1 INTRODUCTION

In 1963, the powerful radio source 3C 273 was identified with a star-like, thirteenth magnitude object with a strongly redshifted ($z = 0.158$) optical spectrum (Schmidt, 1963). Assuming this redshift was due to the Hubble expansion of the Universe, this finding implied 3C 273 was at an enormous distance (500 mega-parsecs) and was 10 times brighter than the most luminous galaxies. Variability on time-scales of weeks suggested that the source was very compact (i.e. smaller than a light-week). It was quickly realised that 'quasars'[1], and other lower-luminosity classes of active galactic nuclei (AGN)[2], are powered by the release of gravitational potential energy as mass is accreted onto a super-massive black hole (BH) at the centre of a galaxy (e.g. Hoyle and Fowler, 1963; Salpeter, 1964; Lynden-Bell, 1969; Lynden-Bell and Rees, 1971).

$z = \frac{\Delta\lambda}{\lambda_0}$

Super-massive: $10^{6-9}\ M_\odot$

Begining in the early 1990s, inactive super-massive BHs were found in the centres of many nearby massive galaxies (e.g. Kormendy and Richstone, 1995; Ferrarese and Ford, 2005; Kormendy and Ho, 2013). This proved that quasar activity is a stage in the life of all massive galaxies (e.g. Lynden-Bell, 1969). Shortly after, it was discovered that the BH mass and the properties of the host-galaxy bulge were strongly correlated (e.g. the M_{BH}-σ relation; Ferrarese and Merritt, 2000; Gebhardt et al., 2000; Graham et al., 2001; Tremaine et al., 2002; Marconi and Hunt, 2003; Aller and Richstone, 2007; Gültekin et al., 2009). This was an unexpected finding, given that the sphere-of-influence of a BH, $\lesssim$ 100 parcsecs (e.g. Kormendy and Ho, 2013), is many orders of magnitude smaller than a typical galactic bulge. It suggests that the BH and the host galaxy bulge grow synchronously, with the energetic output of the rapidly-accreting BH coupling with the gas in the host galaxy and regulating star formation and the growth of the BH itself (e.g. Silk

σ: stellar velocity dispersion of host galaxy bulge

Sphere-of-influence: where the gravity of the BH dominates over the other mass components

1 The term 'quasar' originated as a contraction of 'quasi-stellar radio source', although 90 per cent of quasars are now known to radio-quiet

2 Throughout this thesis we use the terms 'quasar' and 'Active Galactic Nucleus (AGN)' interchangeably to describe active super-massive black holes, although the term quasar is generally reserved for the luminous ($L_{Bol} > 10^{12} L_\odot$) subset of AGNs.

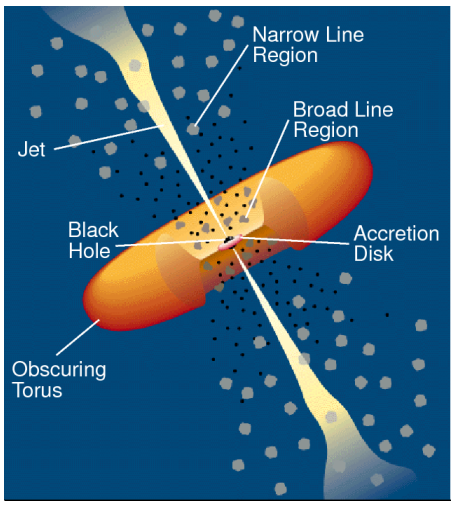

Figure 1.1: Cartoon picture of the inner regions on an AGN. Credit: Urry and Padovani, (1995).

and Rees, 1998; King, 2003; Di Matteo, Springel, and Hernquist, 2005; King and Pounds, 2015).

