

# CONTENTS

---

1	CORRECTING C IV-BASED VIRIAL BLACK HOLE MASSES	1
1.1	Introduction	1
1.2	Quasar Sample	4
1.3	Spectral Measurements	5
1.3.1	C IV	7
1.3.2	H $\alpha$	8
1.3.3	H $\beta$ and [O III]	10
1.3.4	Fitting procedure	10
1.3.5	Spectra removed from sample	11
1.3.6	Emission-line parameter uncertainties	13
1.3.7	Contemporaneity of spectra	14
1.3.8	Quasar monochromatic luminosity	15
1.3.9	Characterising the emission-line widths	15
1.4	An empirical correction to C IV-based virial BH mass estimates	18
1.4.1	H $\alpha$ /H $\beta$ FWHM comparison	18
1.4.2	Measuring the quasar systemic redshift	23
1.4.3	Balmer/C IV line widths as a function of C IV-blueshift	24
1.4.4	C IV-based virial BH mass estimates	29
1.4.5	C IV-derived BH masses at low C IV blueshift	30
1.5	Practical application of the C IV-based correction to virial BH mass estimates	31
1.5.1	Recipe for unbiased C IV based BH masses	31
1.5.2	Systematic trends in residuals	36
1.5.3	Effectiveness of the C IV blueshift based correction to BH masses	38
1.5.4	Comparison to previous prescriptions	42
1.6	Population trends with C IV blueshift	45
1.6.1	The BAL parent population	48

1.6.2	The frequency of quasars with high accretion rates	50
1.7	Conclusions	51
1.8	Catalogue of derived properties	51

## LIST OF FIGURES

---

Figure 1.1	Composite spectra of the C IV emission-line as a function of C IV blueshift for SDSS DR7 quasars. 3
Figure 1.2	Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the C IV FWHM with Shen et al., (2011). 8
Figure 1.3	Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the H $\alpha$ FWHM with Shen and Liu, (2012). 9
Figure 1.4	Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the H $\beta$ FWHM with Shen, (2016). 11
Figure 1.5	Model fits to continuum-subtracted H $\alpha$ , H $\beta$ , and C IV emission in four quasars, chosen to represent the range of S/N and line shapes present in the catalogue. 12
Figure 1.6	The redshift and luminosity distributions of the spectra removed from our H $\alpha$ /C IV (a, b) and H $\beta$ /C IV (c, d) samples. 14
Figure 1.7	The FWHM, dispersion ( $\sigma$ ) and shape (FWHM/ $\sigma$ ) of C IV as a function of the C IV blueshift. 17
Figure 1.8	Comparison of the C IV line profiles of J123611+112922 and J152529+292813. 19
Figure 1.9	Comparison of H $\alpha$ and H $\beta$ FWHM measurements for 99 quasars. 20
Figure 1.10	One- and two-dimensional projections of the MCMC sampling of the posterior distribution from the fit in Figure 1.9. 21

- Figure 1.11 The H $\alpha$  and H $\beta$  emission-line regions in the median composite spectrum, shown as function of the velocity shift from the respective predicted line peak wavelengths. 22
- Figure 1.12 C iv FWHM relative to H $\alpha$  FWHM (a), and C iv based BH mass (BHM) compared to H $\alpha$  based mass (b), both as a function of the C iv blueshift. 25
- Figure 1.13 C iv FWHM relative to H $\beta$  FWHM (a), and C iv based BH mass (BHM) compared to H $\beta$  based mass (b), both as a function of the C iv blueshift. 26
- Figure 1.14 One- and two-dimensional projections of the MCMC sample of the posterior distribution for a linear fit to the FWHM C iv/H $\alpha$  ratio as a function of the C iv blueshift. 27
- Figure 1.15 The distribution of the orthogonal displacement of each data point from the best-fitting linear relationship in the fit to FWHM(C iv)/FWHM(H $\alpha$ ) as a function of the C iv blueshift. 28
- Figure 1.16 Rest-frame EQW versus blueshift of the broad C iv emission-line for 32,157 SDSS DR7 quasars at  $1.6 < z < 3.0$ . 33
- Figure 1.17 Same as Figure 1.12a, with the marker colour representing the H $\alpha$  FWHM. 38
- Figure 1.18 Rest-frame EQW versus blueshift of the broad C iv emission-line for 32,157 SDSS DR7 quasars at  $1.6 < z < 3.0$  and our sample. 39
- Figure 1.19 Comparison of the C iv- and H $\alpha$ -based BH masses before and after applying the C iv blueshift-based correction to the C iv FWHM. 41
- Figure 1.20 Comparison of the C iv and H $\alpha$  line dispersion,  $\sigma$ . 42
- Figure 1.21 Comparison of BH mass estimates derived from C iv and H $\alpha$  as a function of the C iv blueshift. 43

Figure 1.23	The FWHM, dispersion ( $\sigma$ ) and shape (FWHM/ $\sigma$ ) of H $\alpha$ as a function of the C iv blueshift. 46
Figure 1.24	H $\alpha$ -derived Eddington ratio versus C iv blueshift. 47

---

In Figures ??, 1.16, 1.18 and ?? regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator.

## LIST OF TABLES

---

Table 1.1	The numbers of quasars with reliable H $\alpha$ and H $\beta$ line measurements, and the spectrographs and telescopes used to obtain the near-infrared spectra 5
Table 1.2	The number of spectra removed from our sample by the cuts described in Section 1.3.5. 11
Table 1.3	The fractional error on the corrected BH mass as a function of C iv blueshift for different uncertainties in the quasar systemic redshift. 32
Table 1.4	The format of the table containing the emission-line properties from our parametric model fits. The table will be available in machine-readable form online. 53

## ACRONYMS

---

AGN	Active Galactic Nuclei
NLR	Narrow Line Region
BLR	Broad Line Region
EV <sub>1</sub>	Eigenvector 1
ICA	Independent Component Analysis
PCA	Principal Component Analysis
SDSS	Sloan Digital Sky Survey
BOSS	Baryon Oscillation Spectroscopic Survey
UV	Ultra-Violet
EQW	Equivalent width
S/N	Signal-to-noise ratio
BH	Black Hole
SED	Spectral Energy Distribution
IR	Infrared
NIR	Near-infrared
FWHM	Full-Width-at-Half-Maximum

# CORRECTING C IV-BASED VIRIAL BLACK HOLE MASSES

---

## 1.1 INTRODUCTION

The goal of better understanding the origin of the correlation between the masses of super-massive BHs and the masses of host-galaxy spheroids has led to much work focussing on the properties of quasars and AGN at relatively high redshifts,  $z \gtrsim 2$ . Extensive reverberation-mapping campaigns have been used to calibrate single-epoch virial-mass estimates which use the velocity widths of the hydrogen Balmer emission-lines and the nuclear continuum luminosity to provide reliable BH masses. Single-epoch virial BH mass estimates using H $\beta$  are possible up to redshifts  $z \sim 0.7$ , and the technique has been extended to redshifts  $z \sim 1.9$  via the calibration of the broad Mg II  $\lambda\lambda 2796, 2803$  emission-line (McLure and Jarvis, 2002; Onken and Kollmeier, 2008; Wang et al., 2009; Rafiee and Hall, 2011). At redshifts  $z \gtrsim 2$ , however, ground-based statistical studies of the quasar population generally have no access to the rest-frame optical and near-ultraviolet spectral regions.

The C IV  $\lambda\lambda 1548, 1550$  emission doublet is both relatively strong in the majority of quasars and visible in modern optical spectra, such as those provided by SDSS, to redshifts exceeding  $z \sim 5$ . C IV-derived BH masses have therefore become the standard (e.g. Vestergaard and Peterson, 2006; Park et al., 2013) for both individual quasars and in studies of quasar population demographics.

Currently, the number of reverberation mapped quasars is small ( $\sim 50$  quasars; Park et al., 2013) and restricted to low redshifts and luminosities. The luminosities of quasars at redshifts  $z \gtrsim 2$  are much greater than in the reverberation mapped sample, and the reliability of the existing calibration involving C IV FWHM velocity measurements and ultraviolet luminosity is not established definitively when extrapolating to high-redshifts and luminosities. While some authors have found good agreement between BH mass-estimates based on C IV and H $\beta$  (e.g. Vestergaard and Peterson, 2006; Assef et al., 2011; Tilton and Shull, 2013), others have questioned the consistency

(e.g. Baskin and Laor, 2005; Trakhtenbrot and Netzer, 2012; Shen and Liu, 2012).

In contrast to a number of low-ionisation emission-lines, such as  $\text{Mg II}$ , the  $\text{C IV}$  emission has long been known to exhibit significant asymmetric structure, with an excess of flux to the blue of the predicted rest-frame transition wavelength (Gaskell, 1982). More recent work (e.g. Sulentic, Marziani, and Dultzin-Hacyan, 2000; Richards et al., 2011) has established that the extent of ‘blueshifts’ in the  $\text{C IV}$  emission correlates with a number of properties of quasar SEDs. A fundamental assumption on which single-epoch virial BH mass estimates are based is that the widths of the broad emission-lines are directly related to the virial motions of the emitting clouds moving in the gravitational potential of the central BH. While the physical origin of the blueshifted emission has not been established there is a consensus that the associated gas is not tracing virial-induced velocities. A favoured interpretation associates the blueshifted emission with out-flowing material (see Netzer, 2015, for a recent review), reaching velocities significantly larger than virial-induced velocities associated with the BH (e.g. Sulentic et al., 2007; Richards et al., 2011). These outflows, most likely, result from the presence of a radiation line-driven accretion-disc wind (e.g. Konigl and Kartje, 1994; Murray et al., 1995; Proga, Stone, and Kallman, 2000; Everett, 2005; Gallagher et al., 2015; Higginbottom and Proga, 2015).

