1 INTRODUCTION

1.1 DISCOVERY

In 1963, the powerful radio source 3C 273 was identified with a star-like, thirteenth magnitude object with a strongly redshifted[1] optical spectrum (Schmidt, 1963). This finding implied 3C 273 was at an enormous distance (at least by the standards of the time). For an object this distant to appear so bright it must be extremely luminous: 10 times more luminous, in fact, that the largest galaxies known. At the same time, its rapid variability meant that it could be no bigger than a light-week across. 3C 273 was therefore a new kind of exotic object, at the edge of the known Universe. Understandably, its discovery caused a lot of excitement, both for astronomers and the wider public. It was quickly realised that 'quasars'[2], and other lower-luminosity classes of active galactic nuclei (AGN)[3], are powered by the release of gravitational potential energy as mass is accreted onto a super-massive[4] 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).

1.2 THE AGN-HOST GALAXY CONNECTION

Beginning 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, rather than being rare and exotic objects, quasar activity is in fact a stage in the life of all massive galaxies (e.g. Lynden-Bell, 1969). Shortly after, it was discovered that the BH mass was tightly correlated with properties of the host-galaxy bulge (e.g. the stellar velocity dis-

1 $z = c(\Delta\lambda/\lambda) = 0.158$, where c is the speed of light.

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

3 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.

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

persion, σ, which is proportional to the bulge mass; e.g. 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[5] of a BH ($\lesssim 100$ parsecs; e.g. Kormendy and Ho, 2013), is many orders of magnitude smaller than the dimensions of a typical galactic bulge. The tight correlation between their masses suggested that the BH and the host-galaxy bulge grow synchronously. Both the density of quasars and the cosmic star formation history evolve strongly with redshift and peak at $2 \lesssim z \lesssim 3$ (e.g. Boyle and Terlevich, 1998; Brandt and Hasinger, 2005; Richards et al., 2006). The similarity of their cosmic evolution is further evidence for the existence of an intimate connection between quasars and their host galaxies.

In a currently favoured model, rapid BH fuelling and star-formation are triggered by a gas-rich galaxy merger (e.g. Hopkins et al., 2006), satellite accretion or secular processes (e.g. Fanidakis et al., 2012). The energetic output of the rapidly-accreting BH couples with the gas in the host-galaxy and regulates star formation and the growth of the BH itself (e.g. Silk and Rees, 1998; King, 2003; Di Matteo, Springel, and Hernquist, 2005; King and Pounds, 2015). This process, which is referred to as 'quasar feedback', is also commonly 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.

1.3 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.

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

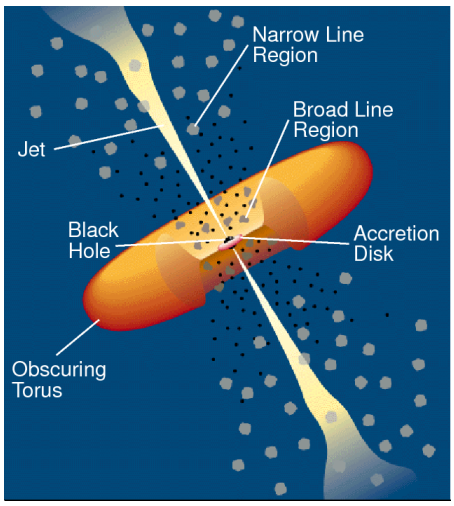

Figure 1.1: Cartoon picture of the inner regions on an AGN. Credit: Urry and Padovani, (1995).

1.3.1 *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 to soft-X-ray wavelengths.

1.3.2 *The broad line region*

One of the pre-eminent features of many AGN spectra are broad optical and ultraviolet 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 photo-ionised 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 $\mathrm{km\,s^{-1}}$.

1.3.3 *The dusty torus*

Further out from the central engine on parsec-scales are dusty, molecular clouds which are co-planar with the accretion disc. These dusty clouds are generally referred to as the 'torus'. The torus is a central feature of orientation-based unification

schemes (e.g. Antonucci, 1993) in which Type I and Type II AGN (i.e. AGN which possess/lack broad emission-lines) are intrinsically identical objects that are viewed in different orientations. In a Type II AGN, the system is observed in an edge-on configuration and, as a result, emission from the accretion disc and BLR is obscured by the dusty torus. 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 a gross over-simplification. 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 more realistic 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).

