

## NARROW LINE REGION PROPERTIES

---

### 1.1 INTRODUCTION

X-ray and UV spectroscopy has revealed high-velocity outflows to be nearly ubiquitous in high accretion rate quasars. Strong evidence for high-velocity outflows include BALs, NALs and blueshifts in high-ionisation broad emission-lines. This suggests that the energy released by quasars can have a dramatic effect on their immediate environment.

The BH mass and the mass of the host galaxy spheroid are strongly correlated. This suggests that the BH and bulge grown 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 and Rees, 1998; King, 2003; Di Matteo, Springel, and Hernquist, 2005; King and Pounds, 2015). If this picture is true, then quasars should be capable of driving powerful outflows over galactic scales. In recent years, a huge amount of resources have been devoted to searching for observational evidence of these galaxy-wide, quasar-driven outflows. This has resulted in recent detections of outflows in quasar-host galaxies using tracers of atomic, molecular, and ionised 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; Ciccone et al., 2014).

One particularly successful technique has been using forbidden emission-lines to probe conditions in the AGN NLR. Because of its high equivalent width,  $[\text{O III}]\lambda\lambda 4960, 5008$  is the most studied of the narrow AGN emission-lines. The  $[\text{O III}]$  emission is found to consist of two distinct components: a narrow, ‘core’ component, with a velocity close to the systemic redshift of the host galaxy, and a broader ‘wing’ component, which is normally blueshifted. The general consensus is that the core component is dominated by the gravitational potential of the host galaxy whereas the broad, blueshifted wing traces outflowing gas. The relative balance between the core and wing

components varies significantly from object to object, and may depend on properties of the AGN (e.g. luminosity; Shen and Ho, 2014).

Observations of broad velocity-widths and blueshifts in narrow emission-lines stretch back several decades (e.g. Weedman, 1970; Stockton, 1976; Heckman et al., 1981; Veron, 1981; Feldman et al., 1982; Heckman, Miley, and Green, 1984; Vrtilek, 1985; Whittle, 1985; Boroson and Green, 1992). However, these studies rely on small samples, which are often unrepresentative of the properties of the AGN population. More recently, the advent of large optical spectroscopic surveys (e.g. SDSS) have facilitated studies of the NLR in tens of thousands of AGN (e.g. Boroson, 2005; Greene and Ho, 2005; Zhang et al., 2011; Mullaney et al., 2013; Zakamska and Greene, 2014; Shen and Ho, 2014). This has provided constraints on the prevalence and drivers of ionised outflows. At the same time, there is strong evidence from spatially resolved spectroscopic observations that these outflows are extended over galaxy scales (e.g. Greene et al., 2009; Greene et al., 2011; Hainline et al., 2013; Harrison et al., 2012; Harrison et al., 2014).

However, these studies do not cover the redshift range when star formation and BH accretion peaked ( $2 \lesssim z \lesssim 4$ ), which is when quasar feedback is predicted to be at its most effective. At these redshifts bright optical emission-lines including the [O III] doublet are redshifted to near-infrared wavelengths, where observations are much more challenging. As a consequence, studies at high redshifts have typically relied on relatively small numbers of objects (e.g. Netzer et al., 2004; Sulentic et al., 2004; Shen, 2016). These studies find [O III] to be broader in more luminous AGN, suggesting that AGN efficiency in driving galaxy-wide outflows increases with luminosity (e.g. Netzer et al., 2004; Nesvadba et al., 2008; Kim et al., 2013; Brusa et al., 2015; Carniani et al., 2015; Perna et al., 2015; Bischetti et al., 2016). The fraction of objects with very weak [O III] emission also appears to increase with redshift and/or luminosity (e.g. Netzer et al., 2004).

Other recent studies have looked at the [O III] emission properties of extreme objects - e.g. heavily obscured quasars (Zakamska et al., 2016) and the most luminous quasars (Bischetti et al., 2016) - at redshifts  $z \sim 2$ . When detected, the [O III] emission in these objects is extremely broad and strongly blueshifted. These observations are consistent with galaxy formation mod-

Table 1.1: The numbers of quasars with [O III] line measurements and the spectrographs and telescopes used to obtain the near-infrared spectra.

Spectrograph	Telescope	Number
FIRE	MAGELLAN	31
GNIRS	GEMINI-N	28
ISAAC	VLT	7
LIRIS	WHT	7
NIRI	GEMINI-N	29
NIRSPEC	Keck II	3
SINFONI	VLT	80
SOFI	NTT	76
TRIPLESPEC	ARC-3.5m	27
TRIPLESPEC	P200	45
XSHOOTER	VLT	21
Total		354

els that predict AGN feedback to be strongest in luminous, dust-obscured quasars.

In this 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.

## 1.2 QUASAR SAMPLE

From our near-infrared spectroscopic catalogue (Chapter ??), we have selected 354 quasars which have spectra covering the strong, narrow [O III] doublet. The broad Balmer H $\beta$  line has also been observed for all but two of the sample. For 165 quasars, the spectra extend to the broad H $\alpha$  emission-line at 6565Å, and in 260 objects optical spectra, including C IV, are also available (mostly from SDSS/BOSS). The sample covers a wide range in redshifts ( $1.5 \lesssim z \lesssim 4$ ) and luminosities ( $45.5 \lesssim \log L_{\text{Bol}} \lesssim 49 \text{ erg s}^{-1}$ ). The spectrographs and telescopes used to obtain the near-infrared spectra are summarised in Table 1.1.

### 1.3 PARAMETRIC MODEL FITS

In this section, we describe how emission-line parameters are derived. Our approach is to model the spectra using a power-law continuum, an empirical Fe II template (taken from Boroson and Green 1992) and multiple Gaussian components to model the emission from the broad and narrow emission-line regions. Non-parametric properties are then derived from the best-fitting model. This approach, which is commonly adopted in the literature (e.g. Shen et al., 2011; Shen and Liu, 2012; Shen, 2016), is more robust when analysing spectra with limited S/N (in comparison to measuring line properties directly from the data) and allows different emission-lines to be de-blended.

The same approach was used to model the H $\beta$ /[O III] complex in Chapter ?? . However, a number of small adjustments have been made to the model (Section 1.3.3). H $\alpha$  emission-line properties (used to estimate the quasar systemic redshift) are also re-derived in this Chapter using a slightly modified model (Section 1.3.4) to the one adopted in Chapter ?? . C IV emission-line properties (used to infer the strength of BLR outflows) are taken directly from Chapter ?? .

#### 1.3.1 *Transforming spectra to the quasar rest-frame*

Before a spectrum can be modelled, it must first be transformed to the quasar rest-frame. The redshift used in this transformation is either derived from the peak of the broad H $\alpha$  emission ( $\sim 40$  per cent of our sample), from the peak of the broad H $\beta$  emission ( $\sim 40$  per cent) or from the peak of the narrow [O III] emission (20 per cent). The rest-frame transformation is only required to be accurate to within  $\sim 1000 \text{ km s}^{-1}$  of the systemic redshift for our fitting procedure to work. In later sections, more precise estimates of the systemic redshift will be calculated using our parametric model fits.

#### 1.3.2 *Removing Fe II emission*

Fe II emission is generally strong in the vicinity of H $\beta$ /[O III]. Therefore, before H $\beta$ /[O III] is modelled, we first model and subtract the continuum and Fe II emission using the procedure described in Chapter ?? .

We encountered 24 objects for which Fe II emission appears to be present in the spectrum even after the subtraction proce-

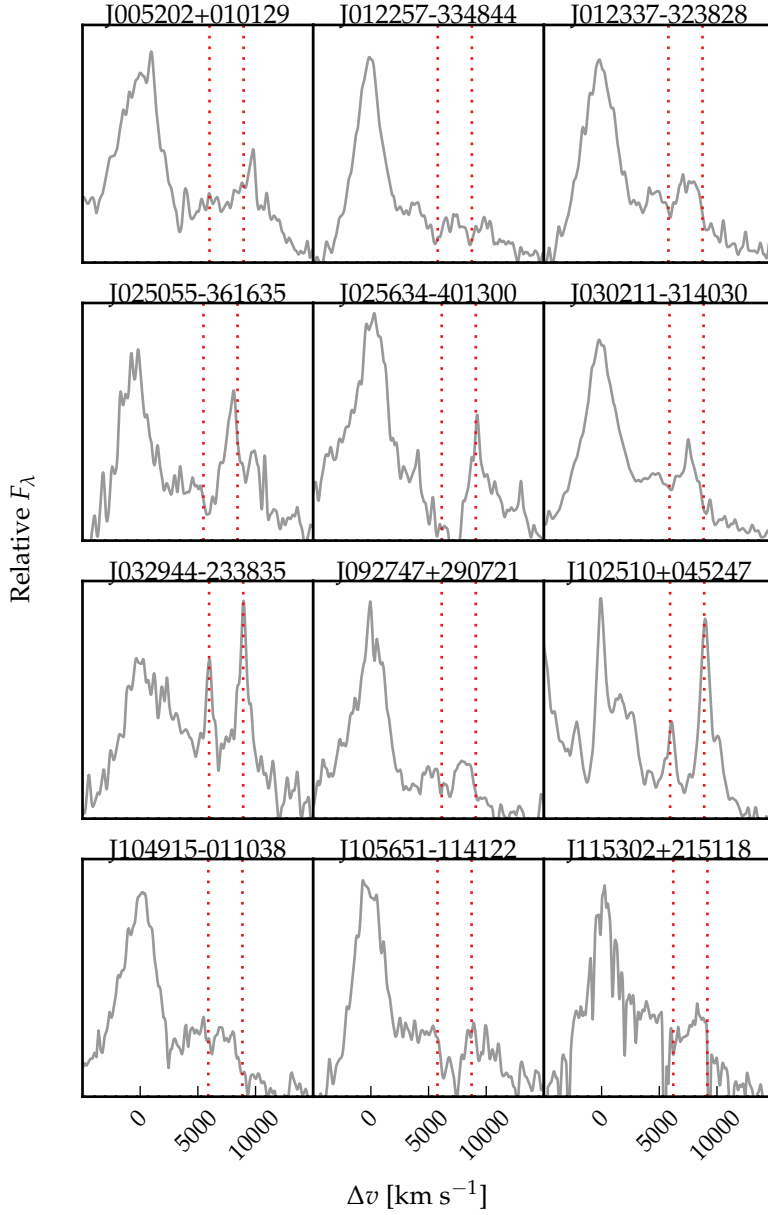


Figure 1.1: Spectra of the 24 objects for which significant Fe II emission is still visible following our Fe II-subtraction procedure. Spectra have been smoothed via convolution with a  $100 \text{ km s}^{-1}$  Gaussian kernel. The vertical lines indicate the expected positions of the [O III] doublet (which is generally very weak) with the systemic redshift defined using the peak of the broad H $\beta$  emission.

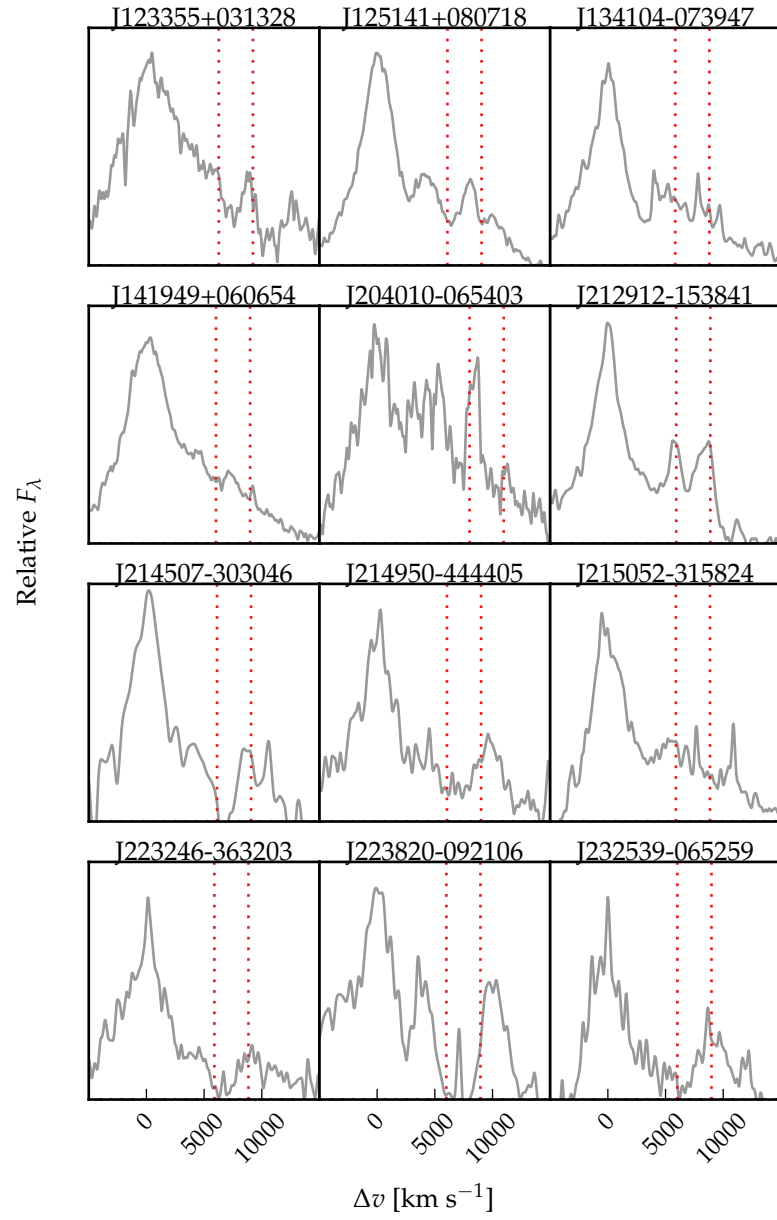


Figure 1.1: Continued.

Table 1.2: Summary of models used to fit the  $H\beta$  emission, and the number of quasars each model is applied to.