Figure 1.1 shows the shape of the  $\text{C IV}$ -emission in composite spectra constructed from SDSS DR7 quasars as a function of  $\text{C IV}$  blueshift. The profiles show how, at large values of blueshift ( $\gtrsim 2000 \text{ km s}^{-1}$ ) the  $\text{C IV}$ -profile is displaced to the blue by amounts comparable to the FWHM of the profile. At fixed emission-line EQW, virtually the entire  $\text{C IV}$ -profile appears to shift blueward and the change in line shape is not simply an enhancement of flux in the blue wing of a still identifiable symmetric component. While gravity almost certainly plays a key role, determining the escape velocity for out-flowing material for example, it is clear that the virial assumption, on which single-epoch BH mass measurements are predicated, is not straightforwardly applicable for the  $\text{C IV}$  emission-line in quasars exhibiting large blueshifts. In general, researchers studying quasar demographics at high-redshift adopt estimates of BH masses based on the width of  $\text{C IV}$ -emission, without reference to the blueshift of the  $\text{C IV}$ -emission (e.g. Vestergaard, 2004; Kollmeier et al., 2006; Gav-



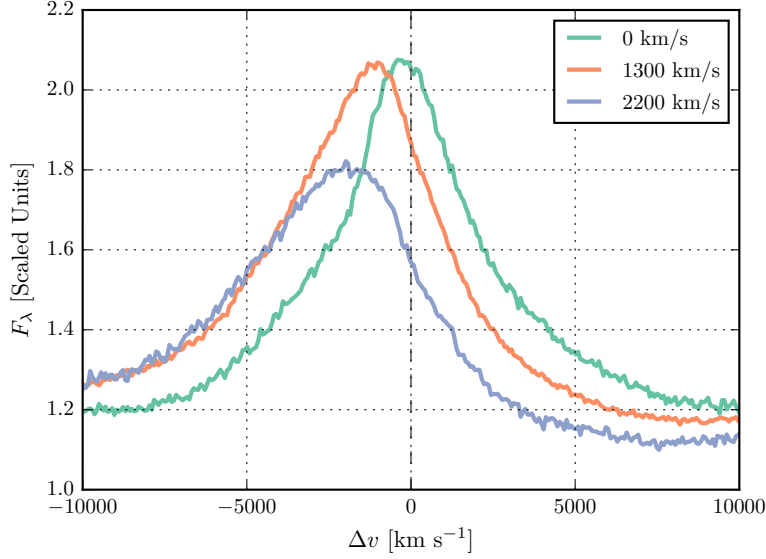


Figure 1.1: Composite spectra of the C IV emission-line as a function of C IV blueshift for SDSS DR7 quasars. Quasars classified as BALs, or possessing strong associated absorbers have been excluded, and the composite-spectra shown are derived using an arithmetic mean of a minimum of 200 spectra at each blueshift. Virtually the entire C IV-profile appears to shift blueward and the change in line shape is not simply an enhancement of flux in the blue wing of a still identifiable symmetric component. In order of increasing C IV blueshift, the composite spectra have FWHM 4870, 5610, and 6770 km s<sup>-1</sup> and EQW 33.1, 31.6, and 28.8 Å.

ignaud et al., 2008; Vestergaard et al., 2008; Vestergaard and Osmer, 2009; Kelly et al., 2010; Kelly and Shen, 2013). As a consequence, BH masses derived from C IV emission-line velocity-widths are systematically biased compared to masses from the Balmer lines (e.g. Shen et al., 2008; Shen and Liu, 2012).

As highlighted by Richards et al., (2011), the sample of reverberation mapped quasars includes a restricted range of the C IV emission-line shapes seen in the quasar population. In particular, the reverberation mapped objects generally possess high C IV EQWs and low C IV-blueshifts. Nevertheless, the derived scaling relations based on the reverberation-mapped sample are regularly applied to the quasar population with low C IV EQWs and/or large C IV-blueshifts, where any non-virial outflow-related contribution to the dynamics is significant.

In recent literature, attempts have been made to minimise the influence of the systematic non-virial contribution to the

C iv emission on estimates of the BH mass. Strategies include (i) significantly reducing the dependence of the derived masses on the emission-line velocity width (e.g. from the  $V^2$  dependence predicted assuming a virialized broad line region to just  $V^{0.56}$  in Park et al. 2013; see also Shen and Liu 2012), (ii) adopting a measure of emission-line velocity-width that is relatively insensitive to changes in the core of the emission-line profile (e.g. Denney et al., 2013) and (iii) estimating the amplitude of the non-virial contribution to the C iv emission-line via comparison with other ultraviolet emission-lines (e.g. Si iv+O iv  $\lambda$ 1400 in Runnoe et al. 2013 and Brotherton et al. 2015). The increased number of quasars with high-quality spectra that cover both the observed-frame optical (where the redshifted C iv appears) and near-infrared (where H $\beta$  and H $\alpha$  lie) enables us to take a rather different approach in this Chapter. We will use properties of the C iv emission-line itself to reduce, or even remove, the systematic bias in the BH mass estimates. Specifically, using the low-ionisation Balmer lines H $\alpha$  and H $\beta$  as reliable proxies for the virial velocity, we will measure empirically the systematic bias in C iv-based virial BH mass estimates as a function of the C iv emission-line blueshift.

## 1.2 QUASAR SAMPLE

We have compiled a sample of 307 non-BAL quasars at redshifts  $1.5 < z < 4$  with both optical and near-infrared spectra. Reliable emission-line properties were measured for 230 quasars (Section 1.3.5), with 164 possessing H $\alpha$  line measurements and 144 H $\beta$  line measurements. This will allow us to directly compare virial BH mass estimates based on the C iv line-width with estimates based on the line-widths of the low-ionisation Balmer lines H $\alpha$  and H $\beta$ . The sample is considerably larger than previous studies of the rest-frame optical spectra of high- $z$  quasars (e.g. Shen and Liu, 2012). As we demonstrate in Section 1.5.3, the quasars have C iv blueshifts of up to  $\sim 5000 \text{ km s}^{-1}$ , and span the full range observed in the population.

The near-infrared data has been described in Chapter ?? and the telescopes/spectrographs used are summarised in Table 1.1. Corresponding optical spectroscopy was obtained from the SDSS (70 quasars), BOSS (126 quasars) and Hamburg-ESO surveys (15 quasars), and with VLT/UVES (11 quasars) and VLT/XSHOOTER (8 quasars). Many of the quasars in the SDSS DR7 catalogue have been re-observed as part of BOSS. As the

Table 1.1: The numbers of quasars with reliable  $H\alpha$  and  $H\beta$  line measurements, and the spectrographs and telescopes used to obtain the near-infrared spectra

Spectrograph	Telescope	$H\alpha$ Sample	$H\beta$ Sample
FIRE	MAGELLAN	18	19
GNIRS	GEMINI-N	22	17
ISAAC	VLT	0	4
LIRIS	WHT	15	0
NIRI	GEMINI-N	0	12
SINFONI	VLT	2	25
SOFI	NTT	47	23
TRIPLESPEC	ARC-3.5m	33	20
TRIPLESPEC	P200	23	19
XSHOOTER	VLT	4	7
Total		164	144

BOSS-spectra typically have higher S/N than the SDSS DR7 spectra, we have used the BOSS spectra when available. Once more, further details are provided in Chapter ?? . We have subdivided our sample into two overlapping groups: quasars with reliable  $H\alpha$  line measurements (the ‘ $H\alpha$  sample’) and quasars with reliable  $H\beta$  measurements (the ‘ $H\beta$  sample’).

### 1.3 SPECTRAL MEASUREMENTS

Conventionally, single-epoch virial estimates of the BH mass are a function of the line-of-sight velocity width of a broad emission-line and the quasar luminosity. The velocity width is a proxy for the virial velocity in the broad line region (BLR) and, as revealed in reverberation-mapping studies, the luminosity is a proxy for the typical size of the BLR (the  $R - L$  relation; e.g. Kaspi et al., 2000; Kaspi et al., 2007). Most reverberation mapping campaigns have employed  $H\beta$  time-lags and velocity widths, but the line-widths of  $H\alpha$  and  $Mg\ II\lambda 2800$  have been shown to yield consistent BH masses (e.g. McLure and Jarvis, 2002; Greene and Ho, 2005; Onken and Kollmeier, 2008; Shen et al., 2008; Wang et al., 2009; Rafiee and Hall, 2011; Mejía-Restrepo et al., 2016). In Section 1.4.1, we verify that the  $H\alpha$  and  $H\beta$  line-widths yield consistent BH for the 99 quasars in our sample with measurements of both.

In our work, a robust measure of the  $C\ IV$  emission-line ‘blueshift’ provides the basis for the corrected  $C\ IV$  velocity-

width measurements, and hence BH masses. The effectiveness of the scheme is validated via a direct comparison of the C IV velocity-widths to the Balmer emission velocity-widths in the same quasars. Our process is as follows. First, an accurate measure of the quasar's systemic redshift is required, for which we adopt the centre of the Balmer emission, where the centre,  $\lambda_{\text{half}}$ , is the wavelength that bisects the cumulative total flux. Balmer emission centroids are available for all quasars in the catalogue but we verify that the measure is relatively unbiased through a comparison of the centroids to the wavelengths of the peak of the narrow [O III]  $\lambda\lambda 4960, 5008$  doublet for the subset of spectra where both are available (Section 1.4.2). Second, the blueshift of the C IV emission-line is determined. Again, we adopt the line centroid to provide a robust measure of the C IV emission blueshift. The blueshift (in  $\text{km s}^{-1}$ ) is defined as  $c \times (1549.48 - \lambda_{\text{half}}) / 1549.48$  where  $c$  is the velocity of light and  $1549.48 \text{ \AA}$  is the rest-frame wavelength for the C IV doublet<sup>1</sup>. Positive blueshift values indicate an excess of emitting material moving towards the observer and hence out-flowing from the quasar.

Emission-line velocity widths are derived from the full-width-at-half-maximum (FWHM) of the lines but we also compute the line dispersion (calculated from the flux-weighted second moment of the velocity distribution) as some authors have claimed this provides a better estimate of the virial velocity (Denney et al., 2013).

To minimise the impact of the finite S/N of the quasar spectra and the presence of absorption features superposed on the broad emission-lines we first fit a parametric model to the continuum and the emission-lines. The particular form of the model parametrizations is not important and the fits are used only to provide robust line parameters, such as the centroid  $\lambda_{\text{half}}$ , and FWHM, which are measured non-parametrically from the best-fitting model. The models used and the fitting procedure are described below. The issues involved in deriving parameters for broad emission-lines from spectra of modest S/N – for example, subtraction of narrow line emission, subtraction of Fe II emission – have been covered comprehensively by other authors (e.g. Shen et al., 2011; Shen and Liu, 2012; Den-

<sup>1</sup> The adopted C IV rest-frame wavelength assumes an optically thick BLR, in which case the contribution from each component is equal. Adopting a 2 : 1 ratio (appropriate for an optically thin BLR) changes the blueshifts by  $\sim 80 \text{ km s}^{-1}$ .

ney et al., 2013; Shen, 2016) and, as far as possible, we follow standard procedures described in the literature.