The number density of quasars, which evolves strongly with redshift, peaks at redshifts $2 \lesssim z \lesssim 3$ (e.g. Brandt and Hasinger, 2005; Richards et al., 2006) and the most massive ($M_{BH} \gtrsim 10^9 M_\odot$) present-day BHs experienced much of their growth during this epoch. The cosmic star formation rate history closely follows the cosmological evolution of the quasar luminosity function (e.g. Boyle and Terlevich, 1998), which establishes a further connection between BH and galaxy properties. Quasar feedback has also been invoked to reproduce the high-mass end of the galaxy luminosity function in cosmological simulations (e.g. Kauffmann and Haehnelt, 2000). The insight that quasars may play a crucial role in the evolution of galaxies has led to an explosion of interest in their properties in recent years.

1.1 AGN: CURRENT PARADIGM

The basic features of the current AGN paradigm are widely accepted, although many of the details are unknown. The basic features are: a hot accretion disc surrounding a super-massive BH, rapidly orbiting clouds of ionised gas, and a dusty, obscuring structure (generally referred to as the 'torus'). Collimated jets of relativistic plasma and/or associated lobes are also seen in the 10 per cent of quasars that are radio-loud (e.g. Peterson, 1997). A cartoon picture illustrating the basic structure of an AGN is shown in Figure 1.1.

1.1.1 *Accretion disc*

Material is pulled towards a super-massive BH and sheds angular momentum through viscous and turbulent processes in a hot accretion disc (e.g. Begelman, 1985). The accretion disc reaches temperatures of $\sim 10^6$ K, and radiates primarily at ultraviolet (UV) to soft-X-ray wavelengths.

1.1.2 *The broad line region*

One of the pre-eminent features of many AGN spectra are broad optical and UV emission lines produced in the broad line region (BLR). The BLR consists of gas clouds at distances from several light-days to several light-months that are photoionised by the ultraviolet continuum emission emanating from the accretion disc. Because of the close proximity to the central super-massive BH, bulk motions are dominated by gravity and radiation pressure from the accretion disc. The very broad emission line widths are assumed to Doppler-broadened, and imply line-of-sight velocities of many thousands of $\mathrm{km\,s^{-1}}$.

1.1.3 *The dusty torus*

Further out from the central engine are dusty, molecular clouds which are co-planar with the accretion disc. These dusty clouds are generally referred to as the 'torus'. In a Type II AGN, the system is observed in an edge-on configuration and, as a result, emission from the accretion and BLR is obscured by the dusty torus (e.g. Antonucci, 1993). Although this simple picture (shown in Figure 1.1 as well as in countless other publications) is a useful starting point, the idea of a torus as a static, doughnut-like structure is almost certainly incorrect. For example, the problem of maintaining the large scale height required to explain the observed fraction of Type 1/2 AGN has long been recognized (e.g. Krolik and Begelman, 1988). In one set of models, the torus is the dusty part of an accretion disc wind that extends beyond the dust-sublimation radius (e.g. Konigl and Kartje, 1994; Everett, Gallagher, and Keating, 2009; Gallagher et al., 2012; Everett, 2005; Keating et al., 2012; Elitzur and Shlosman, 2006).

1.1.4 *The narrow line region*

Further away from the central BH and beyond the dusty torus is the narrow emission line region (NLR). Like the BLR, the NLR is ionised by radiation from the accretion disc. Unlike the BLR, densities in the NLR are low enough that forbidden transitions are not collisionally suppressed. Emission line widths are typically hundreds of $\mathrm{km\,s^{-1}}$in the NLR.

1.2 WINDS AND OUTFLOWS IN AGN

Quasars are very powerful sources of radiation, and are embedded in matter-rich environments at the centres of galaxies. Strong winds, driven by some combination of gas pressure, radiation pressure, and magnetic forces, are to be expected under these conditions (e.g. Blandford and Payne, 1982; Proga, Stone, and Kallman, 2000; Everett, 2005). In line with these expectations, evidence for outflowing gas is common in the spectra of quasars.