Model	Fix centroids?	Number
2 broad Gaussians + 1 narrow Gaussian	No	9
2 broad Gaussians	No	274
2 broad Gaussians	Yes	39
1 broad Gaussian	N/A	8

Figure ??). In these objects the relative strengths of the  $\text{Fe II}$  lines differ significantly from those of I Zw 1, on which the Boroson and Green, (1992)  $\text{Fe II}$  template we use is based. The residual  $\text{Fe II}$  emission is at rest-frame wavelengths very close to zero-velocity wavelengths of the  $[\text{O III}]$  doublet, which is generally very weak in these objects. The Gaussians we fit to the spectra to model  $[\text{O III}]$  are therefore strongly biased, and the  $[\text{O III}]$  emission properties we infer from the model are in error by large factors.

For example, J223819-092106 was analysed by Shen, (2016) using a very similar model. Shen, (2016) reported the  $[\text{O III}]$  emission in this object to be shifted by  $\sim 7500 \text{ km s}^{-1}$  relative to the Hewett and Wild, (2010) systemic redshift. Our analysis suggests that emission which was modelled by Shen, (2016) as  $[\text{O III}]$  is more likely to be poorly-subtracted  $\text{Fe II}$  emission. Because the derived  $[\text{O III}]$  emission properties can be strongly biased in objects so-affected, these objects are flagged and are excluded from our analysis in the remainder of this Chapter (leaving 330 objects in our sample).

### 1.3.3 Modelling $H\beta/[\text{O III}]$

The  $H\beta$  and  $[\text{O III}]$  emission is fit using a similar procedure to the one described in Chapter ???. However, we make a number of modifications to the parametric model employed, which we will now describe.

#### $H\beta$

In general,  $H\beta$  is modelled by two Gaussians with non-negative amplitudes and FWHM greater than  $1200 \text{ km s}^{-1}$ . In 10 objects  $H\beta$  is modelled with a single Gaussian and in 41 objects  $H\beta$  is modelled with two Gaussians, but the velocity centroids of the two Gaussians are constrained to be equal. These spectra gen-

erally have low S/N, and adding extra freedom to the model does not significantly decrease the reduced  $\chi^2$ . In addition there are cases where the blue wing of the H $\beta$  emission is below the lower wavelength limit of the spectra; in these cases models with more freedom are insufficiently constrained by the data.

Contributions to the H $\beta$  emission from the NLR is weak in the vast majority of our sample, and in general we do not include an additional Gaussian component to model this emission. In nine objects features in the model - data residuals suggest that a narrow emission component is significant, and an additional narrow Gaussian is included for these quasars. It is likely that there is some not insignificant contribution from the NLR in other quasars in our sample. If this is the case then measures of the H $\beta$  velocity width will be biased to lower values on average. However, our systemic redshift estimates that use the peak of the H $\beta$  emission (Section 1.3.8) will not be affected. The H $\beta$  models, and the numbers of quasars each model is applied to, are summarised in Table 1.2.

### [O III]

Each component of the [O III] doublet is fit with one or two Gaussians, depending on the fractional reduced  $\chi^2$  difference between the one- and two-component models. Concretely, if the addition of the second Gaussian decreases the reduced  $\chi^2$  by more than 5 per cent then the double-Gaussian model is accepted. One hundred and twenty-eight spectra are fit with a single Gaussian and 140 with two Gaussians. The peak flux ratio of the [O III] 4960 Å and 5008 Å components are fixed at the expected 1:3 ratio and the width and velocity offsets are set to be equal<sup>1</sup>.

In 62 objects with very weak [O III] (mean EQW  $\sim 2$  Å) we found that the Gaussian model has a tendency to fit features to the noise. In some cases this can lead to large errors on the [O III] line properties. To avoid this problem, we instead fit a fixed [O III] template to the spectra, with the overall scaling of this template the only free-parameter in the fit. This template is generated by running our line-fitting routine on a median composite spectrum of the 268 quasars with reliable [O III] line measurements. The spectra used to construct the composite

<sup>1</sup> For J003136+003421, a significantly better fit ( $\Delta\chi^2_V \sim 25\%$ ) is obtained when the peak flux ratio constraint relaxed; the peak ratio of the best-fitting model is 1:2.13.



Table 1.3: Summary of models used to fit the [O III] emission, and the number of quasars each model is applied to.

Model	Number
2 Gaussians	140
1 Gaussian	128
Template	62

were first de-redshifted and continuum- and Fe II-subtracted. The models we use to fit [O III], and the numbers of quasars each model is applied to, are summarised in Table 1.3.

In Figure 1.2 we show example fits to 15 objects, chosen at random. The median reduced chi-squared value is 1.31 and, in general, there are no strong features observable in the spectrum minus model residuals.

#### 1.3.4 Modelling $H\alpha$

There are 165 quasars in our sample with spectra covering the  $H\alpha$  emission-line. In Section 1.3.8, we use the peak of the  $H\alpha$  emission as one estimate of the quasar systemic redshift. In this section we describe how the  $H\alpha$  emission was modelled.

The continuum emission is first modeled and subtracted using the procedure described in Section ???. We then test five different models with increasing degrees of freedom to model the  $H\alpha$  emission. The models are summarised in Table 1.4. They are (1) a single broad Gaussian; (2) two broad Gaussians with identical velocity centroids; (3) two broad Gaussians with different velocity centroids; (4) two broad Gaussians with identical velocity centroids, and additional narrower Gaussians to model narrow  $H\alpha$  emission, and the narrow components of [N II] $\lambda\lambda$ 6548,6584 and [S II] $\lambda\lambda$ 6717,6731; (5) two broad Gaussians with different velocity centroids, and additional narrower Gaussians. If used, the width and velocity of all narrow components are set to be equal in the fit, and the relative flux ratio of the two [N II] components is fixed at the expected value of 2.96.

In order to determine which model is selected for each spectrum, we use the following procedure. Each of the five models are fit to every spectrum and the reduced- $\chi^2$  recorded. Initially, the model with the smallest reduced- $\chi^2$  is selected. We then measure how the reduced- $\chi^2$  changes as the complexity of the model is decreased (i.e. considering the models in Table 1.4 in descending order). If it results in an increase in the reduced- $\chi^2$

Table 1.4: Summary of models used to fit the H $\alpha$  emission, and the number of quasars each model is applied to.

Model	Components	Fix centroids?	Number
1	1 broad Gaussian	N/A	10
2	2 broad Gaussians	Yes	71
3	2 broad Gaussians	No	32
4	2 broad Gaussians + narrow Gaussians	Yes	51
5	2 broad Gaussians + narrow Gaussians	No	53

which is less than 10 per cent relative to the best fitting model, then the simpler model is selected.

### 1.3.5 *Deriving emission-line properties from the best-fitting models*

All [O III] line properties are derived from the [O III] $\lambda$ 5008 emission, but, as described above, the kinematics of [O III] $\lambda$ 4960 are constrained to be identical in our fitting routine.

We do not attach any physical meaning to the individual Gaussian components used in the model. Decomposing the [O III] emission into a narrow component at the systemic redshift and a lower-amplitude, blueshifted broad component is subject to large uncertainties and is highly dependent on the spectral S/N and resolution. Furthermore, there is no theoretical justification that the broad component should have a Gaussian profile.

We therefore choose to characterize the [O III] line profile using a number of non-parametric measures, which are commonly used in the literature (e.g. Zakamska and Greene, 2014; Zakamska et al., 2016). A normalised cumulative velocity distribution is constructed from the best-fitting model, from which the velocities below which 5, 10, 25, 50, 75, 90, and 95 per cent of the total flux accumulates can be calculated. These velocities are then adjusted so that the peak of the [O III] emission is at 0 km s<sup>-1</sup>.

The width of the emission-line can then be defined using either  $w_{50}$  ( $\equiv v_{75} - v_{25}$ ),  $w_{80}$  ( $\equiv v_{90} - v_{10}$ ) or  $w_{90}$  ( $\equiv v_{95} - v_5$ ). In terms of the FWHM,  $w_{50} \simeq \text{FWHM}/1.746$ ,  $w_{80} \simeq \text{FWHM}/0.919$ ,  $w_{90} \simeq \text{FWHM}/0.716$ , assuming a Gaussian line profile.  $w_{90}$  is relatively more sensitive to the wings of the line

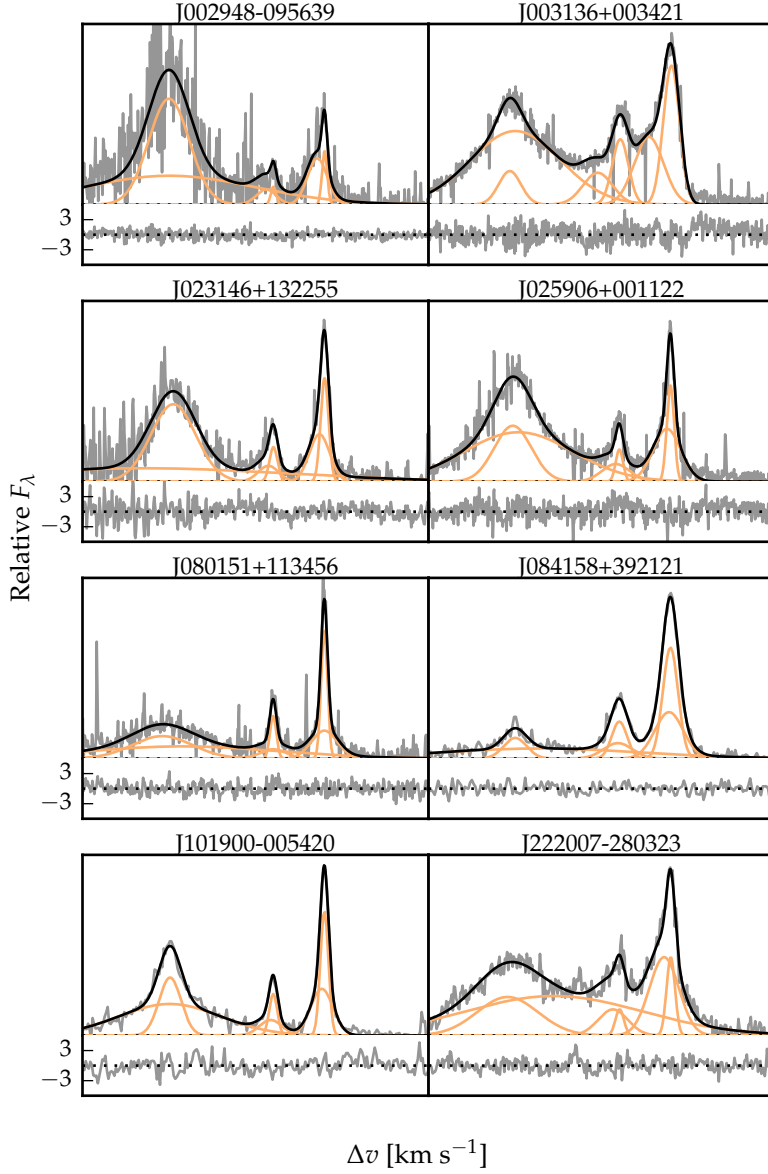


Figure 1.2: Model fits to the continuum- and Fe II-subtracted H $\beta$ /[O III] emission in 15 quasars, chosen at random. The data is shown in grey, the best-fitting model in black, and the individual model components in orange. The peak of the [O III] emission is used to set the redshift, and  $\Delta v$  is the velocity shift from the rest-frame transition wavelength of H $\beta$ . Below each spectrum we plot the data minus model residuals, scaled by the errors on the fluxes.

Table 1.5: The format of the table containing the emission-line properties from our parametric model fits.

Column	Name	Units	Description
1	UID		Catalogue name
2	OIII_V5	$\text{km s}^{-1}$	[O III] $v_5$
3	OIII_V5_ERR	$\text{km s}^{-1}$	Uncertainty in $v_5$
4	OIII_V10	$\text{km s}^{-1}$	[O III] $v_{10}$
5	OIII_V10_ERR	$\text{km s}^{-1}$	Uncertainty in $v_{10}$
6	OIII_V25	$\text{km s}^{-1}$	[O III] $v_{25}$
7	OIII_V25_ERR	$\text{km s}^{-1}$	Uncertainty in $v_{25}$
8	OIII_V50	$\text{km s}^{-1}$	[O III] $v_{50}$
9	OIII_V50_ERR	$\text{km s}^{-1}$	Uncertainty in $v_{50}$
10	OIII_V75	$\text{km s}^{-1}$	[O III] $v_{75}$
11	OIII_V75_ERR	$\text{km s}^{-1}$	Uncertainty in $v_{75}$
12	OIII_V90	$\text{km s}^{-1}$	[O III] $v_{90}$
13	OIII_V90_ERR	$\text{km s}^{-1}$	Uncertainty in $v_{90}$
14	OIII_V95	$\text{km s}^{-1}$	[O III] $v_{95}$
15	OIII_V95_ERR	$\text{km s}^{-1}$	Uncertainty in $v_{95}$
16	z_OIII		[O III] redshift
17	z_OIII_ERR		Uncertainty in [O III] redshift
18	OIII_W50	$\text{km s}^{-1}$	[O III] $w_{50}$
19	OIII_W50_ERR	$\text{km s}^{-1}$	Uncertainty in [O III] $w_{50}$
20	OIII_W80	$\text{km s}^{-1}$	[O III] $w_{80}$
21	OIII_W80_ERR	$\text{km s}^{-1}$	Uncertainty in [O III] $w_{50}$
22	OIII_W90	$\text{km s}^{-1}$	[O III] $w_{90}$
23	OIII_W90_ERR	$\text{km s}^{-1}$	Uncertainty in [O III] $w_{50}$
24	OIII_A		[O III] asymmetry
25	OIII_A_ERR		Uncertainty in [O III] asymmetry
26	OIII_EQW	$\text{\AA}$	[O III] EQW
27	OIII_EQW_ERR	$\text{\AA}$	Uncertainty in [O III] EQW
28	OIII_LUM	$\text{erg s}^{-1}$	[O III] luminosity
29	OIII_LUM_ERR	$\text{erg s}^{-1}$	Uncertainty in [O III] luminosity
30	EQW_FE_4434_4684	$\text{\AA}$	Fe II EQW
31	EQW_FE_4434_4684_ERR	$\text{\AA}$	Uncertainty in Fe II EQW
32	HB_VPEAK	$\text{km s}^{-1}$	H $\beta$ peak velocity
33	HB_VPEAK_ERR	$\text{km s}^{-1}$	Uncertainty in H $\beta$ peak velocity
34	HA_VPEAK	$\text{km s}^{-1}$	H $\alpha$ peak velocity
35	HA_VPEAK_ERR	$\text{km s}^{-1}$	Uncertainty in H $\alpha$ peak velocity
36	HB_Z		H $\beta$ redshift
37	HB_Z_ERR		Uncertainty in H $\beta$ redshift
38	HA_Z		H $\alpha$ redshift
39	HA_Z_ERR		Uncertainty in H $\alpha$ redshift
41	OIII_FE_FLAG		Bad Fe II subtraction
42	OIII_EXTREM_FLAG		Extreme [O III] emission

profile, whereas  $w_{50}$  is relatively more sensitive to the core. We also define the relative asymmetry of the line as:

$$A = \frac{(v_{90} - v_{\text{peak}}) - (v_{\text{peak}} - v_{10})}{(v_{90} - v_{10})}. \quad (1.1)$$