### 1.3.1 C iv

We first define a power-law continuum,  $f(\lambda) \propto \lambda^{-\alpha}$ , with the slope,  $\alpha$ , determined using the median values of the flux in two continuum windows at 1445-1465 and 1700-1705 Å. The continuum emission is subtracted from the spectra, which is then transformed from wavelength units into units of velocity relative to the rest-frame line-transition wavelength for the C iv doublet. The parametric model is ordinarily fit within the wavelength interval 1500-1600 Å (corresponding to approximately  $\pm 10\,000\text{ km s}^{-1}$  from the rest-frame transition wavelength), a recipe that is commonly adopted (e.g. Denney et al., 2013). The line-window was extended if more than 5 per cent of the total flux in the profile was present blueward of the short wavelength limit. Narrow absorption features, which are frequently found superimposed on C iv emission, were masked out during the fit.

The C iv emission was fit with sixth-order Gauss-Hermite (GH) polynomials, using the normalisation of van der Marel and Franx, (1993) and the functional forms of Cappellari et al., (2002). We allowed up to six components, but in many cases a lower order was sufficient (40 and 45 per cent were fit with second- and fourth-order GH polynomials respectively). GH polynomials were chosen because they are flexible enough to model the often very asymmetric C iv line profile. The flip-side of this flexibility, however, is that the model has a tendency to over-fit when spectra possess low S/N. The fits were therefore carefully checked visually and the number of components reduced if over-fitting was evident.

We find that using the commonly employed three-Gaussian component model, rather than the GH polynomials, resulted in only marginal differences in the line parameters. Our best-fit parameters are also in good agreement with Shen et al., (2011), who employ a multi-Gaussian parametrization. In Figure 1.2 we compare our measurements of the C iv FWHM from the 71 SDSS DR7 spectra in our sample with the measurements published in Shen et al., (2011). There is a very strong agreement between our measurements, with a scatter of 0.05 dex ( $200\text{ km s}^{-1}$ ).

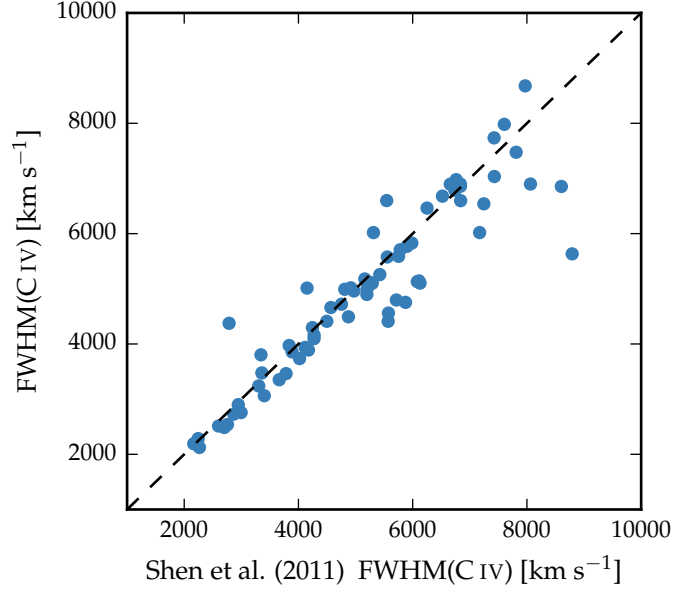


Figure 1.2: Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the C IV FWHM with Shen et al., (2011).

### 1.3.2 $H\alpha$

A power-law continuum is fit using two continuum windows at 6000-6250 and 6800-7000 Å. The continuum-subtracted flux is then fit in the wavelength interval 6400-6800 Å. We adopt a rest-frame transition wavelength of 6564.89 Å to transform wavelengths into equivalent Doppler velocities. The broad component of  $H\alpha$  is fit using one or two Gaussians, constrained to have a minimum FWHM of  $1200 \text{ km s}^{-1}$ . When two Gaussians are used, the velocity centroids are constrained to be the same.

The emission-line profiles of both  $H\beta$  and  $H\alpha$  frequently include a significant narrow component from the physically more extended NLR. Additional Gaussian components were included in our parametric model to fit the narrow component of  $H\alpha$  as well as  $[\text{N II}]\lambda\lambda 6548, 6584$  and  $[\text{S II}]\lambda\lambda 6717, 6731$ . This resulted in a better fit to the observed flux in 50 per cent of cases. We impose a  $1200 \text{ km s}^{-1}$  upper limit on the FWHM of all narrow lines and the amplitudes of all components must be non-negative. The relative flux ratio of the two  $[\text{N II}]$  components is also fixed at the expected value of 2.96. In 70 per cent of the spectra the  $[\text{O III}]\lambda\lambda 4960, 5008$  doublet is detected at moderate S/N in the  $H\beta$  region. In these cases the peak of the  $[\text{O III}]$

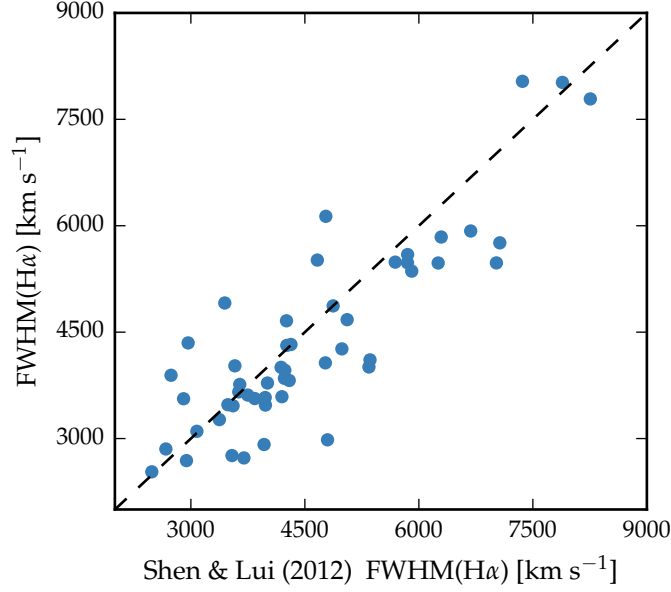


Figure 1.3: Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the  $H\alpha$  FWHM with Shen and Liu, (2012).

is used to fix the velocity offsets and the FWHMs of the narrow line components in the  $H\alpha$  region. For spectra where the  $[O\text{III}]$  doublet does not constrain the velocity and FWHM accurately, the narrow emission in the  $H\alpha$  and  $H\beta$  regions are fitted independently but, for each region, the individual narrow-line velocity offsets and the FWHMs are constrained to be identical. In these objects the narrow line contribution is generally weak, and so does not have a large effect on the line parameters we measure for the broad component.

The model described above is very similar to the one described in Shen and Liu, (2012) and Shen et al., (2011), the only major differences being that we do not fit the  $H\alpha$  and  $H\beta$  emission regions simultaneously and we fix the centroids of the Gaussian components used to fit the broad emission. In Figure 1.3 we plot our  $H\alpha$  FWHM measurements against the measurements published in Shen and Liu, (2012), for 51 quasars in common to both samples. There is a strong correlation and a scatter of just 0.07 dex.

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

Emission from optical Fe II is generally strong in the vicinity of  $H\beta$ . We therefore fit a combination of a power-law continuum and an optical Fe II template – taken from Boroson and Green, (1992) – to two windows at 4435-4700 and 5100-5535 Å. The Fe II template is convolved with a Gaussian, and the width of this Gaussian, along with the normalisation and velocity offset of the Fe II template, are free variables in the pseudo-continuum fit. We use the same model to fit the broad and narrow components of  $H\beta$  as was used with  $H\alpha$ . Each line in the  $[O\text{ III}]$  doublet is fit with two Gaussians, to model both the systemic and any outflow contributions. The peak flux ratio of the  $[O\text{ III}]$  4960 Å and 5008 Å lines is fixed at 1 : 3. As for the fit to the narrow lines in the spectral region around  $H\alpha$ , the width and velocity offsets of all the narrow components are set to be equal, and an upper limit of  $1200\text{ km s}^{-1}$  is placed on the FWHM.

The parametric model we fit to the  $H\beta/[O\text{ III}]$  emission region was very similar to the model employed by Shen, (2016). In Figure 1.4 we plot our  $H\beta$  FWHM measurements against the measurements published in Shen, (2016), for 39 quasars in common to both samples. As expected, we observe a very tight correlation, with a scatter of 0.04 dex.

### 1.3.4 *Fitting procedure*

Model parameters were derived using a standard variance-weighted least-squares minimisation procedure employing the Levenberg-Marquardt algorithm. Prior to the fit, the spectra were inspected visually and regions significantly affected by absorption or of low S/N were masked out.

In Figure 1.5 we present our parametric fits to the C IV,  $H\alpha$  and  $H\beta$  emission-lines in a handful of quasars, which have been chosen to illustrate the range of spectrum S/N and line shapes in the sample. The Doppler velocities have been shifted so that the  $H\alpha$  emission-line centroid is at  $0\text{ km s}^{-1}$ . The y-axes of the data-minus-model residual plots have been scaled by the spectrum flux errors. The median reduced- $\chi^2$  values in our  $H\alpha$ ,  $H\beta$  and C IV fits are 0.96, 1.58, and 0.91 respectively and, in general, there are no strong features observable in the model residuals. The only significant features seen in the residual C IV spectra correspond to the location of narrow absorption lines which were excluded in the fitting procedure.



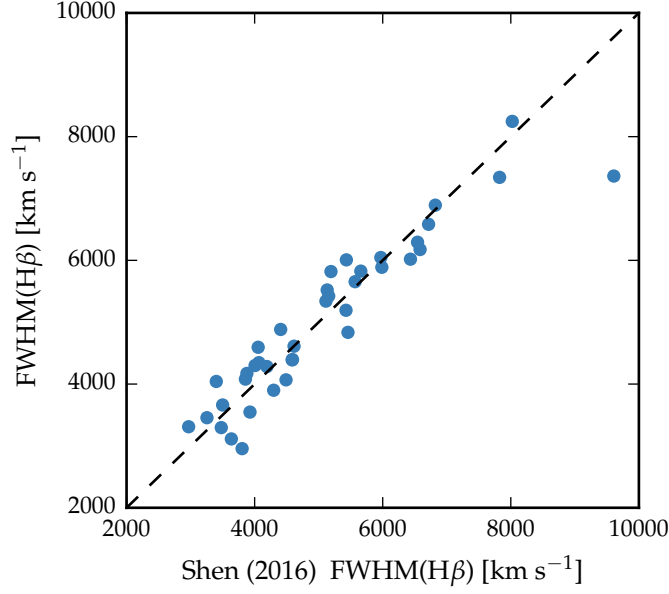


Figure 1.4: Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the  $H\beta$  FWHM with Shen, (2016).