Perhaps the most dramatic evidence of outflows is seen in broad absorption line quasars (BALQSOs; Weymann et al., 1991). BALQSOs are characterised by broad absorption features in the ultra-violet resonance lines of highly ionised N v, C iv and Si iv. The absorption is generally interpreted as evidence for outflows with velocities reaching 60 000 $\mathrm{km\,s^{-1}}$(e.g. Turnshek, 1988). The far-side of the outflow is obscured by the accretion disc, and so only the near-side, which is moving towards the observer, is detected. Therefore, the absorption lines that are detected are always blueshifted. The observed C iv BALQSO fraction in radio-quiet quasars is $\sim$ 15 per cent (e.g. Hewett and Foltz, 2003; Reichard et al., 2003) and the intrinsic fraction in the quasar population has been estimated at $\sim$ 40 per cent (Allen et al., 2011). Outflows can also explain narrow UV and X-ray absorption lines (NALs) which are seen in $\sim$ 60 per cent of Seyfert 1 galaxies (Crenshaw et al., 1999) and some quasars (e.g. Hamann et al., 1997).

Seyfert 1: Class of AGN with broad emission lines and clearly detectable host galaxies

The blueshifting of high-ionisation lines in the BLR (including C iv) can be understood if the lines are produced in outflowing clouds (although see, e.g., Gaskell and Goosmann 2016, for an alternative explanation). The blueshifting of C iv appears to be nearly ubiquitous in the quasar population (e.g. Richards et al., 2002; Richards et al., 2011).

Together, these results suggest that outflows are very common and the energy released by quasars can have a dramatic effect on their immediate surroundings. Accretion-disc wind models have been developed to explain the wide range of emission and absorption line phenomena described above (e.g. Murray et al., 1995; Elvis, 2000; Proga, Stone, and Kallman, 2000; Everett, 2005). UV photons excite the partially-ionised material surrounding the accretion disc. This exerts a pressure on the atoms - a phenomenon known as radiation line-driving - and a wind is blown from the accretion disc.

In models for the co-evolution of quasars and galaxies, the energy released by quasars impacts galaxies on much larger scales than is probed by the emission and absorption diagnostics described above. In recent years, a huge amount of resources have been devoted to searching for observational evidence of galaxy-wide, quasar-driven outflows (for recent reviews, see Alexander and Hickox, 2012; Fabian, 2012; Heckman and Best, 2014). This has resulted in recent detections of outflows in AGN-host galaxies using tracers of atomic, molecular, and ionised gas with enough power to sweep their host galaxies clear of gas (e.g. Nesvadba et al., 2006; Arav et al., 2008; Nesvadba et al., 2008; Moe et al., 2009; Dunn et al., 2010; Alexander et al., 2010; Harrison et al., 2012; Harrison et al., 2014; Nesvadba et al., 2010; Rupke and Veilleux, 2013; Veilleux et al., 2013; Nardini et al., 2015; Feruglio et al., 2010; Alatalo et al., 2011; Cimatti et al., 2013; Cicone et al., 2014).

1.3 MEASURING BLACK HOLE MASSES

The BH mass is one of the most important physical parameters of a quasar and considerable resources have been devoted to measuring the masses of BHs in active galaxies. Large-scale studies of AGN and quasar demographics have become possible through the calibration of single-epoch virial-mass estimators using BH masses from reverberation-mapping (e.g. Peterson, 2010; Vestergaard et al., 2011; Marziani and Sulentic, 2012; Shen, 2013).

1.3.1 *Reverberation mapping*

Under the assumptions that the BLR dynamics are virialised and the gravitational potential is dominated by the BH, the BH

Virial theorem: average kinetic energy equals half of the average negative potential energy

mass is given by:

$$M_{BH} = \frac{V_{Virial}^2 R_{BLR}}{G} \tag{1.1}$$

where V_{virial} is the virial velocity in the BLR and R_{BLR} the characteristic BLR radius. The problem of measuring the mass therefore reduces to the problem of measuring the velocity and orbital radius of the line-emitting clouds in the BLR.

Continuum variability is a common characteristic of quasars, owing to the stochastic nature of the accretion process. Because the BLR is photo-ionized by the continuum, the broad emission lines also vary with some characteristic lag, which is related to the light travel time across the BLR. The reverberation mapping method, first proposed by Blandford and McKee, (1982), uses the time lag between variations in the continuum emission and correlated variations in the broad line emission to measure the typical size of the BLR (e.g. Peterson, 1993; Netzer and Peterson, 1997; Peterson, 2014).