Line-width measures are corrected for instrumental broadening by subtracting the resolution of the spectrograph (Table ??) in quadrature. Because the line profiles are typically non-Gaussian, this deconvolution procedure is only approximate. All of the derived parameters we have calculated are summarised in Table 1.5. The columns are as follows:

- 1 Unique ID: QSOXXX.
- 2-3 Systemic redshift measured at [O III] peak wavelength, and its error.
- 4-17  $v_5, v_{10}, v_{25}, v_{50}, v_{75}, v_{90}$  and  $v_{95}$  velocity of [O III], relative to [O III] peak, and their errors, in  $\text{km s}^{-1}$ .
- 18-23  $w_{50} (\equiv v_{75} - v_{25}), w_{80} (\equiv v_{90} - v_{10})$  and  $w_{90} (\equiv v_{95} - v_5)$  velocity width of [O III], and their errors, in  $\text{km s}^{-1}$ .
- 24-25 Dimensionless [O III] asymmetry  $A$ , and its error.
- 26-27 Rest-frame [O III] EQW, and its error, in  $\text{\AA}$ .
- 28-29  $1-\sigma$  upper-limit on rest-frame [O III] EQW, in  $\text{\AA}$ .
- 30-31 [O III] luminosity, and its error, in  $\text{erg s}^{-1}$ .
- 32-33 4434-4684  $\text{\AA}$  rest-frame Fe II EQW, and its error, in  $\text{\AA}$ .
- 34-35 Velocity of  $\text{H}\beta$  peak, relative to [O III] peak, in  $\text{km s}^{-1}$ , and its error.
- 36-37 Velocity of  $\text{H}\alpha$  peak, relative to [O III] peak, in  $\text{km s}^{-1}$ , and its error.
- 38-38 Redshift of  $\text{H}\beta$  peak, and its error.
- 40-41 Redshift of  $\text{H}\alpha$  peak, and its error.
- 44 Fe II flag.
- 45 Extreme [O III] flag.
- 46-47 C IV  $v_{50}$ , relative to [O III] peak, in  $\text{km s}^{-1}$ , and its error.

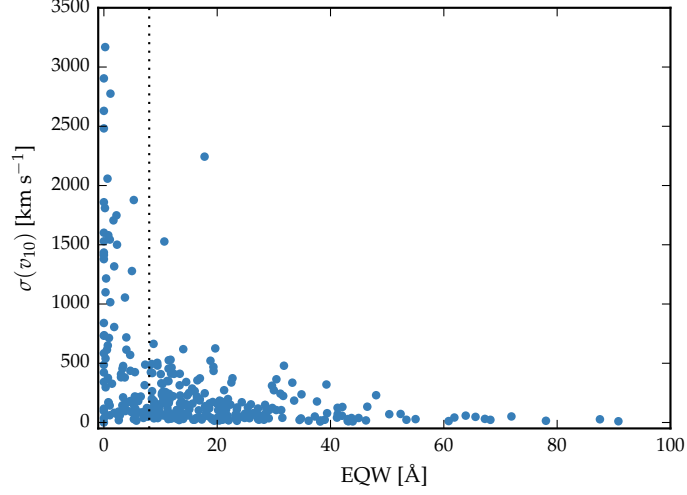


Figure 1.3: Uncertainty in  $v_{10}$  as a function of the EQW, for [O III]. Uncertainties in  $v_{10}$  are large to the left of the vertical line, at  $8\text{\AA}$ . These objects are ignored in our subsequent analysis of the [O III] line shape.

### 1.3.6 Deriving uncertainties on parameters

Our method to estimate realistic uncertainties on emission-line properties derived from the best-fitting model is very similar to the one described in Section ???. Very briefly, random simulations of each spectrum are generated. Our fitting-procedure is run on each simulated spectrum, and the errors on the line parameters are estimated by looking at the distribution of values from the ensemble of simulations. In a slight modification of the procedure in Section ??, the error is defined as half the the 68 (84 - 16) percentile spread in the parameter values.

### 1.3.7 Low EQW [O III]

In Figure 1.3 we show how the uncertainty in  $v_{10}$  depends on the EQW. As the strength of [O III] decreases, the average uncertainty in  $v_{10}$  increases. When the [O III] EQW  $> 80\text{\AA}$ , the mean uncertainty in  $v_{10}$  is  $50\text{ km s}^{-1}$ ; this increases to  $450\text{ km s}^{-1}$  when  $10 < \text{EQW} < 20\text{\AA}$ . As the EQW drops below  $8\text{\AA}$ , typical uncertainties in  $v_{10}$  become very large (exceeding  $1000\text{ km s}^{-1}$  in many objects). Clearly, the emission-line is too weak for properties - in this case  $v_{10}$  - to be reliably measured in many of these objects. Therefore, when the [O III] line properties (e.g. velocity-width, centroid) are analysed in later sections, these ob-

jects with  $EQW < 8\text{\AA}$  will be excluded. This leaves 226 quasars in the sample.

### 1.3.8 Reliability of systemic redshift estimates

In this section, we compare systemic redshift estimates based on [O III], H $\beta$  and H $\alpha$ . The wavelength of each of these lines is measured at the peak of the emission and this measurement is made using the best-fitting parametric model. In the case of the Balmer lines, this model includes both broad and (if present) narrow emission features.

We compare systemic redshift estimates based on [O III] and H $\beta$  (Figure 1.4a), [O III] and H $\alpha$  (Figure 1.4b) and H $\beta$  and H $\alpha$  (Figure 1.4c). [O III], H $\beta$  and H $\alpha$  measurements are available for 226, 418 and 226 objects respectively. We exclude [O III], H $\beta$  and H $\alpha$  measurements when the uncertainties on the peak velocities exceed 200, 300 and 200  $\text{km s}^{-1}$  respectively. This excludes 4, 6 and 12 per cent of the [O III], H $\beta$  and H $\alpha$  measurements respectively. We also exclude [O III] measurements from 16 objects with very broad, blueshifted [O III] emission that is strongly blended with the red wing of H $\beta$  (these objects are discussed in Section 1.6) because these redshifts are almost certainly strongly biased. After these cuts, there are 182, 85 and 162 objects being compared in samples (a), (b) and (c) respectively.

We generate probability density functions using a Gaussian kernel density estimator. The bandwidth, which is optimised using leave-one-out cross-validation, is 170, 120 and 140  $\text{km s}^{-1}$  for samples (a), (b) and (c) respectively. The systematic offset between the H $\alpha$  and H $\beta$  estimates is consistent with being zero, and the scatter is 230  $\text{km s}^{-1}$ . The scatter in these distributions is consistent with previous studies of redshift uncertainties from broad emission-lines (e.g. Shen et al., 2016). The [O III] redshifts appear to be systematically offset in comparison to both H $\alpha$  and H $\beta$ , in the sense that [O III] is blueshifted. This effect is strongest when [O III] is compared to H $\beta$ , in which case [O III] is shifted by  $\sim 100 \text{ km s}^{-1}$  to the blue.

Hewett and Wild, (2010) found that [O III] was blueshifted by  $\sim 45 \text{ km s}^{-1}$  relative to a rest-frame defined using photospheric Ca II  $\lambda\lambda 3935, 3970$  absorption in the host galaxies of  $z < 0.4$  SDSS AGN. They also noted, as we find here, that [O III] is increasingly blue-asymmetric at higher luminosities. Therefore, our finding is consistent with Hewett and Wild, (2010), when the

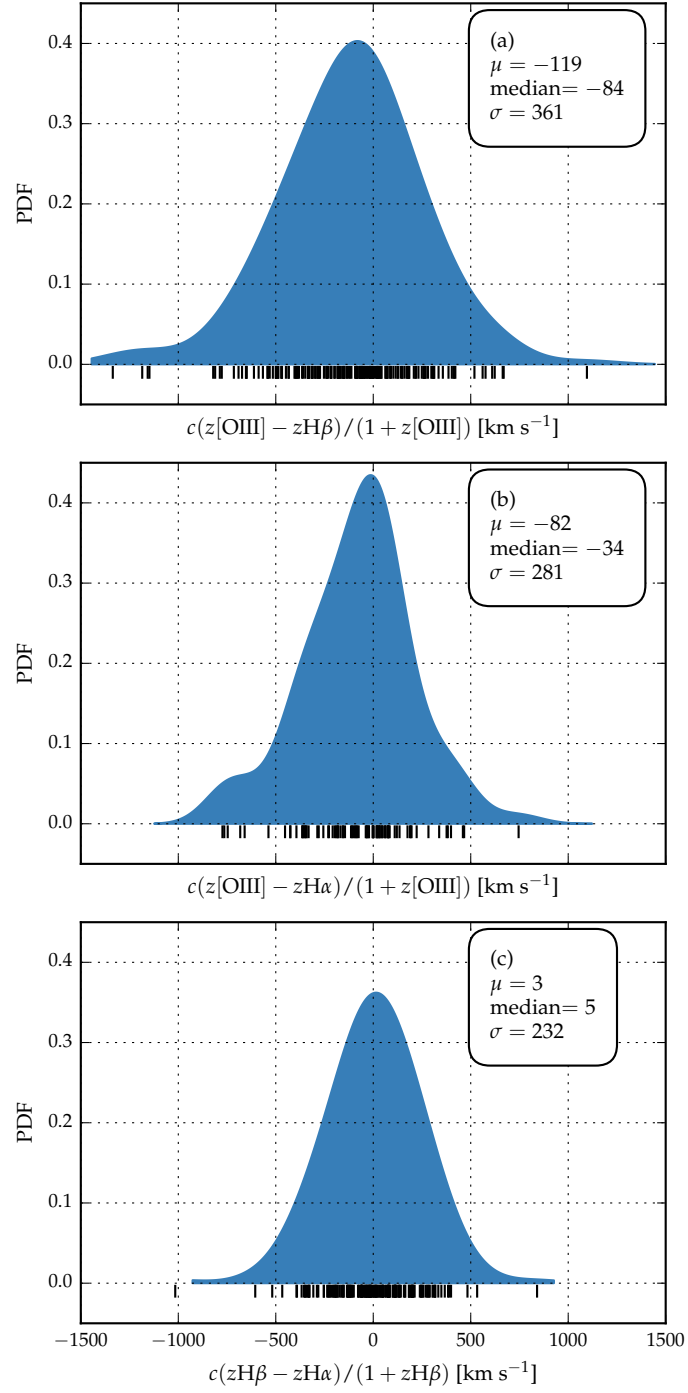


Figure 1.4: Comparison of systemic redshift estimates using [O III], broad H $\beta$  and broad H $\alpha$ . The probability density functions are generated using a Gaussian kernel density estimator with a  $\simeq 150 \text{ km s}^{-1}$  kernel width. The short black lines show the locations of the individual points.



very different luminosities of the two samples are accounted for.

## 1.4 RESULTS

In our sample of 354 quasars, there is a huge diversity in [O III] emission properties (Fig. 1.2). In Figure ??, we present a sub-set of the measurements we have made of the [O III] line.

*Highlight key results more prominently*

The strength of [O III] depends on the covering factor of NLR gas, its density and ionisation parameter. The [O III] EQW follows an approximately log-normal distribution, peaking at  $17\text{\AA}$ . In 10 per cent of our sample [O III] is very weak, with  $\text{EQW} < 1\text{\AA}$ . The average [O III] strength is consistent with earlier studies on smaller samples (e.g. Sulentic et al., 2004; Netzer et al., 2004; Shen, 2016).

The mean and standard deviation of the line width (characterized by  $w_{80}$ ) is  $1535 \pm 562\text{ km s}^{-1}$ , with a median of 1529, minimum of 206 and maximum of 3214. This is consistent with recent near-infrared spectroscopy of  $z > 1.5$  quasars which often report velocity widths  $\gtrsim 1000\text{ km s}^{-1}$  (e.g. Netzer et al., 2004; Kim et al., 2013; Brusa et al., 2015; Shen, 2016). For gas discs rotating in the potential of the massive galaxies line widths do not exceed  $w_{80} \simeq 600\text{ km s}^{-1}$  (Liu et al., 2013). Therefore the [O III] gas cannot be in dynamical equilibrium with the host galaxy. [O III] emission is suppressed by collisional de-excitation in higher-density environments, and so the large velocity widths cannot be due to Doppler broadening in the BLR.

The [O III] asymmetry is shown in Figure 1.5c. In 40 per cent of the sample [O III] is fit with a single Gaussian. The asymmetry is zero in this model and so these objects are excluded. For the objects fit with two Gaussians, [O III] is blue-asymmetric in 90 per cent. This indicates that there is an outflow component in the [O III]-emitting gas. We also find a weak anti-correlation between  $w_{80}$  and the asymmetry (Spearman correlation coefficient: -0.30, p-value:  $4 \times 10^{-4}$ ), in the sense that the broadened lines tend to be more blue-asymmetric.