Table 1.4 includes the line parameters of our best-fitting model for each line.

#### 1.3.5 Spectra removed from sample

Through visual inspection we flagged and discarded the spectra of quasars for which reliable emission-line parameters could not be obtained.

First, we flagged emission-lines in spectra that possessed insufficient S/N. A single minimum S/N threshold was not en-

Table 1.2: The number of spectra removed from our sample by the cuts described in Section 1.3.5.

		H $\alpha$ sample	H $\beta$ sample
Total		194	279
H $\alpha$ /H $\beta$	Wavelength	6	27
	S/N	8	83
C IV	Wavelength	6	5
	S/N	4	12
	Absorption	6	8
Total remaining		164	144

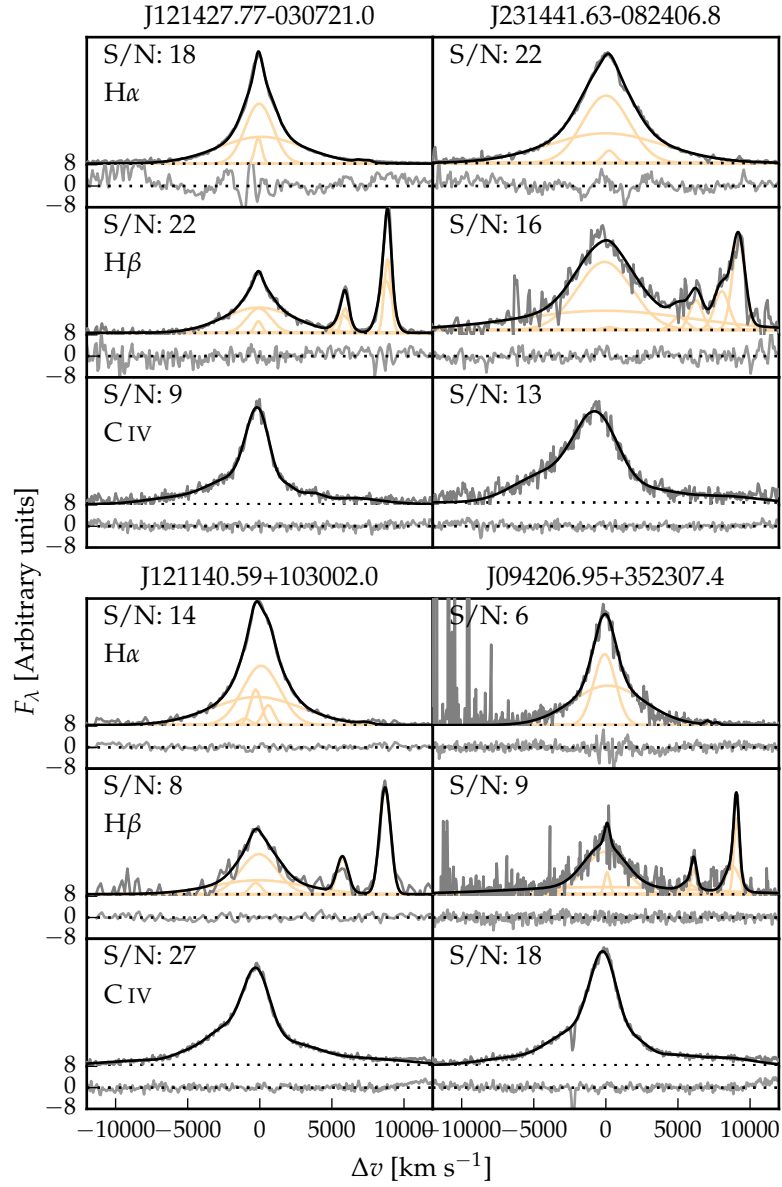


Figure 1.5: Model fits to continuum-subtracted H $\alpha$ , H $\beta$ , and C IV emission in four quasars, chosen to represent the range of S/N (indicated in the figure and given per  $150 \text{ km s}^{-1}$  pixel in the continuum) and line shapes present in the catalogue. The data is shown in grey, the best-fitting parametric model in black, and the individual model components in orange. The centroid of the broad H $\alpha$  emission is used to set the redshift, and  $\Delta v$  is the velocity shift from the line rest-frame transition wavelength. Below each fit we plot the data minus model residuals, scaled by the errors on the fluxes.

tirely effective and, instead, spectra were flagged when it was judged conservatively that no meaningful constraints could be placed on the velocity centroid and/or width of the emission-line.

Second, we flagged emission-lines where significant regions of the continuum and/or emission-line fell outside of the wavelength coverage of the spectra. Reliable continuum definition and subtraction is not straightforward for emission-lines so affected.

Third, we flagged C iv emission-lines because of strong, narrow absorption close to the peak of the line where reliable interpolation across the absorption, using our parametric model, was not possible.

The number of spectra that are removed by each cut is given in Table 1.2 and the distribution in redshift and luminosity is shown in Figure 1.6. Unsurprisingly, there is a preferential removal of intrinsically faint quasars, whose spectra can be of poorer S/N, and a loss of quasars at redshifts  $z \sim 2.6$  where the H $\alpha$  emission falls at the edge of the K-band. H $\beta$  is much weaker than H $\alpha$ , and the H $\beta$  spectra are generally of lower S/N. As a result, the fraction of H $\beta$  spectra that are flagged – 39 per cent – is particularly high.

### 1.3.6 *Emission-line parameter uncertainties*

The  $1\sigma$  error bars calculated from the covariance matrix in least-squares minimisation will underestimate the true uncertainties on the line parameters, since they do not account for systematic errors such as the significant uncertainty introduced in the continuum subtraction procedure. To calculate more realistic uncertainties on our fitted variables we employed a Monte Carlo approach. One thousand artificial spectra were synthesised, with the flux at each wavelength drawn from a Normal distribution (mean equal to the measured flux and standard deviation equal to the known error). Our emission-line fitting recipe was then implemented on each of these mock spectra. The uncertainty in each parameter is given by the spread in the best-fitting values from the one thousand realisations of the fitting routine. In some cases the standard deviation of the parameter distribution was biased by extreme values caused by bad fits<sup>2</sup>. We therefore chose to measure the spread in the parameter distribution by

<sup>2</sup> In the analysis of the real spectra such fits are identified via visual inspection.

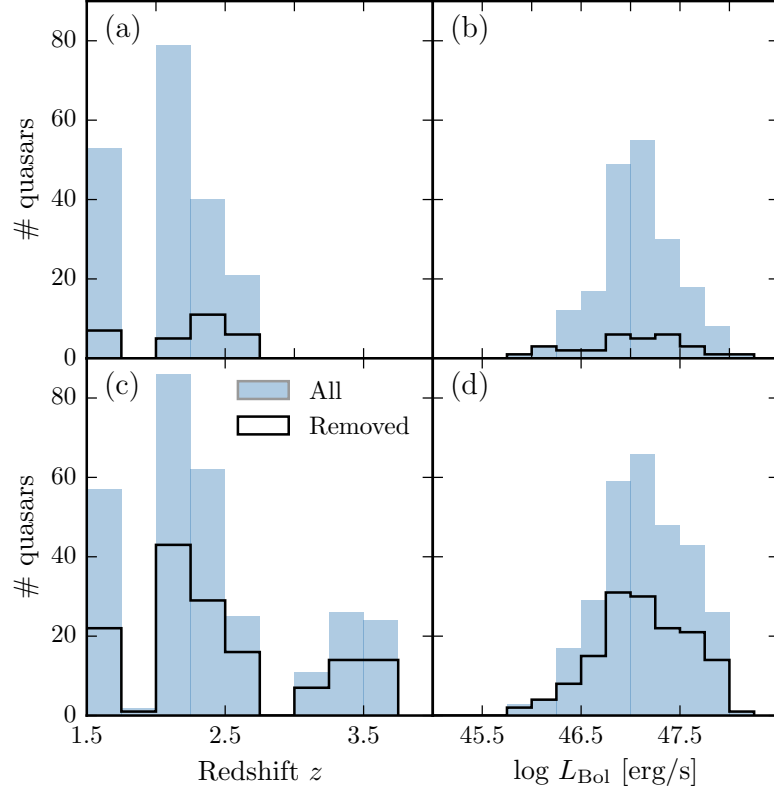


Figure 1.6: The redshift and luminosity distributions of the spectra removed from our H $\alpha$ /C iv (a, b) and H $\beta$ /C iv (c, d) samples.

fitting a composite model with two Gaussian components – one to model uncertainty in the parameter and the other any possible outlier component. The uncertainty in each line parameter was then taken to be the width of the narrower Gaussian. The uncertainties on all derived quantities, such as the BH mass, are propagated through by assuming that the uncertainties are uncorrelated and independent.

### 1.3.7 Contemporaneity of spectra

The epochs of the near-infrared and optical spectra can differ by many years. For example, the NTT SOFI spectra were taken  $\sim 14$  years after the SDSS spectra, and the VLT SINFONI spec-

tra 20 years or more after the Hamburg-ESO observations<sup>3</sup>. If the broad emission-line profiles varied significantly on these time-scales the relation between the C iv and Balmer line-width measurements could be blurred.

Cases do exist of dramatic changes in quasar spectra over short time-scales, but this phenomenon is rare (MacLeod et al., 2016). In our spectroscopic catalogue there are 112 SDSS DR7 quasars which are re-observed in BOSS and included in the DR12 quasar catalogue. The mean time elapsed between the two sets of observations is  $\sim 8$  years. The root-mean-square difference in the C iv FWHM measured from the BOSS and SDSS spectra is a modest  $\simeq 500 \text{ km s}^{-1}$ . Differences in the S/N of the spectra will make a substantial contribution and the scatter due to true variations in the C iv velocity-width will be significantly smaller than  $500 \text{ km s}^{-1}$ . We conclude therefore that any intrinsic changes with time do not materially affect the emission-line measurements.

