNRAS, 452, 575

Silk J., Mamon G. A., 2012, Research in Astronomy and Astrophysics, 12, 917

Silk J., Rees M. J., 1998, Growth Lakeland, 4, 4

Skillman S. W., Warren M. S., Turk M. J., Wechsler R. H., Holz D. E., Sutter P. M., 2014, arXiv:1407.2600v1, p. 26

Smolin L., 1991, Classical and Quantum Gravity, 9, 173

Snyder G. F., et al., 2015, MNRAS, 454, 1886

Soltan A., 1982, MNRAS, 200, 115

Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087

Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Steinhardt C. L., Elvis M., 2010, MNRAS, 402, 2637

Stern D., Djorgovski S. G., Perley R. a., de Carvalho R. R., Wall J. V., 2000, ApJ, 119, 1526

Stern D., et al., 2012, ApJ, 753, 30

Stewart K. R., Bullock J. S., Barton E. J., Wechsler R. H., 2009, ApJ, 702, 1005

T. L., D. S., 2001, A&A, 366, 538

Tadhunter 1989, Nature

Tadhunter C., 2016, A&A Review, 24, 10

Tegmark M., Silk J., Rees M. J., Blanchard A., Abel T., Palla F., 1997, ApJ, 474, 1

Tinsley L., 1978, ApJ, 219, 46

Toomre A., Toomre J., 1972, ApJ, 178, 623

Torrey P., et al., 2015, MNRAS, 447, 2753

Treister E., et al., 2004, ApJ, 616, 123

Treister E., Schawinski K., Urry C. M., Simmons B. D., 2012, The Astrophysical Journal, 758, L39

Tremaine S., et al., 2002, ApJ, 574, 740

Tremonti C. a., et al., 2004, ApJ, 613, 898

Trujillo I., Conselice C. J., Bundy K., Cooper M. C., Eisenhardt P., Ellis R. S., 2007, MNRAS, 382, 109

Trump J. R., et al., 2014, ApJ, 793, 101

Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886

Urrutia T., Lacy M., Becker R. H., 2008, ApJ, 674, 80

Verlinde E., 2017, SciPost Phys. 2, 016, 1

Villforth C., et al., 2014, MNRAS, 439, 3342

Villforth C., et al., 2017, MNRAS, 466, 812

Vogelsberger M., et al., 2014a, MNRAS, 444, 1518

Vogelsberger M., et al., 2014b, Nature, 509, 177

Volonteri M., 2010, Astronomy and Astrophysics Review, 18, 279

Volonteri M., Haardt F., Madau P., 2002, ApJ, 582, 559

Wild V., Heckman T., Charlot S., 2010, MNRAS, 405, 933

Wilson R. W., Penzias A. A., 1967, Science, 156, 1100

Worrall D. M., 2009, A&A Review, 17, 1

Wright E. L., et al., 2010, Astronomical Journal, 140, 1868

Wu X.-B., et al., 2015, Nature, 518, 512

Yunes N., Siemens X., 2013, Living Reviews in Relativity, 16, 124

Zhang W., Woosley S. E., Heger A., 2008, The Astrophysical Journal, 679, 639

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

van den Bergh S., Abraham R. G., Ellis R. S., Tanvir N. R., Santiago B. X., Glazebrook K. G., 1996, The Astronomical Journal, 112, 359