### 1.4.1 Luminosity/redshift-evolution of [O III] properties

We extend the dynamic range of our samples in terms of both luminosity and redshift by supplementing our sample with quasars presented by Mullaney et al., (2013) and Harrison et al., (2016). The Mullaney et al., (2013) catalogue contains [O III] line

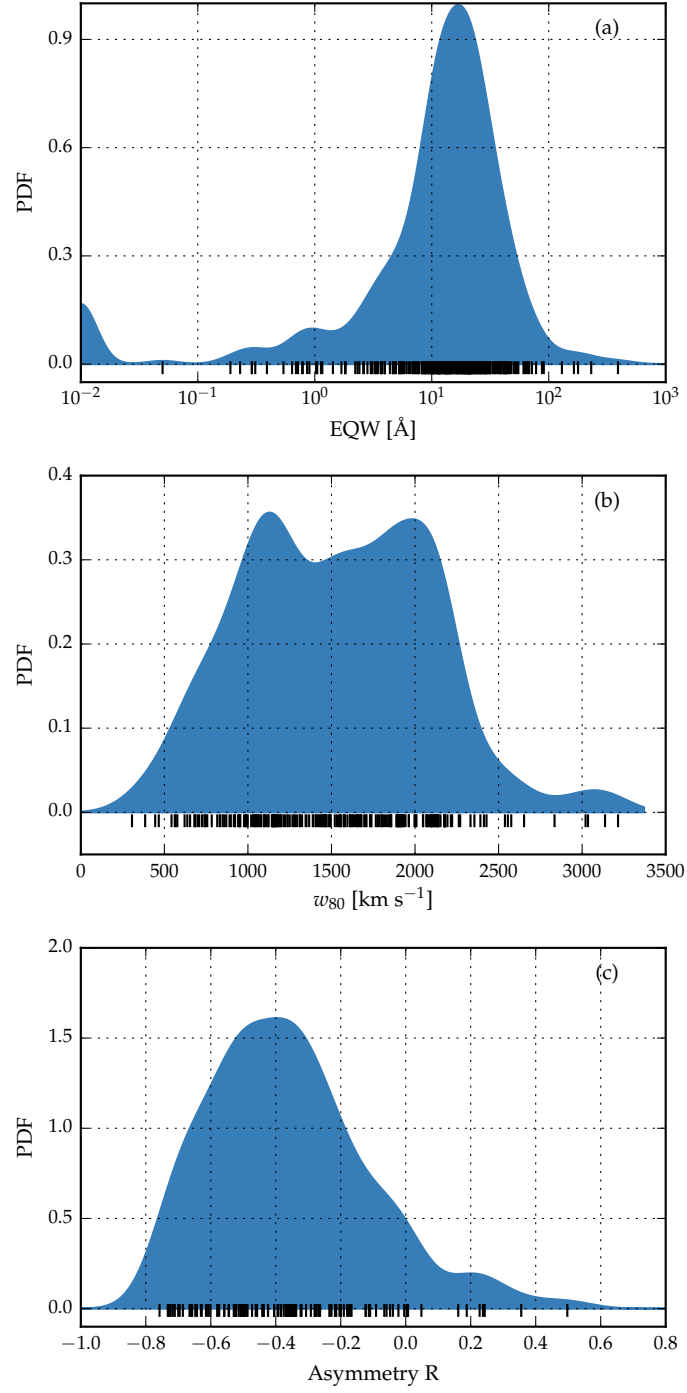


Figure 1.5: Probability density distributions of the [O III] parameters EQW (a),  $w_{80}$  (b) and asymmetry R (c). The  $1200 \text{ km s}^{-1}$  upper limit on the velocity width of the Gaussian functions used to model [O III] is responsible for the peak at  $1200 \text{ km s}^{-1}$  in (b).

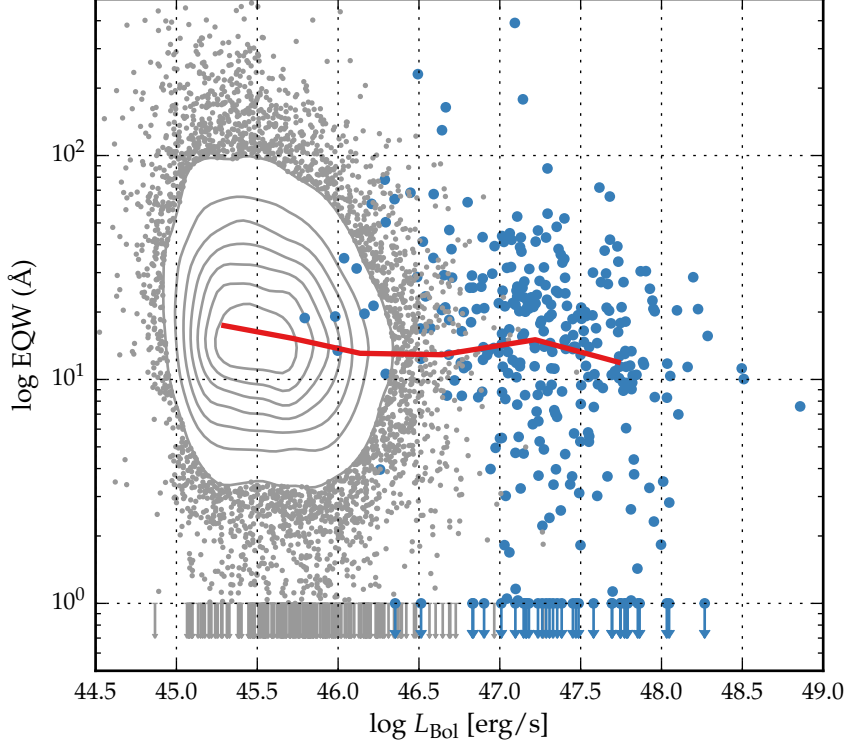


Figure 1.6: The [O III] EQW as a function of the quasar bolometric luminosity for the sample presented in this Chapter (blue circles) and the low- $z$  SDSS sample (grey points and contours). An upper limit at  $\text{EQW} = 1 \text{ \AA}$  indicates points with  $\text{EQW} < 1 \text{ \AA}$ . The red line shows the median [O III] EQW in luminosity bins centred on 45.3, 45.7, 46.1, 46.6, 47.2 and 47.7  $\text{erg s}^{-1}$ , considering only objects with  $\text{EQW} > 1 \text{ \AA}$ .

measurements for  $\sim 25\,000$  optically-selected AGN with SDSS spectra at  $z < 0.4$ . Mullaney et al., (2013) fit [O III] with one or two Gaussians, and then used similar non-parametric measures to the ones we adopt. We select only the Type I AGN from the Mullaney et al., (2013) catalogue. The Harrison et al., (2016) sample contains 40 X-ray selected quasars ( $L_{2-10\text{keV}} \simeq 10^{43-44} \text{ erg s}^{-1}$ ) at intermediate redshifts ( $1.1 \leq z \leq 1.7$ ) observed with the KMOS integral field unit spectrograph on the VLT. We also use the SDSS DR7 quasar catalogue, with properties derived by Shen et al., (2011). [O III] is visible in SDSS spectra up to redshifts  $z = 0.84$ . There are 20,663 quasars in the Shen et al., (2011) catalogue with [O III]  $\text{EQW} > 0 \text{ \AA}$ .

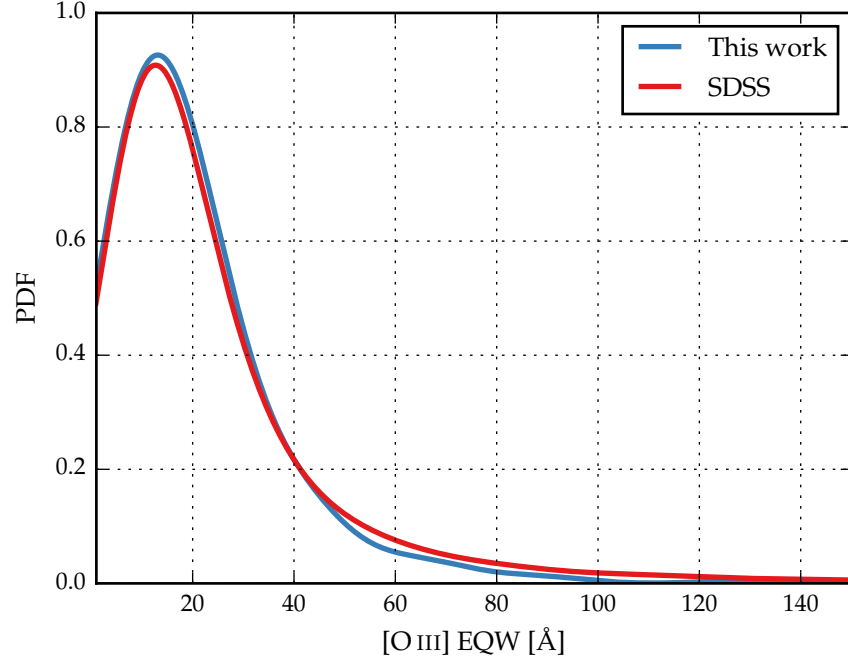


Figure 1.7: [O III] EQW distribution at  $\text{EQW} > 1\text{\AA}$  PDF generated using a Gaussian KDE with a  $8\text{\AA}$  bandwidth (optimized via leave-one-out cross-validation). The two distributions are essentially identical, and peak at  $13\text{\AA}$ . However, 10 per cent of the sample in this work have  $\text{EQW} < 1\text{\AA}$ , compared to just 1 per cent of the SDSS sample.

### *Equivalent width*

In Fig. 1.6 we show the [O III] EQW as a function of the quasar bolometric luminosity. Bolometric luminosity is estimated from the monochromatic continuum luminosity at  $5100\text{\AA}$ , using the correction factor given by Richards et al., (2006). For comparison, we also show the sample from Shen et al., (2011).

Many authors have reported the [O III] EQW to decrease with quasar luminosity (e.g. Brotherton, 1996; Sulentic et al., 2004; Baskin and Laor, 2005). The origin of this correlation - which is known as the Baldwin effect (e.g. Baldwin, 1977; Brotherton, 1996; Zhang et al., 2011; Stern and Laor, 2012) - has not been demonstrated conclusively. One interpretation is that the size of the NLR increases when the quasar luminosity increases, because more ionising photons are available. However, beyond a certain radius there will be no more gas left to ionise, and the size of the NLR will plateau.

We do not observe a Baldwin effect in our sample, despite the luminosity range spanning three dex. This is seen more clearly if we show the distribution of EQWs in the SDSS sample ( $\log L_{\text{Bol}} \sim 45.5 \text{ erg s}^{-1}$ ) and the sample presented in this paper ( $\log L_{\text{Bol}} \sim 47.3 \text{ erg s}^{-1}$ ). Only objects where [O III] is detected with  $\text{EQW} > 1 \text{ \AA}$  are included. The two distributions are essentially identical, and peak at  $13 \text{ \AA}$ .

The fraction of objects with very weak [O III] ( $\text{EQW} < 1 \text{ \AA}$ ) is 10 per cent in our sample, compared to one per cent of the SDSS sample. Therefore, the fraction of objects with very weak [O III] is an order of magnitude larger in the more luminous sample. If we instead define ‘weak’ [O III] emission as having an  $\text{EQW} < 5 \text{ \AA}$ , then 22 per cent of our high-luminosity sample is classified as such, compared to 8 per cent of the SDSS sample.

Netzer et al., (2004), comparing the [O III] properties in a much smaller sample over a comparable luminosity range, reached a similar conclusion. The conclusion reached by Netzer et al., (2004) was that the size of the NLR scaled with the square root of the luminosity of the source of ionising photons in low luminosity AGN, in line with theoretical predictions (e.g. Netzer, 1990). However, extrapolating this relationship to high luminosity quasars leads to the prediction of enormous NLRs with galactic dimensions. If these NLRs had properties similar to the ones in local Seyferts, they would quickly escape the system and disappear. Therefore, the prediction is that no NLR emission should be observed in high-luminosity quasars, which is the case for  $\sim 10$  per cent of our sample. This means that when strong [O III] is detected, it must be from denser gas. Netzer et al., (2004) suggest that this high density gas could be produced in the kiloparsec scale nuclear regions by violent star formation. We find that that virtually all objects which have large C IV blueshifts have very weak [O III] emission.

*Does this  
support/contradict  
Netzer picture?*

#### 1.4.2 Velocity width

In Figure ?? we show the [O III] velocity width as a function of the quasar optical luminosity. A strong correlation exists. On the other hand, we do not find any significant correlation between the [O III] velocity width and the quasar redshift. The lack of any evolution in typical [O III] properties between  $z = 0$  and  $z = 1.5$  has previously been reported (e.g. Harrison et al., 2016); our sample demonstrates that the [O III] properties do not evolve from  $z = 1.5$  all the way to  $z = 4$ .

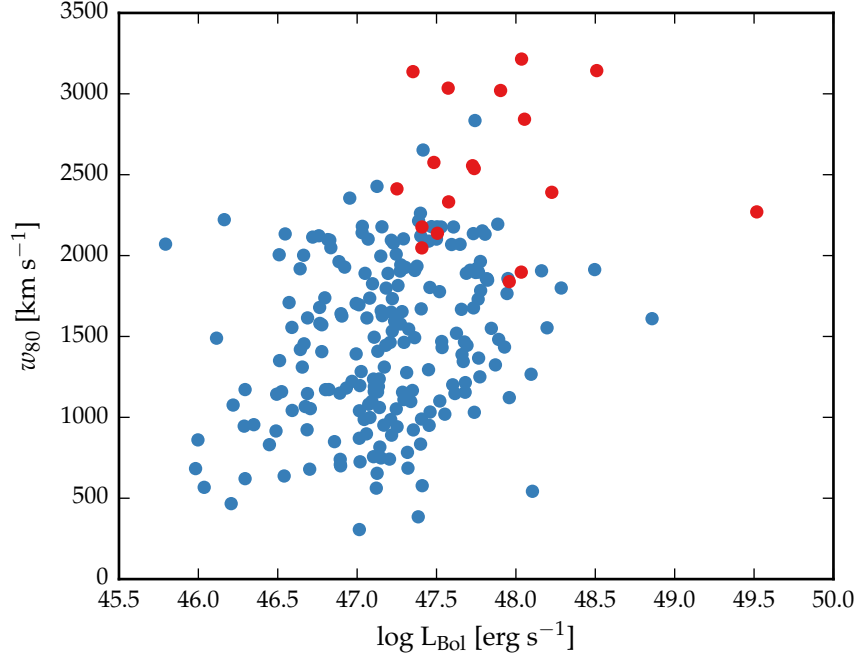


Figure 1.8: [O III] velocity width  $w_{80}$  as a function of quasar bolometric luminosity. Objects with extreme [O III] profiles (Section 1.6) are shown in red.