#### 1.3.8 *Quasar monochromatic luminosity*

Computing virial BH masses also requires the quasar luminosity in an emission-line free region of the continuum adjacent to the broad line being used. The luminosity is used as a proxy for the size of the BLR. The monochromatic continuum flux is generally measured at  $1350 \text{ \AA}$  for C iv and  $5100 \text{ \AA}$  for H $\alpha$  and H $\beta$ . The calculation of these luminosities is described in Chapter ??.

As described in Chapter ??, we estimate the uncertainties on the monochromatic luminosities to be  $\sim 0.3 \text{ dex}$ . Given that the luminosity enters into the calculation of BH mass only as the square-root, the uncertainty on the luminosities does not make a large contribution to the uncertainties in the BH mass estimates.

#### 1.3.9 *Characterising the emission-line widths*

There has been a considerable degree of attention paid to the effectiveness of different velocity-width measures of the C iv-emission; specifically, the line FWHM and the dispersion,  $\sigma$ , derived from the second-moment velocity (e.g. Assef et al., 2011; Denney et al., 2013). The FWHM and line dispersion trace different parts of the broad line velocity field, with the FWHM

---

<sup>3</sup> Time differences in the quasar rest-frame are reduced by a factor of  $(1 + z)$ .

relatively more sensitive to any low-velocity core present and the line dispersion relatively more sensitive to the high velocity wings. In practice, the line dispersion is almost certainly a more robust velocity indicator when the assumptions underlying the virial-origin of the emission-line velocity width are true and the spectral S/N and resolution are adequate. This was demonstrated by Denney et al., (2013) for a sample of quasars possessing a significantly smaller range in C iv-blueshift than investigated here.

In reality, however, as highlighted by Denney, (2012), contributions to the C iv emission-line profile from gas where virial motions do not dominate can be significant. Looking to the future, the results of the new reverberation-mapping projects (Shen et al., 2015; King et al., 2015) will show what fraction of the C iv emission-line, as a function of velocity, does reverberate for quasars with an extended range of C iv emission shapes. The derivation of quantitative corrections to transform velocity-width measures from single-epoch to reverberation-only line profiles should then be possible.

As such information is not yet available, there is a strong rationale for investigating whether the systematic changes in the C iv emission-line profile can be used to improve the single-epoch BH mass estimates derived using the C iv line. In Figure 1.7 we show how the C iv FWHM, line dispersion,  $\sigma$ , and line shape, FWHM/ $\sigma$ , vary as a function of the blueshift. The C iv FWHM is correlated with the blueshift, with the median FWHM of quasars with the largest blueshifts a factor of 2-3 higher than quasars with only moderate blueshifts. The dispersion, however, does not show a similarly strong systematic variation.

Without knowledge of the C iv-blueshifts, the dynamic range present in the FWHM and line dispersion measurements accords with the expectations from the study of Denney et al., (2013); the factor of  $\simeq 4$  spread in the FWHM measurements indicating greater sensitivity to the emission-line profile shape than is the case for the dispersion, which varies by a factor of only  $\lesssim 2$ . Adopting a value of  $1200 \text{ km s}^{-1}$  to define ‘low’ and ‘high’ blueshift, the median C iv-emission dispersion for the low and high-blueshift samples differ by only 10 per cent. It follows, therefore, that while the dispersion provides a relatively line-profile independent measure of the velocity width for quasars where the underlying assumption regarding the virial-origin of the velocity width applies, quasars where the assumption

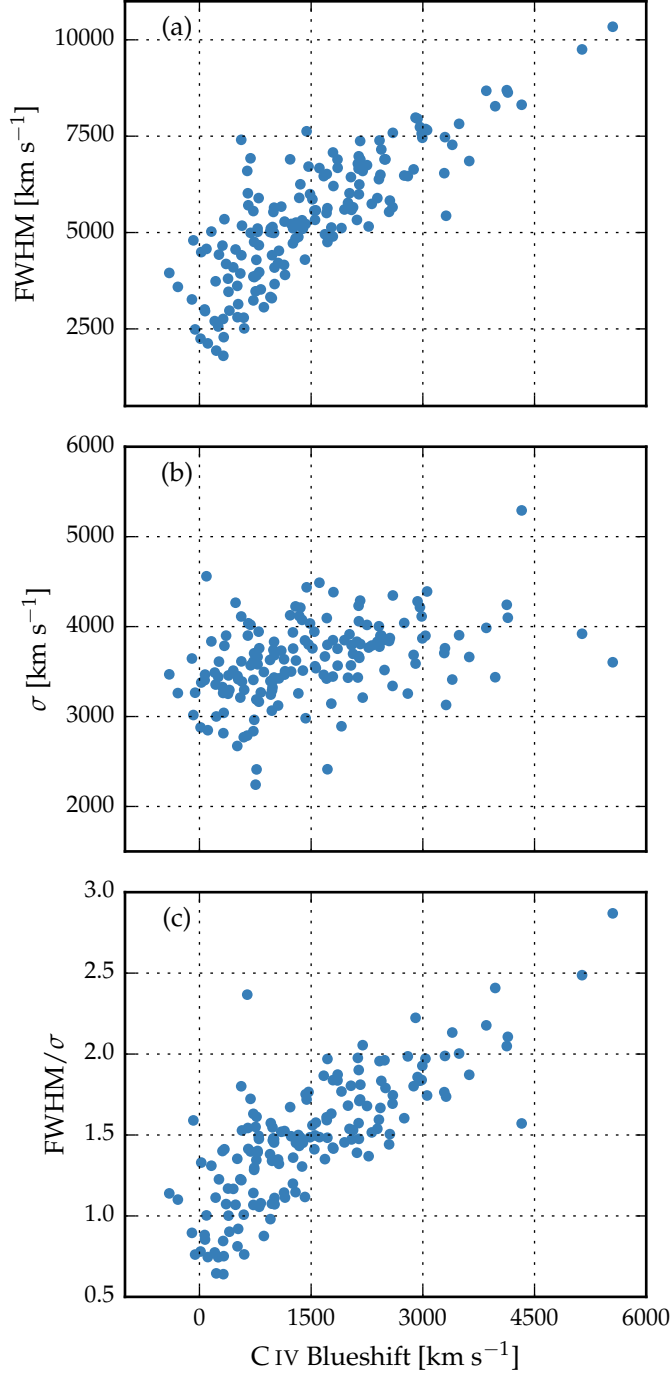


Figure 1.7: The FWHM, dispersion ( $\sigma$ ) and shape ( $\text{FWHM}/\sigma$ ) of C IV as a function of the C IV blueshift.

is not true can be assigned apparently normal velocity-widths and hence potentially incorrect BH masses.

To emphasise this point, in Figure 1.8 we overlay the C iv line profiles of J123611+112922 and J152529+292813, whose dispersions are indistinguishable ( $4168 \pm 271$  and  $4303 \pm 128 \text{ km s}^{-1}$  respectively). Notwithstanding the very similar dispersion values, the emission-line velocity fields differ dramatically and, therefore, the dispersion values cannot be measuring accurately the virial-induced velocity spread of the C iv emission in both quasars.

The analysis here, building on earlier work (including Sulentic et al., 2007; Shen and Liu, 2012), confirms a link between C iv emission-line shape and blueshift, raising the prospect of developing a blueshift-dependent correction to single-epoch BH mass estimates based on the C iv line. Expressed in another way, we are interested in testing if the significant systematic change in line shape as a function of C iv blueshift can be used to provide improved single-epoch BH masses from the C iv emission-line. The tightness of the correlation we observe between the C iv FWHM and blueshift implies that such an approach may be more effective than using the C iv emission-line velocity dispersion without reference to blueshifts. A further practical advantage is that, given the typical S/N of current survey-quality spectra, virial BH mass estimates for high-redshift quasars are usually based on the FWHM rather than the dispersion (e.g. Shen et al., 2011), which, being strongly affected by the continuum placement, is often found to be difficult to measure robustly (e.g. Mejía-Restrepo et al., 2016).

#### 1.4 AN EMPIRICAL CORRECTION TO CIV-BASED VIRIAL BH MASS ESTIMATES

##### 1.4.1 $H\alpha/H\beta$ FWHM comparison

BH mass calibrations which use the width of the broad  $H\beta$  emission-line as a proxy for the virial velocity are widely regarded as the most reliable, since most reverberation mapping employs the  $H\beta$  line and the  $R - L$  relation has been established using  $H\beta$ . When  $H\beta$  is not available,  $H\alpha$  has been shown to be a reliable substitute (e.g. Greene and Ho, 2005; Shen et al., 2011; Shen and Liu, 2012).

In our sample, we have 99 quasars with reliable measurements of both  $H\alpha$  and  $H\beta$  lines. The 99 objects include 21



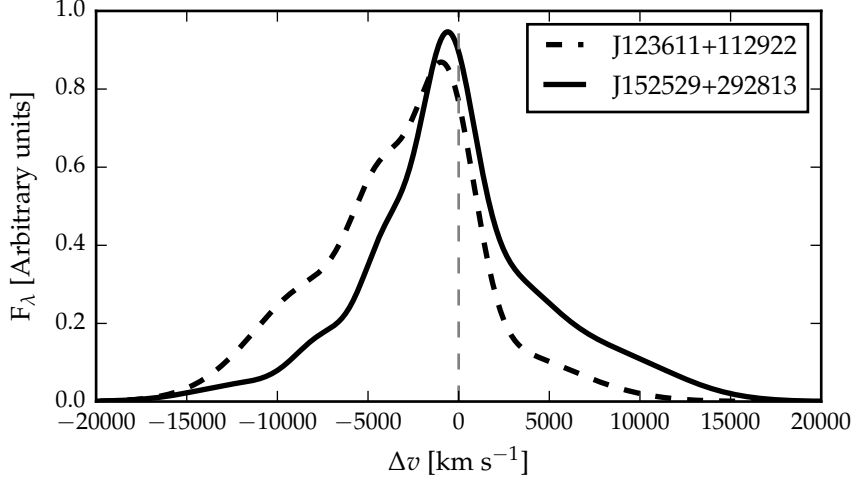


Figure 1.8: Comparison of the C IV line profiles of J123611+112922 and J152529+292813. Notwithstanding the essentially identical dispersion values, the emission-line velocity fields differ dramatically and, therefore, the dispersion values cannot be measuring accurately the virial-induced velocity spread of the C IV emission in both quasars.

quasars which were excluded from the main 308-object catalogue because the C IV FWHM and/or blueshift could not be measured reliably. The line widths are compared in Figure 1.9 and, as expected, a tight correlation is observed. Greene and Ho, (2005), using a sample of 162 quasars with high S/N SDSS spectra at  $z < 0.35$ , established the following relation between the H $\alpha$  and H $\beta$  FWHMs

$$\text{FWHM}(\text{H}\beta) = (1.07 \pm 0.07) \times 10^3 \left( \frac{\text{FWHM}(\text{H}\alpha)}{10^3 \text{ km s}^{-1}} \right)^{(1.03 \pm 0.03)} \quad (1.1)$$

The relation is shown as the dashed line in Figure 1.9. The root-mean-square scatter about this relation is 0.07 dex, compared to the  $\sim 0.1$  dex found by Greene and Ho, (2005). However, we find a systematic offset, in the sense that the H $\beta$  line-widths we measure are on average larger by  $270 \text{ km s}^{-1}$  than predicted by the Greene and Ho, (2005) relation. As our sample covers higher redshifts and luminosities than the sample in Greene and Ho, (2005), we derive a new relation between the H $\alpha$  and H $\beta$  FWHMs.