The typical velocity in the BLR is measured from the Doppler-broadened width of an emission line produced in the BLR. Since the structure and geometry of the BLR is unknown, a virial coefficient, f, is introduced to transform the observed line-of-sight velocity inferred from the line width in to a virial velocity. Unfortunately, f is unknown and likely varies from object to object. In practice, the value of f is empirically determined by requiring that the reverberation-mapping masses are consistent with those predicted from the M_{BH}-σ relation for local inactive galaxies.

Because reverberation mapping depends on temporal resolution rather than spatial resolution, this technique can be applied out to much greater distances than direct dynamical modelling (e.g. Kormendy and Ho, 2013). However, because reverberation mapping relies on dense spectrophotometric monitoring campaigns which span many yearsthe number of AGN with measured lags is limited to $\sim$ 50 AGN (e.g. Kaspi et al., 2000; Peterson et al., 2004; Kaspi et al., 2007; Bentz et al., 2009; Denney et al., 2010; Barth et al., 2011; Grier et al., 2012). This sample is strongly biased to low luminosity Seyfert 1 galaxies, and the maximum redshift is just $z \sim 0.3$. Comprehensive statistical studies of active BHs, particularly during the epoch of peak galaxy formation ($z \gtrsim 2$), therefore require a different approach to measuring BH masses.

1.3.2 *Single-epoch virial estimates*

Reverberation mapping campaigns have also revealed a tight relationship between the radius of the BLR and the quasar optical (or ultraviolet) luminosity (the $R_{BLR} - L$ relation; e.g. Kaspi et al., 2000; Kaspi et al., 2007). A slope of $\simeq 0.5$ is found, which is consistent with naive predictions (e.g. Peterson, 1997). This relation provides a much less expensive method of measuring the BLR radius, and large-scale studies of AGN and quasar demographics have thus become possible through the calibration of single-epoch virial-mass estimators (e.g. Vestergaard, 2002; McLure and Jarvis, 2002; Vestergaard and Peterson, 2006; McGill et al., 2008; Wang et al., 2009; Rafiee and Hall, 2011; Park et al., 2013). Single-epoch virial BH mass estimates normally take the form

$$M_{BH} = 10^{a} \left(\frac{\Delta V}{1000 \text{ km s}^{-1}} \right)^{b} \left(\frac{L_{\lambda}}{10^{44} \text{ erg s}^{-1}} \right)^{c} \tag{1.2}$$

where ΔV is a measure of the line width, L_{λ} is the monochromatic continuum luminosity at wavelength λ, and a, b, and c are coefficients, determined via calibration against a sample of AGN with reverberation-mapping BH mass estimates.

Virial BH masses have made possible large-scale studies of AGN and quasar demographics (e.g. Greene and Ho, 2005; Vestergaard and Peterson, 2006; Vestergaard and Osmer, 2009; Shen et al., 2011; Shen and Liu, 2012; Trakhtenbrot and Netzer, 2012). With these masses, the growth rate of massive BHs can be measured across cosmic time. This conveys important information about the accretion processes occuring in active BHs (e.g. Kollmeier et al., 2006) and is crucial in order to understand the processes responsible for establishing the M_{BH}-σ relation (e.g. Bennert et al., 2011). Single-epoch virial estimates have been used to calculate BH masses in the highest redshift quasars (e.g. a $10^9 M_{\odot}$ BH in a redshift $z \sim 7$ quasar; Mortlock et al., 2011). The fact that a $10^9 M_{\odot}$ BH exits when the age of the Universe is less than 1 Gyr places tight constraints on its accretion history (e.g. Willott, McLure, and Jarvis, 2003).