## 1.5 EIGENVECTOR 1 CORRELATIONS

### 1.5.1 *EV1 trends exist in high-redshift quasars*

The FWHM of the broad H $\beta$  emission-line and the strength of [O III] and the relative strengths of optical Fe II and H $\beta$  have been identified as the features responsible for the largest variance in the spectra of AGN. These parameters form part of EV1, the first eigenvector in a PCA which originated from the work of Boroson and Green, (1992). The underlying driver behind EV1 is thought to be the Eddington ratio (e.g. Sulentic et al., 2000; Shen and Ho, 2014).

In Figure 1.9 we show the [O III] EQW as a function of the H $\beta$  FWHM and the optical Fe II strength. The optical Fe II strength is defined as the ratio of the Fe II and H $\beta$  EQW, where the Fe II EQW is measured between 4434 and 4684 Å. There are 395 objects in our sample with spectra covering H $\beta$  . 303 with partial (>150 Å) coverage of the 4434-4684 region to constrain the Fe II emission. 283 after removing bad fit Fe II. 230 excluding those with more than 25 per cent fractional error in H $\beta$  FWHM or Fe II strength.

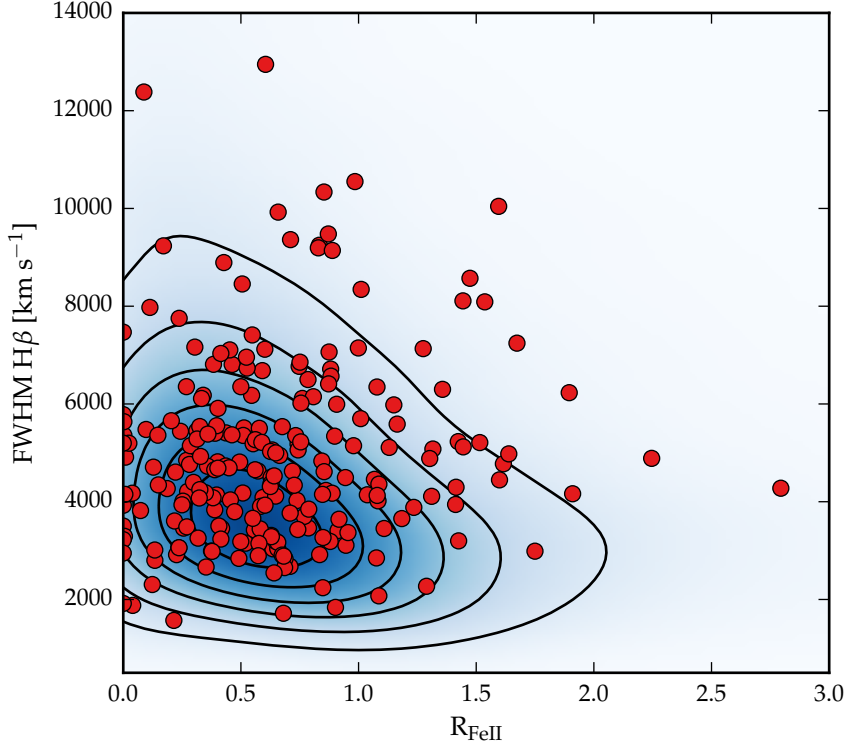


Figure 1.9: EV1 parameter space. The contours and shading show low-redshift, low-luminosity SDSS AGN (with measurements taken from Shen et al. 2011) and the red circles show the high-redshift, high-luminosity objects presented in this Chapter.

The low-redshift SDSS sample, with parameters taken from Shen et al., (2011), is also shown in Figure 1.9. In our sample, these parameters follow very similar correlations to what is observed at low-redshift. In particular, we observe a strong anti-correlation between the [O III] and Fe II EQW. The H $\beta$  FWHM are displaced to higher values, which is consistent with the high-redshift, high-luminosity sample having larger BH masses. Thus, we confirm earlier results using much smaller samples that suggest that the same EV1 correlations exist in high-redshift quasars (e.g. Netzer et al., 2004; Sulentic et al., 2004; Sulentic et al., 2006; Runnoe et al., 2013; Shen, 2016).

### 1.5.2 Connecting EV1 at low and high redshifts

The C IV blueshift and EQW is a diagnostic that similarly spans the diversity of broad emission-line properties in high red-

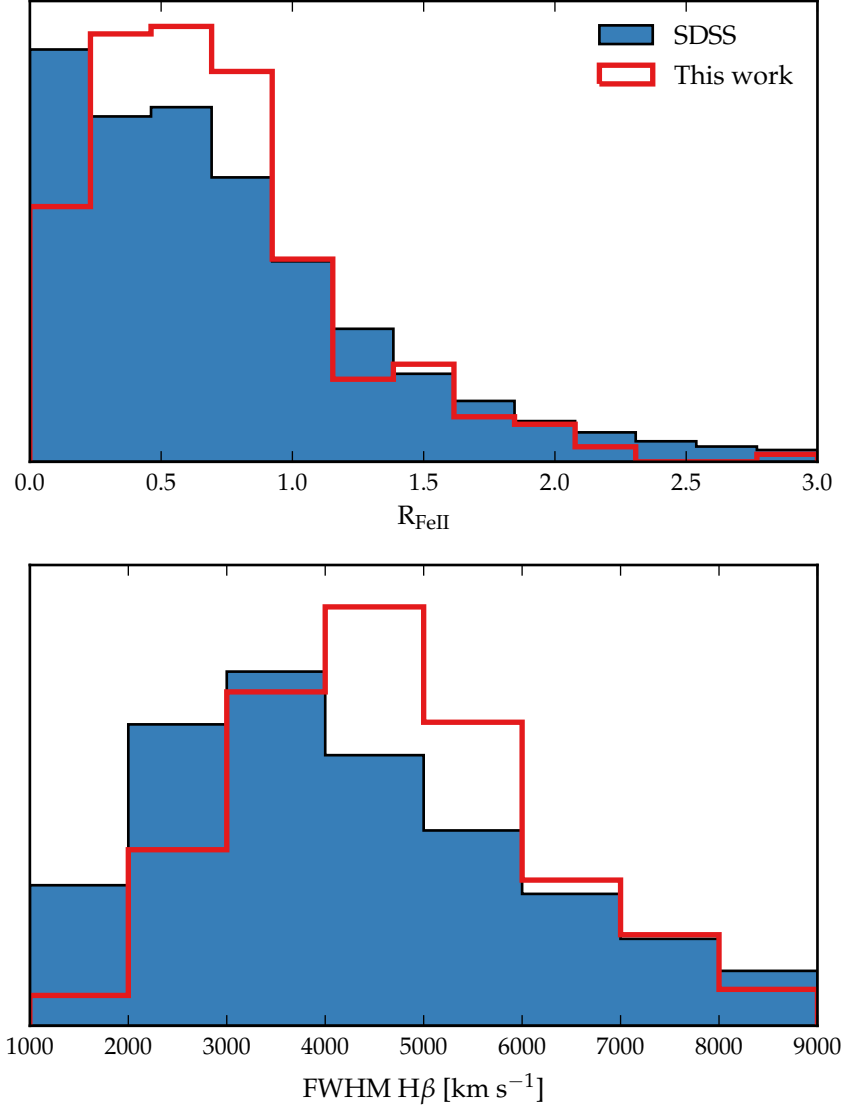


Figure 1.10: EV1 parameters for the high-luminosity quasar sample presented in this work and the low-luminosity SDSS sample.  $H\beta$  FWHM are shifted to higher values in the high-luminosity sample, which is consistent with these quasars having large BH masses. The  $\text{Fe II}$  distributions are similar.

shift quasars (Sulentic et al., 2007; Richards et al., 2011). The similarity of the C IV EQW-blueshift parameter space at high redshift to EV1 parameter space at low redshift suggests that these trends are connected. Because we have optical and near-infrared spectra for XX quasars in our sample, we are able to test how the low-redshift EV1 parameter space maps to the high-redshift C IV parameter space. In Figure 1.11 we show how



the EV1 parameters change as a function of position in the C IV EQW-blueshift parameter space.

Two hundred and thirteen objects are shown in Figure 1.11. Objects for which the C IV line properties could not be measured reliably (see Section ??) have been removed. We consider only objects for which the C IV EQW exceeds 15Å. The C IV blueshift is measured relative to the redshift determined from the peak of [O III], H $\beta$  or H $\alpha$ . We also show the C IV line parameters of 32,157 SDSS DR7 quasars at redshifts  $1.6 < z < 3.0$ . The derivation of the C IV emission properties of these objects is described in Section XX.

Most of the diversity in C IV properties is correlated with the [O III] EQW. This is seen more clearly in Fig. 1.12, in which we plot the [O III] EQW as a function of the C IV blueshift. The correlation seen in Figure 1.12 is very strong. The mean [O III] EQW is 40Å amongst the population of quasars with C IV blueshifts  $< 500 \text{ km s}^{-1}$ , compared to 5Å for quasars with C IV blueshifts  $> 2000 \text{ km s}^{-1}$ . In other words, the NLR emission appears to be missing in quasars with large C IV blueshifts. One possibility is that the NLR gas is swept away on relatively short timescales by quasar outflows; this possibility is explored in more detail in Section XX.

On the other hand, the C IV blueshift and EQW cannot be used to predict the H $\beta$  FWHM. This is consistent with what we found in Chapter ??: objects with large C IV blueshifts have narrow Balmer emission-lines, but objects with modest C IV blueshifts have a wide range of Balmer line widths.

## 1.6 EXTREME [O III] EMITTERS

Figure ?? shows the spectra of 18 objects which we visually identified as having broad, blueshifted [O III] emission which is heavily blended with the red wing of H $\beta$ . Because the emission is so heavily-blended, it is difficult to determine unambiguously what combination of H $\beta$ , [O III] and Fe II is responsible for the unusual plateau-like emission observed in these objects. Therefore, uncertainties on the [O III] emission properties are generally high in these objects.

In Figure 1.8 we show that the luminosities of all of these objects are larger than the sample median. The mean luminosity of the quasars with extreme [O III] emission is  $10^{47.9} \text{ erg s}^{-1}$ , compared to  $10^{47.2} \text{ erg s}^{-1}$  for the rest of the sample. Figure 1.8 demonstrates that these quasars occupy a unique region of the

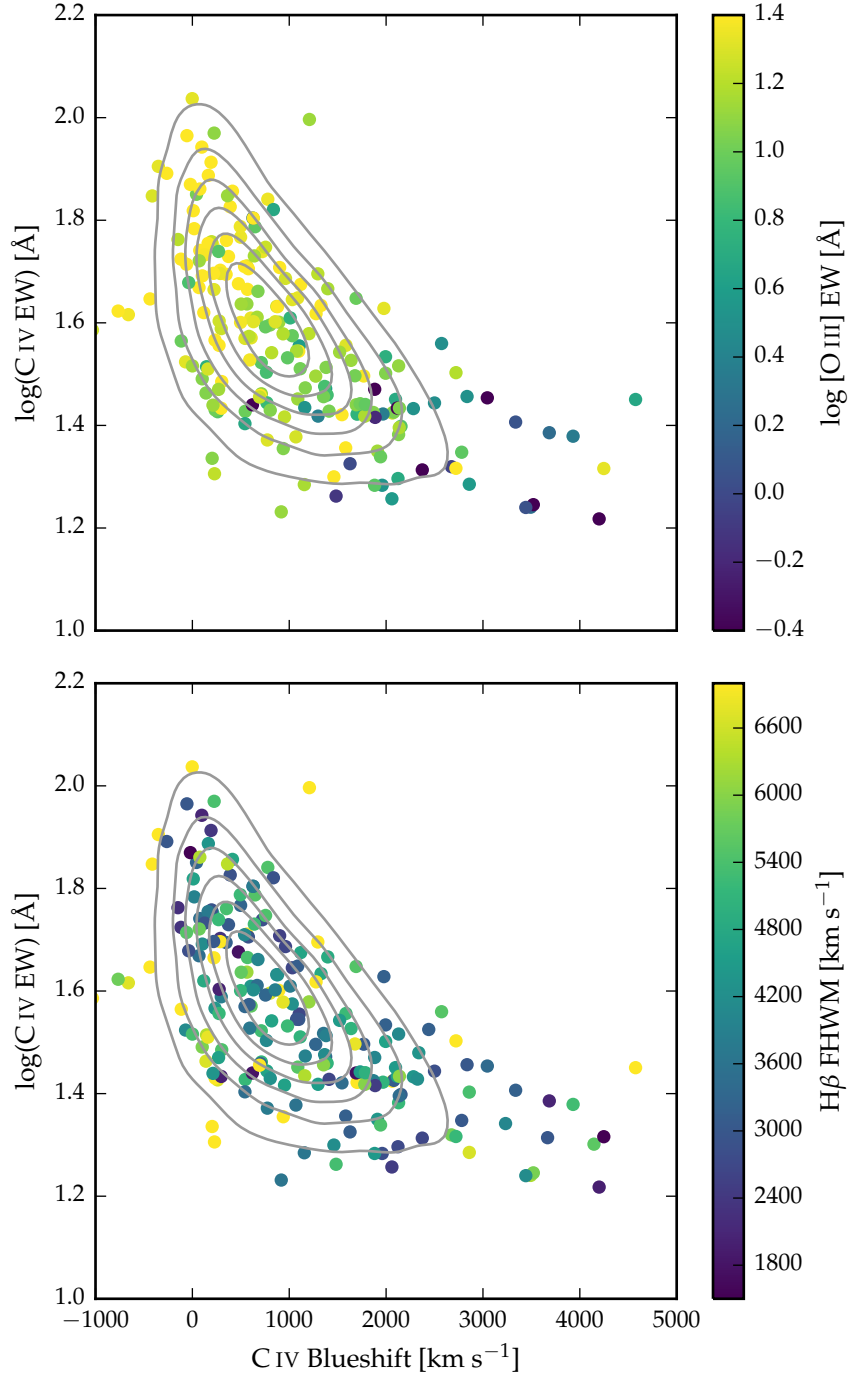


Figure 1.11: The high-redshift EV1 parameter space of C IV blueshift and EQW. Our sample is shown with points, and quasars from the full SDSS catalogue are shown with grey contours. The [O III] EQW varies systematically with position in the C IV blueshift-EQW parameter space (a) but the Hβ FWHM shows significantly less systematic variation (b).