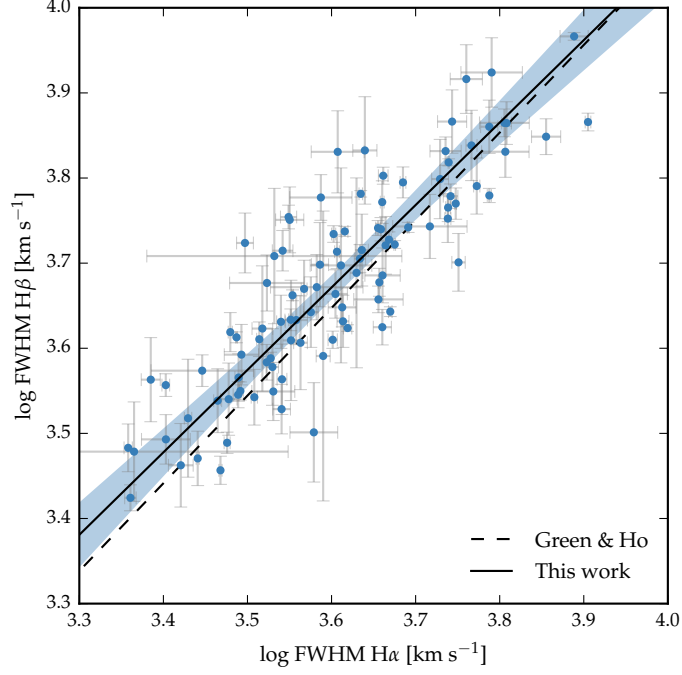


Figure 1.9: Comparison of H $\alpha$  and H $\beta$  FWHM measurements for 99 quasars. The solid line is our best-fitting power-law model, and the blue-shaded region shows the 2- $\sigma$  uncertainties on the model parameters. The dashed line is the relation found by Greene and Ho, (2005) using a sample of  $z < 0.35$  SDSS AGN.

We assume a relation of the same form used by Greene and Ho, (2005), i.e. a simple power-law, and infer the model parameters by fitting a linear model (with slope  $\alpha$  and intercept  $\beta$ ) in log-log space. The fit is performed within a Bayesian framework described by Hogg, Bovy, and Lang, (2010). Each data point is treated as being drawn from a distribution function that is a convolution of the projection of the point's covariance tensor, of variance  $\Sigma_i^2$ , with a Gaussian of variance  $V$  representing the intrinsic variance in the data. The log-likelihood is then given by

$$\ln \mathcal{L} = - \sum_{i=1}^N \frac{1}{2} \ln \left[ 2\pi \left( \Sigma_i^2 + V \right) \right] - \sum_{i=1}^N \frac{\Delta_i^2}{2[\Sigma_i^2 + V]} \quad (1.2)$$

where  $\Delta_i$  is the orthogonal displacement of each data point from the linear relationship. An advantage of this approach is that it allows a proper treatment of the measurement errors on both variables, which in this case are comparably large. The

model also makes the reasonable assumption that there is an intrinsic scatter in the relationship between the variables that is independent of the measurement errors. Following the suggestion by Hogg, Bovy, and Lang, (2010), the linear model was parametrized in terms of  $(\theta, b_{\perp})$ , where  $\theta$  is the angle the line makes with the horizontal axis and  $b_{\perp}$  is the perpendicular distance from the line to the origin. Uniform priors were placed on these parameters, and the Jeffreys prior (the inverse variance) was placed on the intrinsic variance. The posterior distribution was sampled using a Markov Chain Monte Carlo (MCMC) method using the Python package emcee (Foreman-Mackey et al., 2013).

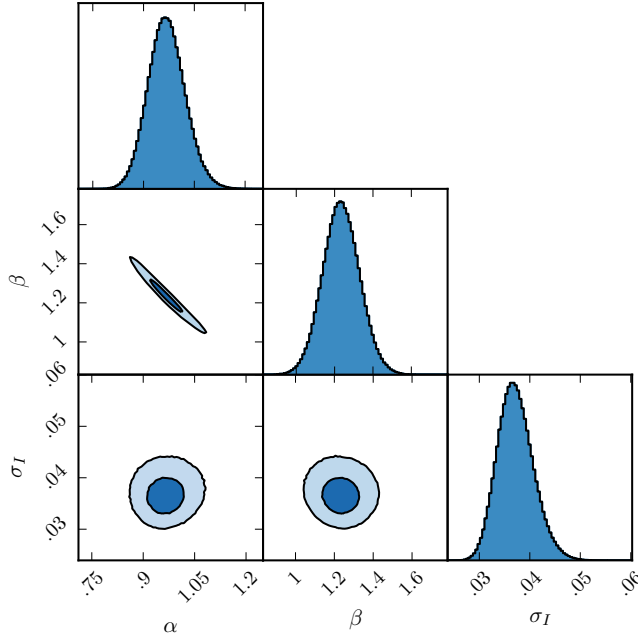


Figure 1.10: One- and two-dimensional projections of the MCMC sampling of the posterior distribution from the fit in Figure 1.9.  $\alpha$  is the power-law index,  $10^{\beta}$  is the normalisation, and  $\sigma_I$  is the intrinsic scatter. In the two-dimensional projections, 1- and 2- $\sigma$  contours are shown.

The one- and two-dimensional posterior distributions are shown in Figure 1.10. The solid line in Figure 1.9 is the maximum likelihood solution

$$\text{FWHM}(\text{H}\beta) = (1.23 \pm 0.10) \times 10^3 \left( \frac{\text{FWHM}(\text{H}\alpha)}{10^3 \text{km s}^{-1}} \right)^{0.97 \pm 0.05} \quad (1.3)$$

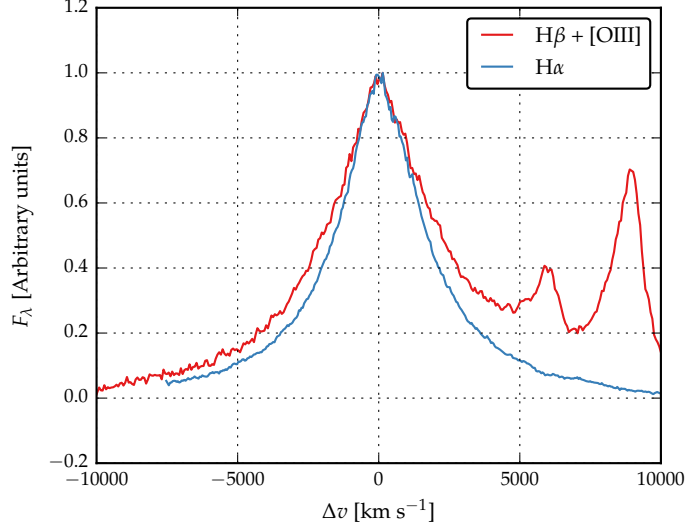


Figure 1.11: The H $\alpha$  (blue) and H $\beta$  (red) emission-line regions in the median composite spectrum, shown as function of the velocity shift from the respective predicted line peak wavelengths. The background continuum and optical Fe II emission has been modelled and subtracted. The line fluxes have been scaled in order for the profile shapes to be readily compared.

and the shaded region shows the  $2\sigma$  uncertainties on the model parameters.

As discussed above, our relation is displaced to slightly higher H $\beta$  FWHM than the Greene and Ho, (2005) relation – the offset is  $210 \text{ km s}^{-1}$  for a quasar with H $\alpha$  FWHM  $4500 \text{ km s}^{-1}$ . We infer a power-law index that, although slightly shallower, is consistent with the Greene and Ho, (2005) index within the quoted uncertainties. The intrinsic scatter in the data,  $\sigma_I$ , we infer from the fit is 0.04 dex. This is smaller than the total scatter seen in Figure 1.9 (0.06 dex), which suggests that measurement errors make a significant contribution to the total scatter in the relation.

For 19 of the 99 quasars with H $\beta$  and H $\alpha$  emission profiles, one of the two Gaussians used to reproduce the H $\beta$  profiles has a FWHM greater than  $20\,000 \text{ km s}^{-1}$  and a fractional contribution to the total H $\beta$  broad line flux greater than 30 per cent (Marziani et al., 2009; Marziani et al., 2013). Such a broad component is not seen in the H $\alpha$  profiles. We constructed composite spectra of the H $\alpha$  and H $\beta$  regions from 217 and 171 quasars respectively. Spectra were first de-redshifted to the quasar rest-

frame, and then interpolated on to a common wavelength grid with a  $1 \text{ \AA}$  resolution. The spectra were scaled by the mean flux in the interval  $4700\text{--}5100 \text{ \AA}$  ( $\text{H}\beta$ ) and  $6400\text{--}6800 \text{ \AA}$  ( $\text{H}\alpha$ ). The  $\text{H}\alpha$  and  $\text{H}\beta$  lines in the median composite spectrum are shown in Figure 1.11. The cores of the two lines are very similar, but  $\text{H}\beta$  has more flux in the wings of the line. This is consistent with what we find in our analysis of the individual spectra.

The very broad  $\text{H}\beta$ -component may be an artifact of the fitting scheme. A particular issue for  $\text{H}\beta$  is the presence of  $\text{Fe II}$  emission, often at a significant level. Furthermore, additional lines could be contributing to the underlying continuum (e.g. the  $\text{He I } \lambda\lambda 4922, 5017$  doublet; Véron, Gonçalves, and Véron-Cetty, 2002; Zamfir et al., 2010). If the  $\text{H}\beta$  FWHM is calculated only from the narrower of the two Gaussian components, then the  $\text{H}\beta$  FWHM decreases by  $630 \text{ km s}^{-1}$  on average, and the new relationship between the  $\text{H}\alpha$  and  $\text{H}\beta$  FWHM is much closer to one-to-one. The  $\text{C IV}$  FWHM relative to the  $\text{H}\alpha/\text{H}\beta$  FWHM will be enhanced by  $\sim 15$  per cent, and so the  $\text{C IV}$ -based BH masses relative to the Balmer-based masses will increase by  $\sim 30$  per cent.