The uncertainties in reverberation mapped BH masses are estimated to be ~ 0.4 dex (e.g. Peterson, 2010), and the uncertainties in virial masses are similar (e.g. Vestergaard and Peterson, 2006). However, the main concern and biggest unknown is the extension of the method to high redshifts. This requires that

the relations calibrated for sub-Eddington BHs with M$\sim 10^7$ are valid for BHs with masses up to 10^{10} that are radiating near the Eddington limit. Furthermore, the vast majority of these measurements are done using Hβ, and so the $R-L$ relation has been established using Hβ. Hβ is redshift beyond the reach of optical spectrographs at $z \gtrsim$ XX, and extending the method to higher redshifts requires the secondary-calibration of other low-ionization emission lines such as Hα (6564.6Å) and Mg II$\lambda\lambda$2796.4,2803.5 (e.g. Vestergaard, 2002; McLure and Jarvis, 2002; Wu et al., 2004; Kollmeier et al., 2006; Onken and Kollmeier, 2008; Wang et al., 2009; Rafiee and Hall, 2011). At redshifts $z \gtrsim 2$, only the high-ionisation C IV emission line is available, but many authors have reported C IV-based masses to be inconsistent with masses based on other lines. These issues are explored in detail in Chapter **??**.

1.4 THE ADVENT OF LARGE SURVEYS

The Palomar-Green (PG) Bright Quasar Survey (BQS; Schmidt and Green, 1983), the first large-area quasar survey, identified 114 quasars via their UV excess. Boroson and Green, (1992) were among the first to use this large sample to analyse quasar spectroscopic properties in a systematic way. With the advent of CCD technology came a new generation of surveys, most notably the Sloan Digital Sky Survey (SDSS; York et al., 2000) and the 2QZ survey (Croom et al., 2004). SDSS, and the next-generation Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al., 2013), now contain spectra of $\sim$ 200 000 AGN and quasars. These large, uniform data sets have revolutionised the study of AGNs and quasars by facilitating statistical studies of the properties of quasars and AGNs covering wide ranges in redshift and luminosity. This has greatly enhanced our understanding of AGNs and their connections to their host galaxies, especially when combined with multi-wavelength coverage. This has been achieved in recent years with sensitive, wide-field photometric surveys, including the UV/optical SDSS, the near-infrared UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al., 2007) and the mid-infrared Wide-field Infrared Explorer (WISE; Wright et al., 2010).

1.5 OVERVIEW OF THESIS

1.5.1 *Chapter 1: A near-infrared spectroscopic database of high-redshift quasars*

With optical spectra, such as those provided by SDSS, we can estimate BH masses and measure properties of outflows. However, optical lines including Hβ and [O III] are shifted to infrared wavelengths at redshifts $z \gtrsim 1$. A complete understanding of quasars during the peak epoch of galaxy formation $z \sim 2$ requires near-infrared spectroscopy. Spectroscopic observations are challenging at infrared wavelengths, and there are far fewer infrared observations of quasars than optical ones. However, we have constructed a catalogue containing 462 high-redshift quasars with near-infrared spectra. This is the largest sample of its kind, and has facilitated the investigations described in Chapters **??** and **??**.

1.5.2 *Chapter 2: Correcting C IV-Based Virial Black Hole Masses*

C IV has long been known to exhibit significant displacements to the blue and these 'blueshifts' almost certainly signal the presence of strong outflows. As a consequence, single-epoch virial black hole (BH) mass estimates derived from C IV velocity-widths are known to be systematically biased compared to masses from the hydrogen Balmer lines. Using a large sample of 230 high-luminosity ($L_{Bol} = 10^{45.5} - 10^{48}$ erg s^{-1}), redshift $1.5 < z < 4.0$ quasars with both C IV and Balmer line spectra, we have quantified the bias in C IV BH masses as a function of the C IV blueshift. C IV BH masses are shown to be a factor of five larger than the corresponding Balmer-line masses at C IV blueshifts of 3000km s^{-1}and are over-estimated by almost an order of magnitude at the most extreme blueshifts, $\gtrsim$ 5000km s^{-1}. Using the monotonically increasing relationship between the C IV blueshift and the mass ratio BH(C IV)/BH(Hα), we derive an empirical correction to all C IV BH-masses. The scatter between the corrected C IV masses and the Balmer masses is 0.24 dex at low C IV blueshifts ($\sim$ 0km s^{-1}) and just 0.10 dex at high blueshifts ($\sim$ 3000km s^{-1}), compared to 0.40 dex before the correction. The correction depends only on the C IV line properties - i.e. FWHM and blueshift - and can therefore be applied to all quasars where C IV emission line properties have been measured, enabling the derivation of un-biased virial BH mass es-