1.3.4 *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.

1.4 THE ADVENT OF LARGE SURVEYS AND THE SYSTEMATIC STUDY OF AGN PROPERTIES

The Palomar-Green (PG) Bright Quasar Survey (BQS; Schmidt and Green, 1983), the first large-area quasar survey, identified 114 quasars via their ultraviolet brightness. Boroson and Green, (1992) were among the first to use the PG quasar sample to analyse quasar spectroscopic properties in a systematic way. In their landmark study, they used a principle component analysis (PCA) to identify the features responsible for the largest variance in quasar spectra. The first eigenvector of their PCA decomposition - generally referred to as 'eigenvector 1' (EV1) - is correlated with the FWHM of the broad Hβ emission line

and the relative strengths of optical Fe II and Hβ. The underlying driver behind EV1 is thought to be the Eddington ratio[6].

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 $\sim$ 300 000 AGN and quasars. These large, uniform data sets have revolutionised the study of AGNs and quasars by facilitating statistical studies of their properties covering wide ranges in redshift and luminosity.

Emission features in the rest-frame optical band are extremely rich diagnostics of quasar properties. However, at $z \sim 2$, rest-frame optical lines are redshifted to near-infrared wavelengths. Near-infrared spectroscopy is therefore essential for a complete understanding of quasars during the peak epoch of galaxy formation. Spectroscopic observations are challenging at infrared wavelengths, and there are far fewer infrared observations of quasars than optical ones. However, in Chapter **??** of this thesis, I describe how I have constructed a catalogue containing 462 redshift $1.5 < z < 4$ quasars with near-infrared spectra. This is the largest sample of its kind, and has facilitated investigations of quasar BH masses and outflow properties which are described in Chapters **??** and **??**.

1.5 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).

6 Eddington ratio: L/L_{Edd}, where the Eddington luminosity L_{Edd} ($= 3.2 \times 10^4 (M/M_\odot) L_\odot$) is the maximum luminosity set by the balance of outward radiation pressure and inward gravitational force.

1.5.1 *Reverberation mapping*

Under the assumptions that the BLR dynamics are virialised[7] and the gravitational potential is dominated by the BH, the BH mass is given by:

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

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 years, the 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

7 The virial theorem states that the average kinetic energy of a system is equal to half of the average negative potential energy.

et al., 2010; Barth et al., 2011; Grier et al., 2012). This sample is strongly biased to low luminosity Seyfert 1 galaxies[8], 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.

1.5.2 *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 (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) have thus become possible through the calibration of single-epoch virial-mass estimators using the reverberation-mapping measurements (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).

With single-epoch virial masses, the growth rate of massive BHs can be measured across cosmic time. This conveys important information about the accretion processes occurring 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). Recent claims of a BH with mass $10^{10}\,M_\odot$ in a $z = 6.3$ quasar (when the Universe is less than 1 Gyr old; Wu et al., 2015) challenges our understanding of the accretion histories of supermassive BHs (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

8 Seyfert 1: A low-luminosity ($L_{Bol} < 10^{12} L_\odot$) class of AGN with broad emission-lines and clearly detectable host-galaxies.