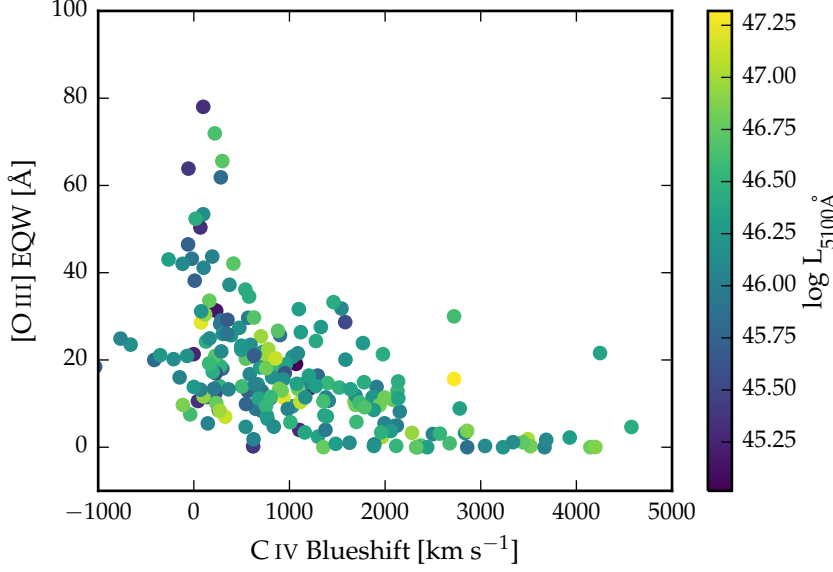


Figure 1.12:  $[\text{O III}]$  EQW as a function of the C IV blueshift. The  $[\text{O III}]$  EQW is strongly anti-correlated with the C IV blueshift. On the other hand, no strong luminosity-dependent trends (indicated by the colours of the points) is evident.

$w_{80}$ - $L_{\text{Bol}}$  parameter space:  $w_{80} \gtrsim 1500 \text{ km s}^{-1}$  and  $L_{\text{Bol}} \gtrsim 10^{47.5} \text{ erg s}^{-1}$ .

A similar  $[\text{O III}]$  emission was also observed in J1201+1206 in a sample of five quasars at redshifts  $2.3 \lesssim z \lesssim 3.5$  with luminosities  $10^{47.5} < L_{\text{Bol}} < 10^{48} \text{ erg s}^{-1}$  observed by Bischetti et al., (2016). These  $[\text{O III}]$  emission-lines are also somewhat similar to the lines observed in a sample of four extremely dust-reddened quasars at  $z \sim 2$  recently identified by Zakamska et al., (2016). The four Zakamska et al., (2016) quasars have  $5 \mu\text{m}$  luminosities of  $\sim 10^{47} \text{ erg s}^{-1}$ , which is comparable to the  $5 \mu\text{m}$  luminosities of the brightest objects in our sample. However, the  $[\text{O III}]$  velocity widths of the Zakamska et al., (2016) objects are extreme in relation to our sample ( $w_{80} \simeq 3500 - 5500 \text{ km s}^{-1}$ ). The extreme nature of the  $[\text{O III}]$  emission in these objects led Zakamska et al., (2016) to propose that these objects are being observed transitioning from a dust-obscured, star-burst phase to a luminous, blue quasar (e.g. Sanders, Hopkins). The fact that we do not see such extreme outflow signatures in our sample, where the mean dust reddening is just 0.03, appears to support this idea.

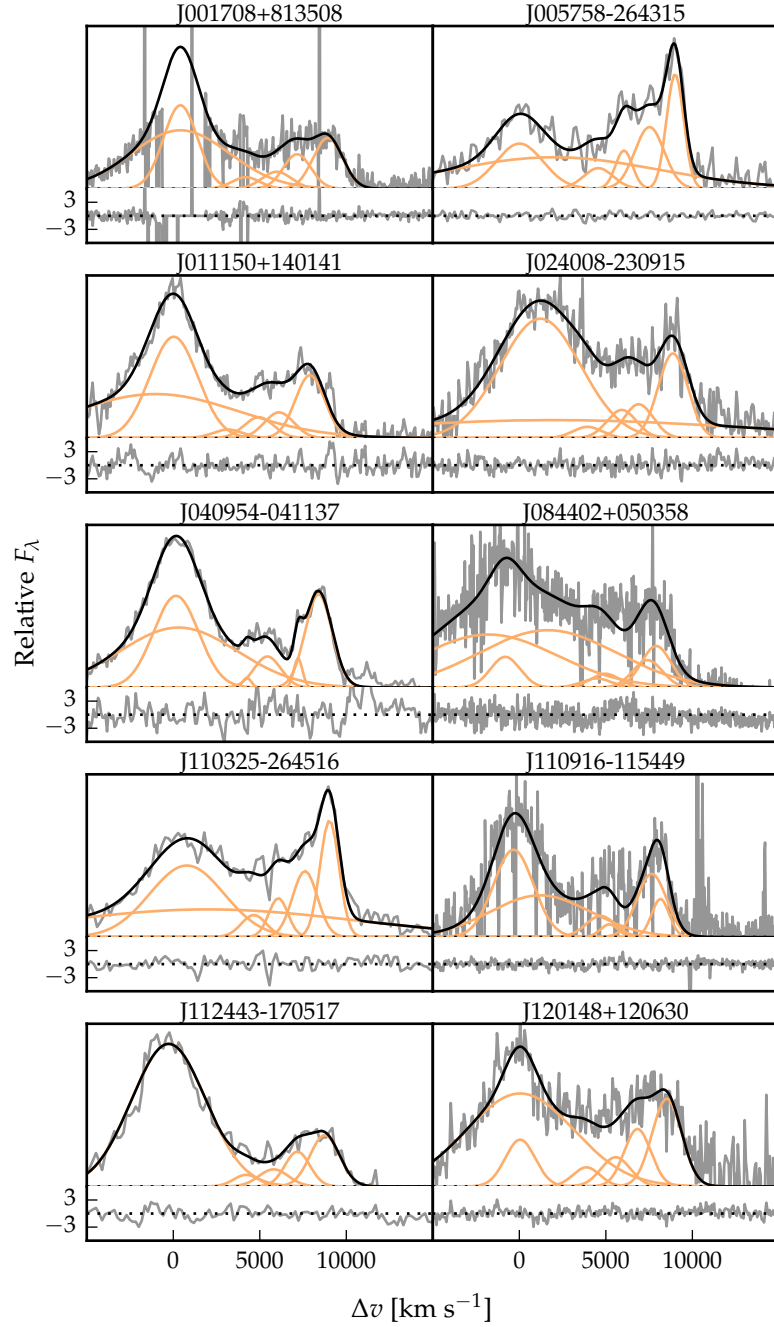


Figure 1.13: Model fits to the continuum- and Fe II-subtracted H $\beta$ /[O III] emission in 18 quasars with extreme [O III] emission profiles. The data is shown in grey, the best-fitting model in black, and the individual model components in orange. The peak of the [O III] emission is used to set the redshift, and  $\Delta v$  is the velocity shift from the rest-frame transition wavelength of H $\beta$ . Below each spectrum we plot the data minus model residuals, scaled by the errors on the fluxes.

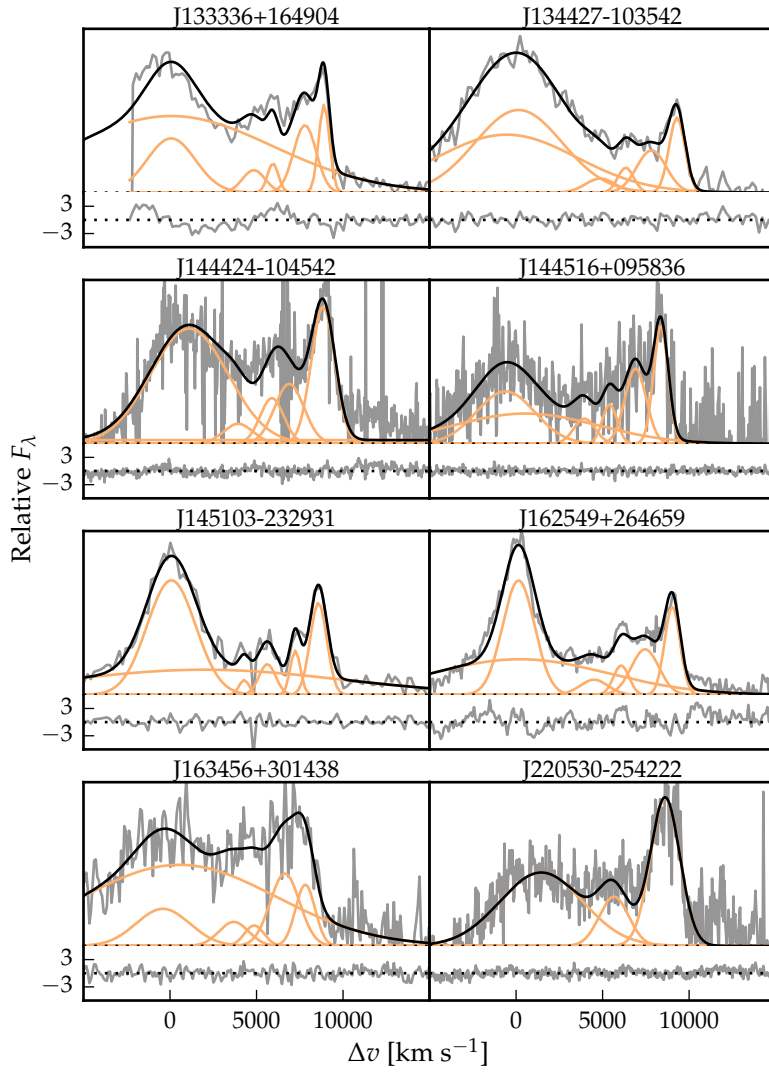


Figure 1.13: Continued.

### 1.6.1 Connections between [O III] and C IV outflows

In Chapter ?? we found that C IV is blueshifted relative to the systemic redshift in a large majority of luminous, high-redshift quasars. In a sub-set, the blueshifting of C IV can reach many thousands of  $\text{km s}^{-1}$ . This is strong evidence that quasars are capable of driving fast outflows, at least over the sub-parsec scale of the BLR. In the present Chapter, we have found that outflows in the NLR, as indicated by broad velocity-widths and asymmetries in [O III], are similarly ubiquitous in the luminous quasar population.

In Section 1.5.2, we found that the [O III] EQW has a very strong dependence on the C IV blueshift. The mean [O III] EQW is  $40\text{\AA}$  amongst the population of quasars with C IV blueshifts  $< 500 \text{ km s}^{-1}$ , compared to  $5\text{\AA}$  for quasars with C IV blueshifts  $> 2000 \text{ km s}^{-1}$ . This establishes a connection between gas in the broad and narrow line regions.

In Figure 1.14 we show the the [O III] blueshift as a function of the C IV blueshift, for the objects where [O III] is detected with  $\text{EQW} > 8 \text{\AA}$ . The [O III] blueshift is defined as  $v_{10}([\text{O III}]) - v_{\text{peak}}([\text{O III}])$  whereas the C IV blueshift is defined as  $v_{50}(\text{C IV}) - v_{\text{peak}}([\text{O III}])$  with the C IV line measurements taken from Chapter ?. We do not show objects for which the errors on the [O III] and C IV blueshifts exceed  $250$  or  $125 \text{ km s}^{-1}$  respectively. These objects, shown in the top two panels of Figure 1.14, have a similar dynamic range to the main sample, meaning our results should not be biased by their exclusion. We also remove the objects with extreme [O III] emission, because the systemic redshift determined from the peak of the [O III] emission is strongly biased in these objects.

Because of the strong anti-correlation between the C IV blueshift and the [O III] EQW, removing objects with weak [O III] eliminates most of the objects with large ( $> 2000 \text{ km s}^{-1}$ ) C IV blueshifts. Nevertheless, [O III] appears to be more blueshifted in quasars with large C IV blueshifts. Although the scatter is large, the correlation appears to be significant (Spearman correlation coefficient:  $0.46$ ,  $p$ -value:  $6e - 7$ ). This suggests a direct connection between the gas kinematics in the broad and narrow line regions.

We considered a number of alternative approaches to parametrising both the [O III] line shape and the systemic redshift. As expected, very similar trends are observed when the [O III] line shape is parametrised using  $v_{25} - v_{\text{peak}}$ ,  $v_{50} - v_{\text{peak}}$ ,

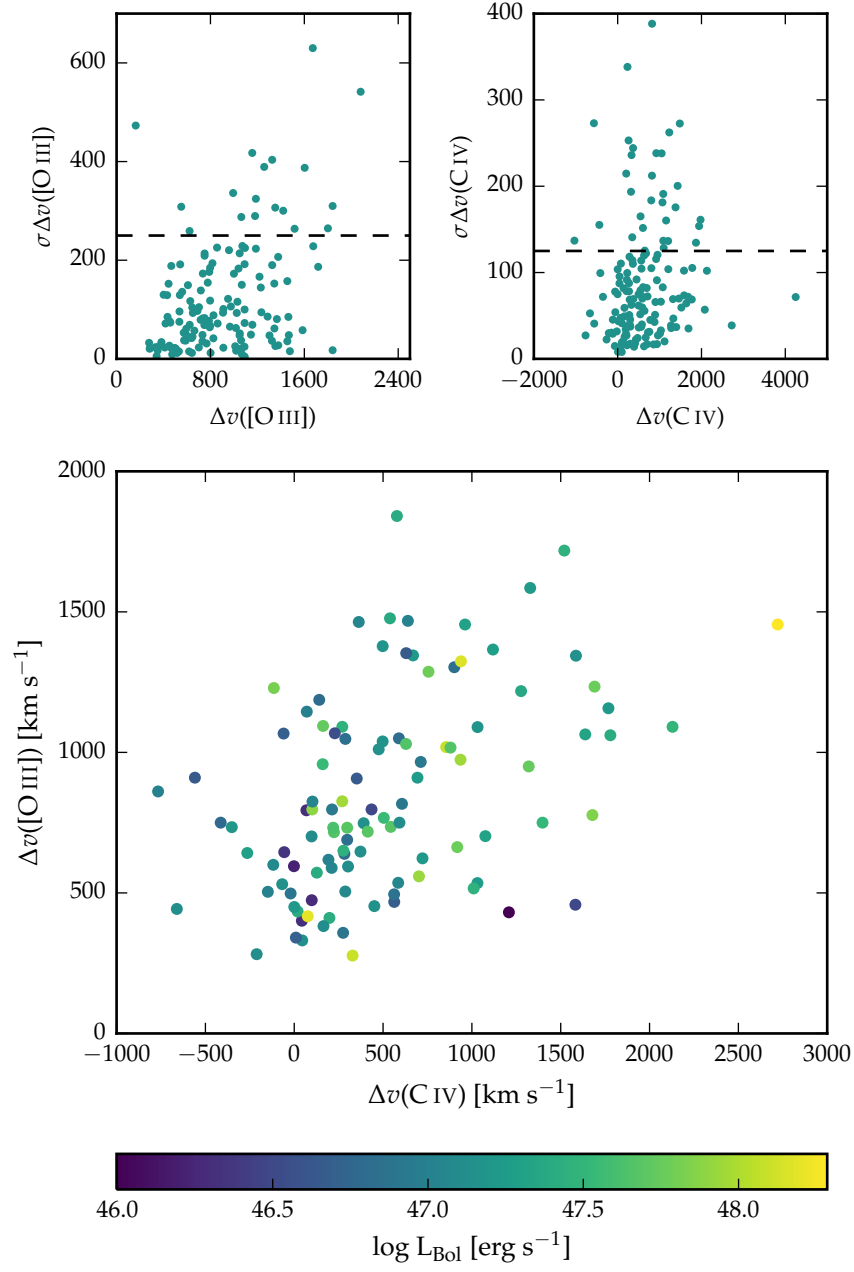


Figure 1.14: The relation between the blueshifts of C IV and [O III].