#### 1.4.2 *Measuring the quasar systemic redshift*

An accurate measure of the quasar’s systemic redshift is required in order for the blueshift of the  $\text{C IV}$  emission-line to be determined. Balmer emission centroids are available for all quasars in the catalogue and so we use this to define the systemic redshift.

For 62 and 86 quasars in the  $\text{H}\alpha$  and  $\text{H}\beta$  samples respectively narrow  $[\text{O III}]$  emission is also detected with sufficient S/N to measure the line centroid. In the model fit to the  $\text{H}\beta$  region the velocity centroids of the broad  $\text{H}\beta$ -line and the core component of the  $[\text{O III}]$  emission were deliberately determined separately. We find the intrinsic difference in the velocity centroids of the  $\text{H}\alpha$  and  $\text{H}\beta$  emission and the narrow  $[\text{O III}]$  emission to have a dispersion of  $300$  and  $400 \text{ km s}^{-1}$ , which is very similar to the value found by Shen et al., (2016). However, the median velocity centroid of the narrow component of the  $[\text{O III}]$  emission is blueshifted by  $250 \text{ km s}^{-1}$  relative to the centroid of the broad Balmer line. Applying our parametric model fitting routine to the composite spectrum from Hewett and Wild, (2010), which is constructed using relatively low redshift SDSS quasars with  $L_{\text{Bol}} \sim 10^{44} \text{ erg s}^{-1}$ , the centroids of the broad component of

H $\beta$  and the narrow component of [O III] are found to be at essentially identical velocities, suggesting that the blueshifting of narrow [O III] could be luminosity dependent.

As described in Section 1.3, the broad components of H $\alpha$  and H $\beta$  were modelled with up to two Gaussians, with identical velocity centroids. We also tested models with no constraints on the centroids of the two broad Gaussians, and measured the systemic redshift from the peak of the composite profile. With this set of models, the median difference between the [O III]- and H $\alpha$ (H $\beta$ ) based redshift estimates is reduced to  $-100(-120) \text{ km s}^{-1}$ , with a  $290(320) \text{ km s}^{-1}$  scatter. This suggests that there is a  $\sim 100 \text{ km s}^{-1}$  systematic error in our Balmer-based redshift estimates. Regardless, since both the systematic offset and the scatter are small in comparison to the dynamic range in C IV blueshifts, the blueshift-based empirical correction we will derive does not depend on whether the broad Balmer emission or the [O III] centroid is used to define the systemic redshift, or how the broad Balmer emission is parameterized.

#### 1.4.3 *Balmer/C IV line widths as a function of C IV-blueshift*

In this Section, we directly compare the C IV and H $\alpha$ /H $\beta$  line widths as a function of the C IV blueshift. Because virial BH mass estimates are generally based on the H $\beta$  FWHM, we first convert our H $\alpha$  FWHM measurements to equivalent H $\beta$  FWHM using Equation. 1.3. In Figures 1.12a and 1.13a we show the C IV FWHM relative to both the (H $\beta$ -scaled) H $\alpha$  FWHM and the H $\beta$  FWHM, as a function of the C IV blueshift.

Employing the same Bayesian fitting framework described in Section 1.4.1, we fit independent linear models to the C IV FWHM relative to the H $\alpha$  and H $\beta$  FWHM as a function of the C IV blueshift. As before, our model has an additional parameter representing any intrinsic scatter in the relationship between the variables which is independent of measurement errors. We also tested a model where some fraction of the data points (which is free to vary) are drawn from an outlier distribution, represented by a broad Gaussian centered on the mean of the data. We found, however, that the inferred outlier fraction was very low (0.004, corresponding to  $\sim 0.7$  data points) and so did not include such a component in our model.

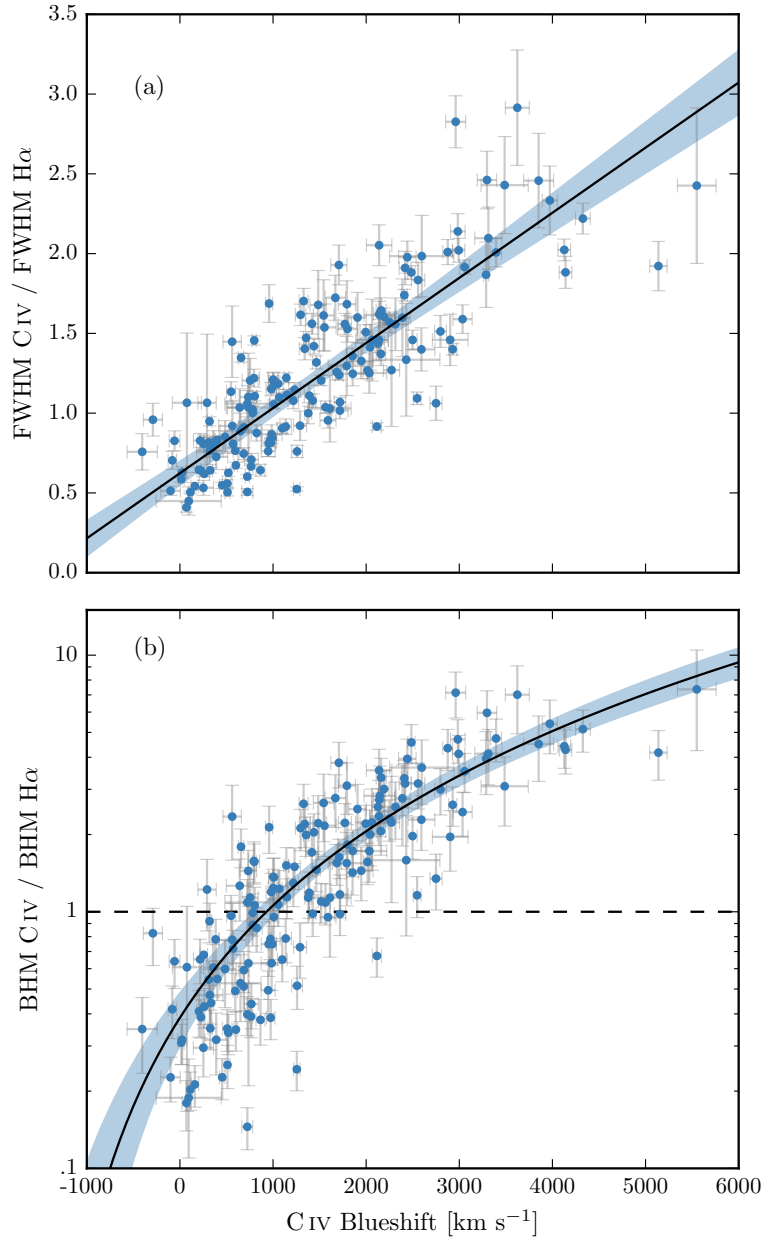


Figure 1.12: C iv FWHM relative to H $\alpha$  FWHM (a), and C iv based BH mass (BHM) compared to H $\alpha$  based mass (b), both as a function of the C iv blueshift. The black line is our best-fit linear model, and the shaded region shows the 2- $\sigma$  uncertainties on the slope and intercept. The H $\alpha$  FWHM have been scaled to match the H $\beta$  FWHM using Equation 1.3.

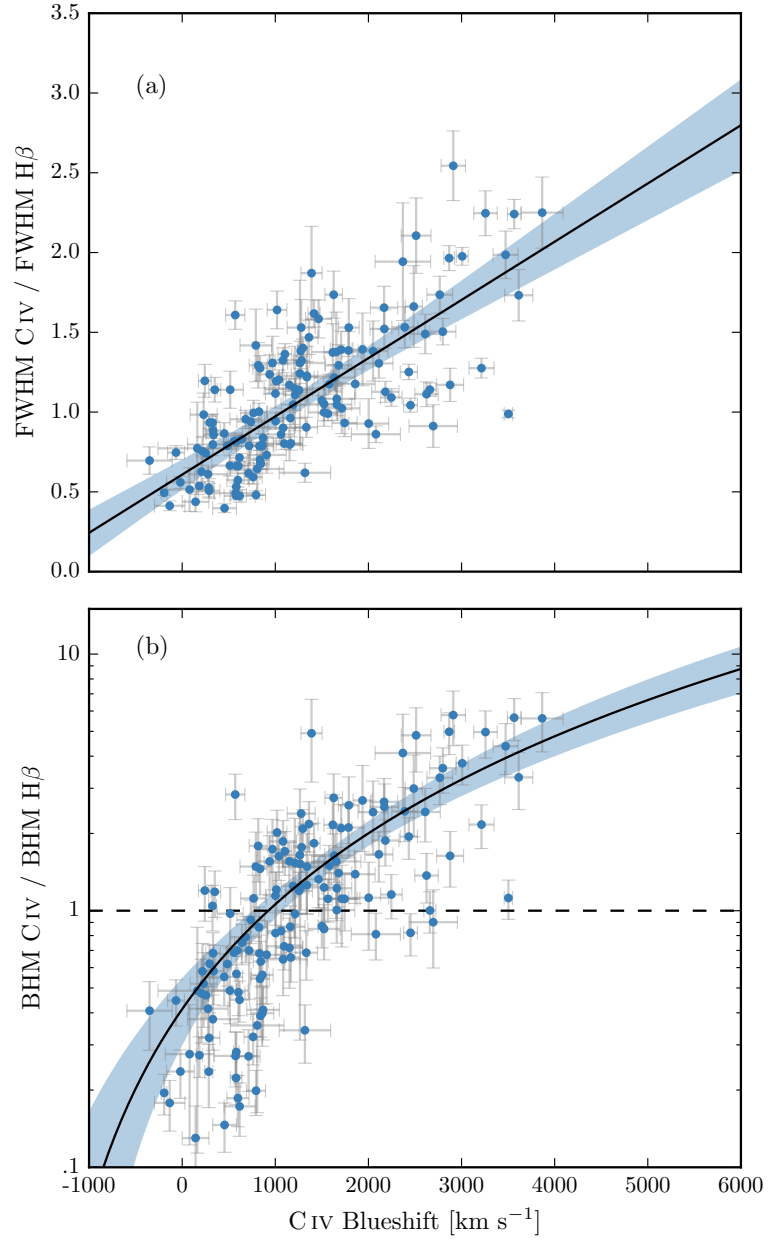


Figure 1.13: C IV FWHM relative to H  $\beta$  FWHM (a), and C IV based BH mass (BHM) compared to H  $\beta$  based mass (b), both as a function of the C IV blueshift.