relations calibrated for sub-Eddington BHs with $M_{BH} \sim 10^7\,M_\odot$ are valid for BHs with masses up to $10^{10}\,M_\odot$ that are radiating near the Eddington luminosity. Furthermore, the vast majority of reverberation mapping measurements are for Hβ and so the $R_{BLR} - L$ relation that underpins the virial method has only been established using this line. Hβ is redshift beyond the reach of optical spectrographs at redshifts $z \gtrsim 0.7$, and extending the method to higher redshifts requires the secondary-calibration of other low-ionization emission-lines such as Hα and Mg II (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 of $z \gtrsim 2$ the low-ionization hydrogen and Mg II emission lines are no longer present in the optical spectra of quasars and it is necessary to employ an emission line in the rest-frame ultraviolet. The strong C IV emission doublet is visible in the optical spectra of quasars to redshifts of $z \sim 5$ and C IV-derived BH masses have become the standard (e.g. Vestergaard and Peterson, 2006; Park et al., 2013). However, 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 BH mass estimates derived from C IV velocity-widths are known to be systematically biased compared to masses from the hydrogen Balmer lines. In Chapter **??**, we use 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 to quantify 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 3000 km s^{-1} and are over-estimated by almost an order of magnitude at the most extreme blueshifts, $\gtrsim 5000$ km 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 ($\sim$ 0 km s^{-1}) and just 0.10 dex at high blueshifts ($\sim$ 3000 km s^{-1}), compared to 0.40 dex before the correction. The correction depends only on the C IV line properties - i.e. the full-width at half-maximum (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 estimates for the majority of high-

luminosity, high-redshift, spectroscopically confirmed quasars in the literature.

1.6 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 for outflows is seen in broad absorption-line quasars (BAL quasars; Weymann et al., 1991). BAL quasars are characterised by broad absorption features in the ultraviolet resonance lines of highly ionised N V, C IV and Si IV. The absorption troughs are thousands of $\mathrm{km\,s^{-1}}$ wide and at significantly blueshifted relative to the quasar rest-frame[9]. The absorption is thought to occur in outflows reaching $60\,000\,\mathrm{km\,s^{-1}}$ (e.g. Turnshek, 1988). The near-universal blueshifting of the observed absorption features can be understood if 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. The observed C IV BAL 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 ultraviolet 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). The blueshifting of high-ionisation lines in the BLR (including C IV) can also 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

9 Much rarer cases of redshifted BAL troughs do also exist (e.g. Hall et al., 2013).

models have been developed to explain the wide range of emission and absorption-line phenomena which are observed (e.g. Murray et al., 1995; Elvis, 2000; Proga, Stone, and Kallman, 2000; Everett, 2005).

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).

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] is the most studied of the narrow AGN emission-lines. By following [O III] to near-infrared wavelengths, we are able to extend these investigations to the redshift range when star formation and BH accretion peaked. In Chapter **??**, 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 (characterised by w_{80}, the width containing 80 per cent of the line power) is $1535 \pm 562\,\mathrm{km\,s^{-1}}$, which is significantly broader than is found in lower-luminosity AGN. We also find blue-asymmetries 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] EQW and the quasar luminosity (the 'Bald-

win' 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 (ICA) can be used to extend these results to spectra with lower signal-to-noise.

1.7 THE SPECTRAL ENERGY DISTRIBUTIONS OF AGN

AGN emit strongly over many decades of the electromagnetic spectrum. Different physical processes dominate AGN spectral energy distributions at different frequencies. X-rays are generated in a hot corona surrounding the accretion disc, the ultraviolet/optical region is dominated by thermal accretion disc emission, re-processed radiation from dust dominates at infrared wavelengths, while radio jets are launched from close to the BH. A complete understanding of AGN properties therefore critically depends on the availability of multi-wavelength data that spans the full SED. This has been achieved in recent years with sensitive, wide-field photometric surveys, including the ultraviolet/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).

To first order, AGNs have remarkably similar SEDs that have very little dependence on luminosity, redshift, BH mass of accretion rate (e.g. Elvis et al., 2012; Hao et al., 2013). In Chapter **??**, we build a simple parametric SED model that is able to reproduce the median optical-infrared colours of tens of thousands of SDSS AGN at redshifts $1 \lesssim z \lesssim 4$. On the other hand, the SED properties of individual quasars do show significant variation. 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 find that the amount of hot dust is strongly correlated with the strength of outflows in the quasar BLR (parametrised using the C IV blueshift). We consider the implications of this result in the context of accretion disc wind models (e.g. Elitzur and Shlosman, 2006).

1.8 NOTE ON CONVENTIONS ADOPTED IN THIS THESIS

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. Unless otherwise stated, optical (i.e. SDSS) magnitudes are given in the AB system and infra-red magnitudes in the Vega system, following the conventions of the original surveys.