$w_{80} = v_{90} - v_{10}$ , or the relative asymmetry  $R$ . The same trend is also observed when the systemic redshift is defined using the peak of the  $H\beta$  emission.

The blueshifting of C IV is known to correlate with luminosity (Richards et al., 2011). In [O III], the blueshifted wing becomes relatively more prominent as the luminosity of the quasar increases (Shen and Ho, 2014). Therefore, it is plausible that the correlation between the C IV and [O III] blueshifts is a secondary effect that is driven by the correlation of each with the luminosity. However, no strong luminosity-dependent trends are apparent in Figure 1.14.

## 1.7 DISCUSSION

### 1.7.1 *Static NLR is removed by outflows*

Is the AGN NLR absent in objects where outflows have reached kilo-parsec scales, sweeping up the low-density material responsible for the [O III]-emission? If the BLR outflows can escape, they are very fast and wouldn't need long to clear out the NLR gas. Estimate a time-scale for how long the NLR would take to be cleared given typical size of galaxy and velocity of outflow.

#### READ ZAKAMSKA DISCUSSION

It has been known for some time that the [O III] EQW is anti-correlated with the strength of optical Fe II, and this trend is thought to be driven by the Eddington ratio. Shen and Ho, (2014) showed that the amplitude of the core [O III] emission decreases faster than the wing component as the Eddington ratio increases. Therefore, the [O III] emission is weaker and more blueshifted in high accretion rate quasars. In Chapter ?? we found that all quasars with strong BLR outflows have high Eddington ratios. In this section, we show that the C IV and [O III] blueshifts are directly linked. This suggests a direct connection between the gas kinematics in the broad and narrow line regions. But we need IFU to find out where this gas is (See text in research proposal).

## 1.8 INDEPENDENT COMPONENT ANALYSIS

ICA better with S/N - can give some examples from PCA work (ask Paul for references). If the [O III] emission is confined to a relatively small number of components can then recover emis-



sion. Can show problem with  $w_{80}$  as an example. Encouraging, but we do need a better training set (training set doesn't span the diversity of properties we see in the high-luminosity sample). The content of the ICA "test" is fine but I would suggest setting up the experiment somewhat differently and orienting the reader at the start. For example I would mention a) the likely lack of sensitivity to S/N [cf standard approach], b) equating components with physical properties of interest and thus potential effectiveness of component weights (e.g. what is to come in Fig. 1.16). Would also though put in the information about using the SDSS objects to generate the component set and how one probably expects the high-luminosity sample to include spectral diversity not present in the SDSS-sample [and hence in the ICA components]. You can, as you have drafted, then say how the analysis could be improved - better training set, removing the component cross-talk (I can tell you how this can be achieved),...

In this section, we consider an alternative approach to the analysis presented in the bulk of this Chapter. We use an independent component analysis (ICA) to separate the spectrum into a linear combination of statistically independent sub-components. Each individual spectrum can then be reconstructed with a linear combination of these components. The goal of this section is to determine whether or not the relative weights of the different components can be used in place of more commonly used emission-line parameters to understand the physical processes occurring in these quasars.

Issues with the parametric model fitting approach adopted above include sensitivity to S/N. We also found the empirical template to be a poor match to the Fe II emission observed in a number of quasars.

### 1.8.1 *The technique*

ICA is a blind source separation technique for separating a signal into linearly mixed statistically independent subcomponents. Unlike the more widely-used principle component analysis technique, ICA produces non-negative components which allows for a physical interpretation of the components and weights. ICA has been successfully applied to model the spectra of emission-line galaxies (Allen et al., 2013). The quasar spectra can be thought of as a set of observations,  $\mathbf{x}$ , which are made up

of statistically independent components,  $\mathbf{c}$ , that are combined by some mixing matrix,  $\mathbf{W}$ :

$$\mathbf{x} = \mathbf{W}\mathbf{c} \quad (1.2)$$

ICA reverses this process and describes how the observed data are generated. Both the independent components and the mixing matrix are unknown, but can be found by solving:

$$\mathbf{c} = \mathbf{W}^{-1}\mathbf{x}. \quad (1.3)$$

*Ask Paul for details.*

The components were solved for using a sample of 2,154 SDSS quasars at redshifts  $XX$ . At these redshifts the SDSS spectrograph covers the rest-frame region  $XX-XX\text{\AA}$  where  $H\beta$  and  $[O\text{III}]$  lie. The individual spectra were first adjusted to give the same overall shape as a model quasar template spectrum. Six positive independent components and four lower-amplitude ‘correction’ components that could be negative were found to be sufficient to reconstruct the spectrum, without over-fitting. Each quasar spectrum  $x_j$  can then be represented as a linear combination of the independent components:

$$x_j = \sum_{i=1}^{10} c_{ij} W_{ij} \quad (1.4)$$

#### *Fitting procedure*

Each of the individual ICA components has been adjusted to give the same overall shape as a quasar template spectrum. We approximate the overall shape of this template by fitting a single power-law to emission-line free windows at 4200-4230, 4435-4700 and 5100-5535  $\text{\AA}$ . We then flatten each of the ICA components by dividing by this power-law. An identical process is performed on each spectrum we fit, so that both the components and the spectrum to be fitted have essentially zero large-scale slope. For each quasar in our sample we perform a variance-weighted least-squares minimisation to determine the optimum value of the components weights. The first six component weights are constrained to be non-negative, and the fit is performed in logarithmic wavelength space, so that each pixel corresponds to a fixed velocity width. The relative shift of the

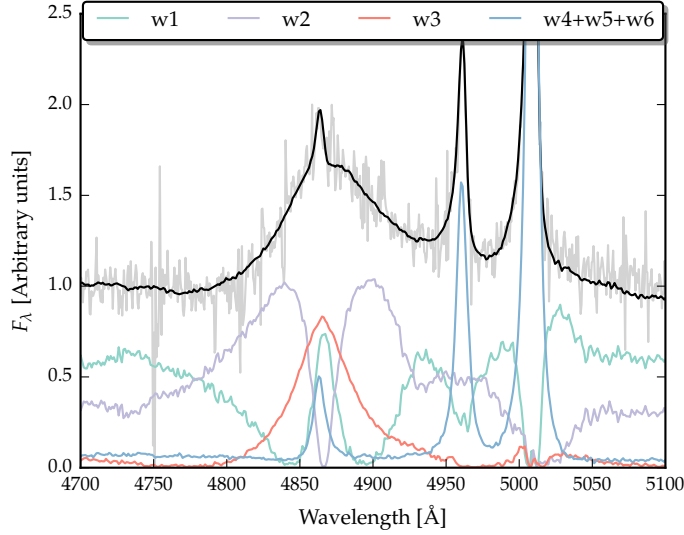


Figure 1.15:  $H\beta/[O\text{ III}]$  emission J002952+020607. The ICA reconstruction is shown in black, and the spectrum in grey. The first three components, and the sum of components four, five and six are shown individually.

ICA components is also allowed to vary in the optimisation procedure, to account for errors in the systemic redshifts used to transform the spectra into rest-frame wavelengths.

### 1.8.2 Quality of fits

In general, the ICA components are able to reconstruct the spectra of the objects in our sample. We also find that that in some cases, the ICA reconstructions are superior at modelling the Fe II emission than the Boroson and Green, (1992) template.

*Need something more quantitative  
Look at chi-squared distribution?  
Doesn't seem that reliable.*

### 1.8.3 Physical interpretation of ICA components

Although the ICA analysis is not based on any physics, there appears to be a direct correspondence between the individual non-negative components and the different emission features which contribute to the spectra (Fig. 1.15). This correspondence is summarised in Table 1.6. The component  $w_1$  seems to correspond to Fe II emission, the components  $w_2$  and  $w_3$  to broad  $H\beta$  emission, the components  $w_4$  and  $w_5$  to narrow  $[O\text{ III}]$  emission at the systemic redshift, and the component  $w_6$  to broad, blueshifted  $[O\text{ III}]$  emission. s

Table 1.6: Physical interpretation of the ICA components.

Component	Origin
$w_1$	Fe II
$w_2$	H $\beta$
$w_3$	H $\beta$
$w_4$	[O III] core
$w_5$	[O III] core
$w_6$	[O III] wing

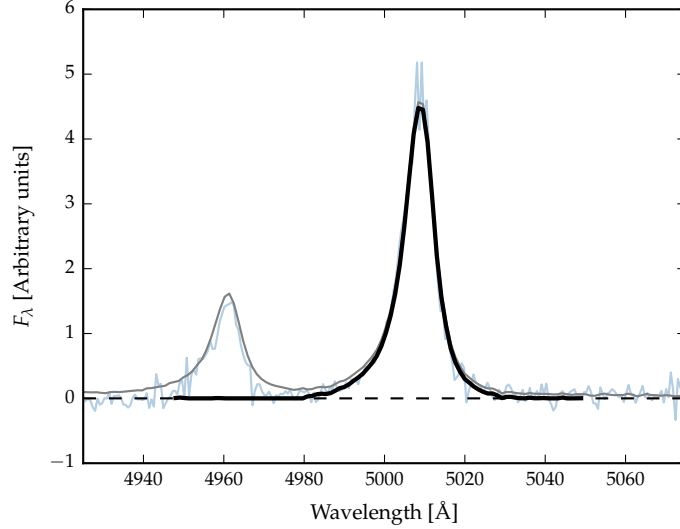


Figure 1.16: [O III] emission in J002952+020607. The data is shown in blue, and the ICA spectrum in grey. The first three ICA components have been subtracted from both the ICA composite and the data. The black curve shows the reconstructed [O III] $\lambda$ 5008 profile.

### *Reconstructing the [O III] profile*

In order to measure non-parametric line parameters, e.g.  $v_{10}$ , we must first reconstruct the [O III] emission. It is fortunate that most of the [O III] emission is in just three of the ICA components; the remaining three contribute very little. Therefore, we can set the first three weights to zero to leave only the [O III] emission. The four correction components are also included.

We define the boundaries of [O III] $\lambda$ 5008 as being between 4950 and 5500Å. The blue limit is close to the peak of the [O III] $\lambda$ 4960 line, and so to recover the intrinsic profile we instead use the blue wing of [O III] $\lambda$ 4960. We use the emission from 4980-5050Å, and from 4900-(4980-(5008.2-4960.3)). The

blue window is then shifted by (5008.2-4960.3) to reconstruct the blue wing of the [O III] $\lambda$ 5008 line. We then subtract a constant, because the flux does not always go to zero (suggests that there is probably flux which is not due to [O III] emission in components four to six).

An examples of a reconstructed [O III] emission-line is shown in Figure 1.16. At present I am summing the flux all the way from 4950Å. However, this is quite a lot of flux to sum up, and we can't ascribe this flux to the wing of the [O III] emission with any certainty. This is borne out by the fact that there are quite large differences between, for example,  $v_{10}$  measured from the Gaussian fit and  $v_{10}$  measured from the ICA fit.

Unfortunately, there are systematic differences between the line-width estimates from the Gaussian reconstructions and the ICA reconstructions, particularly for broad-line objects. The current way of doing the ICA reconstruction of the [O III] line ignores any cross-talk between the components and there is potentially flux being ascribed to the line that could be coming from some other component. We can solve this by finding some more representative broad [O III] lines in SDSS from which to derive the components as well as producing a set of components for [O III] only. Therefore we don't use these reconstructions and leave this for future work.

#### 1.8.4 ICA fits

In Figure 1.17 we show the relative weights of each of the six positive ICA components. Also shown are the same measurements for a sample of low-redshift, low-luminosity AGN. We want to examine whether or not there are systematic differences between these two samples.