FWHM: full-width at half-maximum

timates for the majority of high-luminosity, high-redshift, spectroscopically confirmed quasars in the literature.

1.5.3 *Chapter 3: Quasar-drive outflows in the narrow line region*

The advent of large optical spectroscopic surveys (e.g. SDSS) have facilitated studies of the NLR in tens of thousands of AGN which have provided constraints on the prevalence and drivers of ionised outflows. Because of its high equivalent width, [O III]λλ4960,5008 is the most studied of the narrow AGN emission-lines. By following [O III] to near-infrared wavlengths, we are able to extend these investiations to the redshift range when star formation and BH accretion peaked. We analyse the [O III] properties of a sample of 354 high-luminosity, redshift $1.5 < z < 4$ quasars. To date, this is the largest study of the NLR properties of high redshift quasars.

In our sample, there is a huge diversity in [O III] emission properties. The mean and standard deviation of the line width (characterized by w_{80}) is 1535 ± 562 km s^{-1}, which is significantly broader than is found in lower-luminosity AGN. We also find blue-assymetries in the [O III] emission to be stronger at higher luminosities. The fraction of objects with no or very weak [O III] is an order of magnitude larger than in a sample of low-luminosity SDSS AGN. However, among the objects which do have strong [O III] emission we do not observe any correlation between the [O III] the [O III] EQW and the quasar luminosity (the 'Baldwin' effect), despite a dynamic range in luminosity spanning three orders of magnitude.

We confirm earlier results using much smaller samples that the EV1 correlations found in low-luminosity AGN also exist in high-redshift quasars. The C IV EQW-blueshift parameter space similarly spans the diversity of broad emission-line properties in high redshift quasars. We find a strong anti-correlation between the C IV blueshift and the [O III] EQW, which directly connects the low-redshift EV1 and high-redshift C IV parameter spaces. One possible explanation for this correlation is that the NLR gas is being swept away on relatively short time-scales by quasar-driven outflows. In quasars for which [O III] is detected, we find that the blueshifting of [O III] and C IV are correlated. This establishes a connection between quasar-driven outflows in the broad and narrow line regions. Finally, we demonstrate that Independent Component Analysis can be used to extend these results to spectra with lower signal-to-noise.

1.5.4 *Chapter 4: Outflows and hot dust emission*

To first order, AGNs have remarkably similar SEDs over many decades in luminosity and redshift. Using photometric data from SDSS, UKIDSS and WISE, we build a simple parametric SED model that is able to reproduce the median colours of tens of thousands of AGN with a large dynamic range in redshift and luminosity. On the other hand, the SED properties of individual quasars do vary significantly. In particular, the spread at $1-2$ μm is large and suggests the presence of real variation in the hot dust temperature and luminosity among the quasars. We parameterise the shape of the near-infrared SED using a single blackbody, which we use to find the temperature and relative luminosity of the hot dust. We find no correlation between the hot dust properties and the redshift, BH mass and Eddington ratio. However, we do find a significant correlation between the amount of hot dust and the strength of outflows in the quasar BLR. In a sample of low-redshift AGN, we find that the hot dust abundance is strongly correlated with EV1 trends. We consider the implications of these results in the context of accretion disc wind models (e.g. Elitzur and Shlosman, 2006).

Throughout this thesis, we adopt a ΛCDM cosmology with $h_0 = 0.71$, $\Omega_M = 0.27$, and $\Omega_\Lambda = 0.73$. All wavelengths and equivalent width measurements are given in the quasar rest-frame, and all emission line wavelengths are given as measured in vacuum.