We see that [O III] core emission is weaker in the more luminous sample, but the strength of the wing component is similar. Shen and Ho, (2014) showed that the strength of the core [O III] component decreases with quasar luminosity and optical Fe II strength faster than the wing component, leading to overall broader and more blueshifted profiles as luminosity and Fe II strength (or C IV blueshift) increases. Shen and Ho, (2014) suggested that a stable NLR is being removed by the outflowing material. Similarly, Zhang et al., (2011) found that the more the peak of the [O III] line is blueshifted, the more the core component decreases dramatically, while the blue wing changes much less. Therefore, there is an anti-correlation between the strength

*Paul - include figure because description is complicated. But I might ditch this altogether.*

*Paul: Think you need to say something about the SDSS quasar sample which was used to generate the ICA components. I would at that point herald that the limited spectral diversity in the SDSS sample is likely to prove an issue, even if the ICA approach looks promising.*

*Paul: Paragraph belongs with material above*

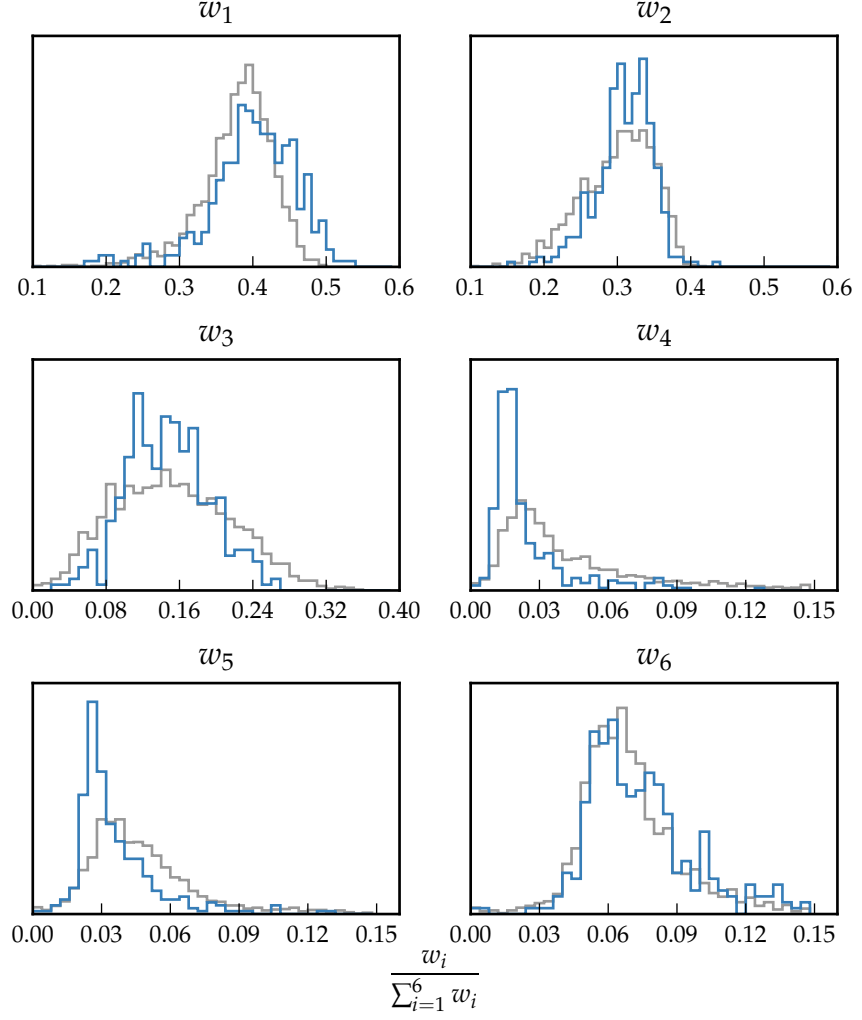


Figure 1.17: The relative weight in each of the six positive ICA components for the high-luminosity (blue) and low luminosity samples (grey). In the high-luminosity sample Fe II emission is stronger (component  $w_1$ ). The core [O III] emission (components  $w_4$ ,  $w_5$ ) is weaker but the strength of the blueshifted wing ( $w_6$ ) is the same.

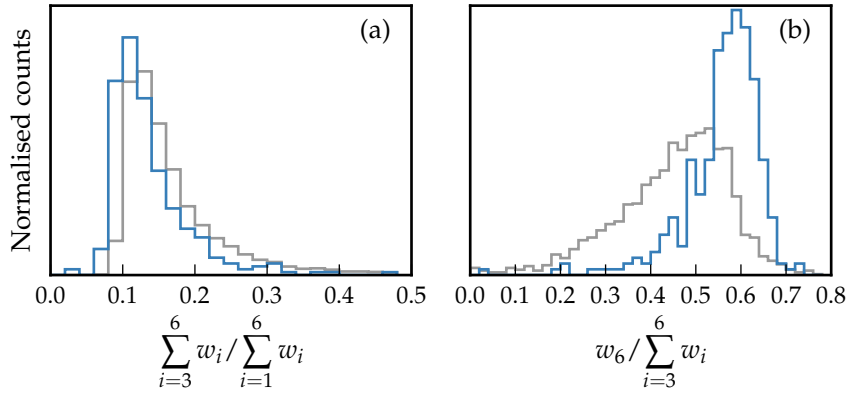


Figure 1.18: The relative weight in the three ICA components corresponding to [O III] emission (*left*) and the relative weight of the component most closely related to blueshifted [O III] emission relative to all three [O III] components (*right*). [O III] emission is weaker in the high-luminosity sample, but the relative contribution from the blueshifted component to the total [O III] emission is higher.

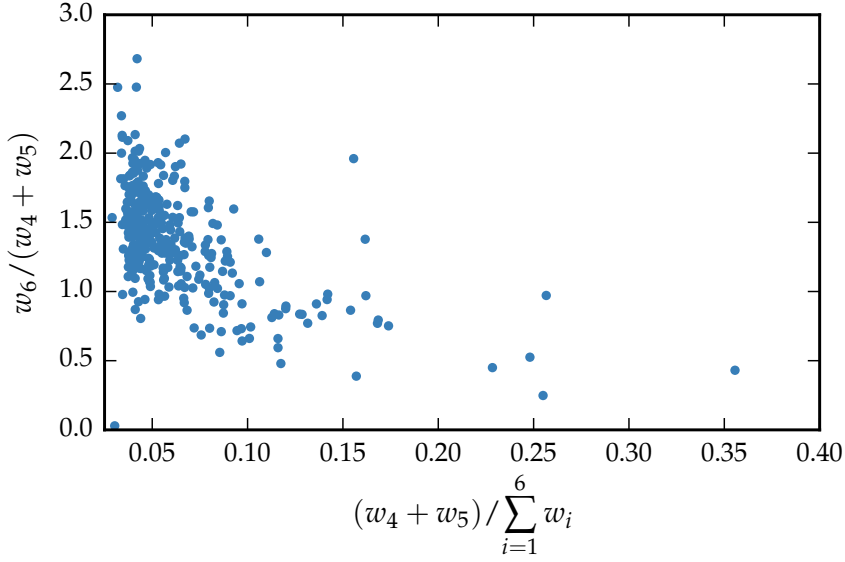


Figure 1.19: Weight in the [O III] wing relative to the weight in the [O III] core emission versus the strength of the core [O III] emission. The blue-asymmetry of the [O III] emission increases as the strength of the core component decreases.

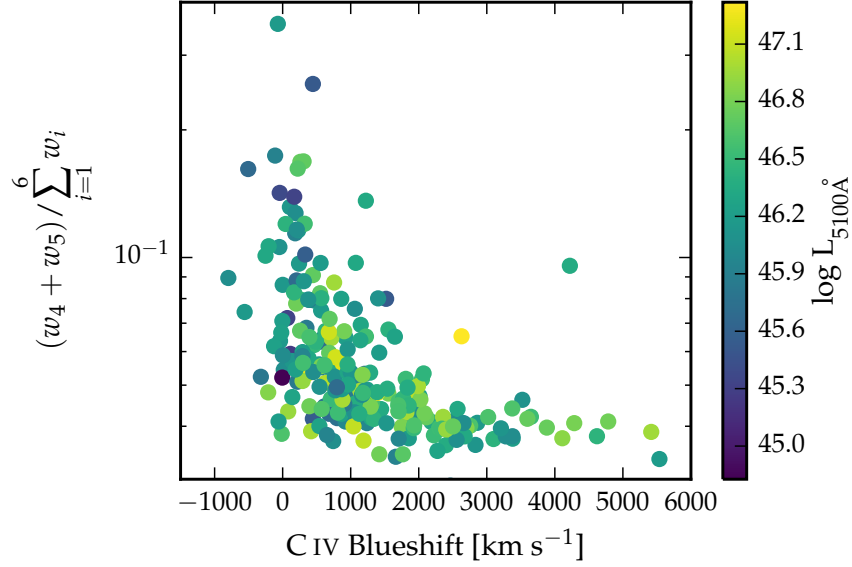


Figure 1.20: The ICA component weight  $w_4$ , which is a proxy for the strength of core [O III], as a function of the C IV blueshift. The C IV blueshift is measured relative to the near-infrared ICA redshift.

of the core component and the relative strength of the wing component (Figure 1.19).

To show this phenomenon more clearly, we plot the relative [O III] strength and the [O III] wing/core ratio in the high/low luminosity samples (Figure 1.19). We see that [O III] is weaker in the high luminosity sample, but that the wing component is much stronger relative to the core component. .

*Similar to  
behaviour of C IV?  
Would suggests  
that the mechanism  
producing the two  
correlations is the  
same*

#### EV1 correlations

In Figure 1.20 we show how the [O III] strength varies as a function of the C IV blueshift. There is a very well defined relation: when C IV is strongly blueshifted [O III] is very weak. This is very similar to what we found when we used Gaussian functions to model the emission. The correlation between C IV blueshift and [O III] EQW is shown in a different way in Figure 1.21. Here we divide our sample into four bins according to the C IV blueshift. From the quasars in each C IV blueshift bin we then find then generate an ICA spectrum using the median weights from each quasar. The differences in the spectra as a function of the C IV blueshift are dramatic. [O III] becomes progressively weaker and more blueshifted. The anti-correlation



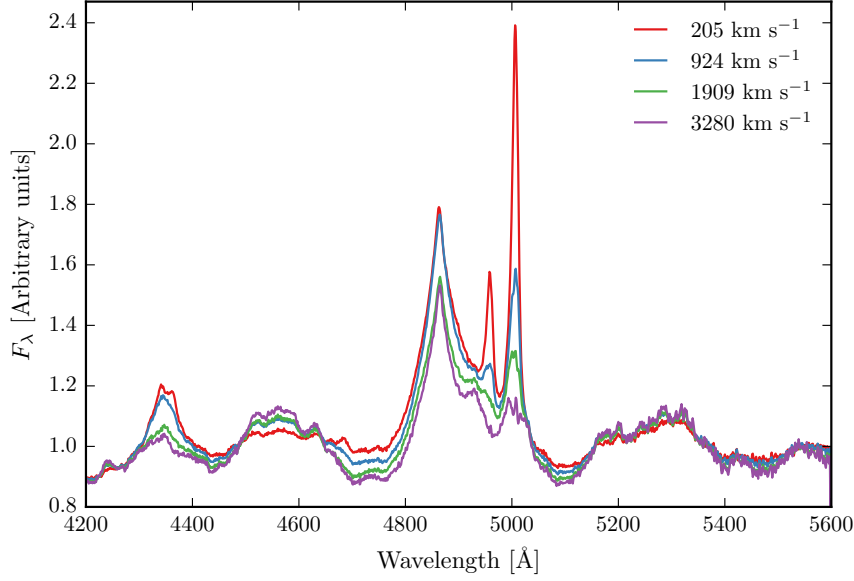


Figure 1.21: Median ICA-reconstructed spectra as a function of the C IV blueshift.

with Fe III and the blue-ward Fe II also clear, but there is no change in the redward Fe II.

#### Updating EV1

The ICA can be thought of as update on EV1. EV1 is a quest for HR diagram for quasars - clearly very important. The spectral diversity is encapsulated in the EV1 components. Most of the variance in EV1 is the anti-correlation between the strengths of [O III] and Fe II. So at one end we have objects with strong Fe II and weak [O III], and at the other end objects with weak Fe II and strong [O III]. Other properties, including the C IV blueshift and the H $\beta$  FWHM, also change systematically. Our work shows that the ICA component weights change systematically along the EV1 sequence.

Accurate systemic redshift estimates are essential in a number of applications, and researchers have devoted a large amount of telescope time to obtaining near-infrared spectra to access [O III] for this purpose. HI, CO and absorption line measures of the host galaxy rest frame suggest that [O III] usually gives consistent results within 200 km s<sup>-1</sup> (de Robertis 1985; Whittle 1985; Wilson & Heckman 1985; Condon et al. 1985; Stripe 1990; Alloin et al. 1992; Evans et al. 2001). However, our work shows that at high luminosities this can result in large er-

*Just present this as an idea for future work right at the end rather than having this sandwiched in the middle.*

rors (profile can be dominated by blueshifted component, Fe II emission can be improperly subtracted, or [O III] might not be detected at all. [O III] is weaker and broader so it is more difficult to detect and measure [O III] accurately for these luminous quasars (to for instance obtain reliable redshift estimates based on [O III])

#### 1.8.5 *Future work*

Pros:

It is less sensitive to the spectral S/N, and the component weights do not need to be constrained. It is therefore much simpler to apply than fitting multiple Gaussians.

Cons:

The components were calculated using a set of lower-redshift, lower-luminosity AGN, and quasar spectra are known to vary systematically as a function of luminosity. For example, the [O III] line is typically broader in more luminous quasars. Because there are so few objects with very broad [O III] in the low-redshift sample, the ICA reconstruction fails to reproduce the broadest [O III] profiles in our sample.

Cross-talk between components.

The size of the narrow line region is roughly expected to scale as  $L^{0.5}$  (e.g. Netzer et al., 2004). However, for high luminosity quasars with strong [O III] this gives NLR sizes which are unreasonably large ( $\sim 100$  kpc; Netzer et al., 2004).

See extra text from Brotherton paper. I could be confused here, but I think the Netzer argument goes that the nlr size increase with luminosity because there are more ionising photons. but then you run out of nlr to ionise. the luminosity of the quasar keeps increasing but the luminosity of the nlr flattens out. so the eqw starts to decrease. but we see a huge scatter in eqw at high luminosities. we can relate this to the C IV blueshift, which I don't think Netzer will have been able to.

We see a correlation between the [O III] velocity width and asymmetry. As the line gets broader it gets more blue-asymmetric. One interpretation of this is that the strength of the narrow core is decreasing, leading to a broader and more blueshifted profile (e.g. Shen and Ho, 2014).

[O III] is broader, which is consistent with these quasars having more massive BHs. [O III] also shows stronger blue asymmetries, suggesting that outflows are stronger/more prevalent at these higher luminosities/redshifts. The luminous

blueshifted broad wing and the extremely broad profile reveals high-velocity outflowing ionized gas. Our results therefore suggest that kilo-parsec-scale outflows in ionized gas are common in this sample of high-luminosity, high-redshift quasars.

## 1.9 RADIO

Radio properties would be interesting. Have another look, and at least say we tried and there aren't enough objects.