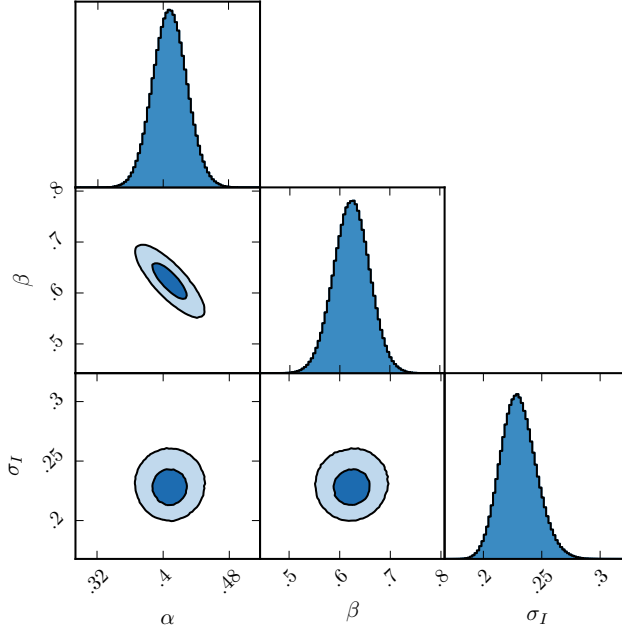


Figure 1.14: One- and two-dimensional projections of the MCMC sample of the posterior distribution for a linear fit to the FWHM C iv/H $\alpha$  ratio as a function of the C iv blueshift. In the two-dimensional projections we show 1- and 2- $\sigma$  contours. The posterior distribution for the linear fit to the FWHM C iv/H $\beta$  ratio, which we do not show, has a very similar appearance.

In Figure 1.14 we show the one- and two-dimensional projections of the posterior distribution from the linear fit to the FWHM C iv/H $\alpha$  ratio. The projections from the FWHM C iv/H $\beta$  fit, which we do not show, have very similar appearances. In Figure 1.12a we plot the maximum likelihood model and the 2 $\sigma$  uncertainties on the model parameters. The maximum likelihood line is given by

$$\text{FWHM}(\text{C iv, Corr.}) = \frac{\text{FWHM}(\text{C iv, Meas.})}{(0.41 \pm 0.02) \left( \frac{\text{C iv Blueshift}}{10^3 \text{ km s}^{-1}} \right) + (0.62 \pm 0.04)} \quad (1.4)$$

for the C iv/H $\alpha$  fit and

$$\text{FWHM}(\text{C iv, Corr.}) = \frac{\text{FWHM}(\text{C iv, Meas.})}{(0.36 \pm 0.03) \left( \frac{\text{C iv Blueshift}}{10^3 \text{ km s}^{-1}} \right) + (0.61 \pm 0.04)}$$

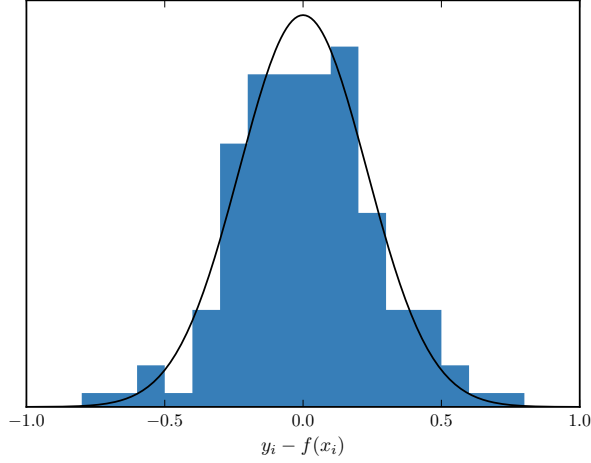


Figure 1.15: The distribution of the orthogonal displacement of each data point from the best-fitting linear relationship in the fit to  $\text{FWHM}(\text{C IV})/\text{FWHM}(\text{H}\alpha)$  as a function of the C IV blueshift (blue histogram). The black curve is a Normal distribution with a width equal to the intrinsic scatter in the population inferred from the fit. The two distributions are well-matched, which demonstrates that our model is a good representation of the data and the measurement errors on the data points are small relative to the intrinsic scatter.

(1.5)

for the C IV/H $\beta$  fit. The intercepts of the two relations are consistent, while the difference between the slopes is only marginally inconsistent given the quoted uncertainties.

The intrinsic scatter in the data about the linear relation we infer is  $0.23 \pm 0.02$  and  $0.25 \pm 0.02$  for the H $\alpha$  and H $\beta$  fits respectively. The intrinsic scatter for the H $\alpha$  fit is represented by the Normal probability density distribution shown in Figure 1.15. In the same Figure, we show the distribution of the orthogonal displacement of each data point from the best-fitting linear relationship. The two distributions are well-matched, which demonstrates that our model is a good representation of the data and the measurement errors on the data points are small relative to the intrinsic scatter.

The overall (intrinsic and measurement) scatter about the best-fitting model is slightly higher when the C IV line-widths are compared to H $\beta$  (0.12 dex) than when compared to H $\alpha$  (0.10 dex). This is likely due, at least in part, to the generally higher S/N of the H $\alpha$  emission. In addition, contributions

from the strong [O III] doublet in the vicinity of H $\beta$  make deblending the H $\beta$  emission more uncertain. As a consequence, for quasars where H $\alpha$  and H $\beta$  are both measured, the mean uncertainty on the H $\alpha$  FWHM is  $130 \text{ km s}^{-1}$ , compared to  $340 \text{ km s}^{-1}$  for H $\beta$ .

In the next Section, we use both the H $\alpha$  and H $\beta$  lines to calculate unbiased BH masses. However, we use the H $\alpha$  measurements to derive an empirical C IV blueshift based correction to the C IV masses (Equation 1.6) because of the issues related to the accurate modelling of the H $\beta$ -profile just described. An extra advantage, which is evident in Figures 1.12a and 1.13a, is that the H $\alpha$  sample has a better C IV blueshift coverage. However, as can be seen from the similarity of Equations 1.4 and 1.5, our results would not change significantly were we instead to use the H $\beta$  sample.

#### 1.4.4 C IV-based virial BH mass estimates

Virial BH masses were calculated using the widely adopted Vestergaard and Peterson, (2006) calibrations. The Vestergaard and Peterson, (2006) C IV FWHM calibration uses the monochromatic continuum luminosity at  $1350 \text{ \AA}$  to predict the BLR radius and corresponds to ( $a = 6.66$ ,  $b = 2$ ,  $c = 0.53$ ) in Equation ???. For the H $\beta$  calibration, Vestergaard and Peterson, (2006) use the monochromatic continuum luminosity at  $5100 \text{ \AA}$  and calibration coefficients corresponding to ( $a = 6.91$ ,  $b = 2$ ,  $c = 0.5$ ). BH masses are computed using the line and continuum properties given in Table 1.4, and we convert our H $\alpha$  emission-line velocity-width measures to predicted H $\beta$  widths using Equation 1.3.

In Figures 1.12b and 1.13b the C IV-based estimates are compared to the H $\alpha$ /H $\beta$  estimates as a function of the C IV blueshift. There is a strong systematic error in the C IV-based masses as a function of blueshift, which is a direct consequence of the FWHM trend described in the previous Section. The C IV emission-based BH masses are in error by a factor of more than five at  $3000 \text{ km s}^{-1}$  in C IV emission blueshift and the overestimate of the BH masses reaches a factor of 10 for quasars exhibiting the most extreme blueshifts,  $\gtrsim 5000 \text{ km s}^{-1}$ .

The virial product is the product of the virial velocity squared and the BLR radius (e.g. Shen, 2013), and is proportional to the BH mass. We use the corrected C IV FWHM given by Equation 1.4 as an indicator of the virial velocity, and adopt the same

R – L relation for the 1350 Å continuum luminosity as Vestergaard and Peterson, (2006) (i.e.  $R \propto L^{0.53}$ ). To find the constant scaling factor necessary to transform the virial product into a BH mass we compute the inverse-variance weighted mean difference between the virial products and the H $\alpha$ -based masses. The virial BH mass can then be expressed in terms of the corrected C IV FWHM and monochromatic continuum luminosity at 1350 Å

$$\text{MBH}(\text{C IV, Corr.}) = 10^{6.71} \left( \frac{\text{FWHM}(\text{C IV, Corr.})}{10^3 \text{ km s}^{-1}} \right)^2 \left( \frac{\lambda L_{\lambda}(1350 \text{ Å})}{10^{44} \text{ erg s}^{-1}} \right)^{0.53} \quad (1.6)$$

Given measured C IV emission-line FWHM and blueshift, Equations 1.4 and 1.6 can then be used to provide an unbiased estimate of the quasar BH mass.

#### 1.4.5 C IV-derived BH masses at low C IV blueshift

Reverberation mapping measurements of nearby AGN have revealed the BLR to be stratified, with high-ionisation lines, including C IV, emitted closer to the BH than low-ionisation lines, including H $\alpha$  and H $\beta$  (e.g. Onken and Peterson, 2002). Vestergaard and Peterson, (2006) found that the C IV-emitting region is at approximately half the radius of the H $\beta$ /H $\alpha$  emitting region. Given the  $\Delta V \propto R_{\text{BLR}}^{-0.5}$  virial relation, this leads to the prediction that the C IV line widths should be  $\simeq 1.4$  times broader than H $\alpha$  for a given BH mass, and this expectation is reflected in the Vestergaard and Peterson, (2006) BH mass calibrations.

In our sample, the median C IV/H $\alpha$  FWHM ratio is  $0.97 \pm 0.31$  for the 77 quasars with C IV blueshifts  $< 1200 \text{ km s}^{-1}$ . This is significantly smaller than the predicted value of 1.4. As a direct consequence of the empirically small C IV/H $\alpha$  FWHM ratio, the C IV-derived BH mass estimates are systematically lower than the corresponding H $\alpha$ -derived masses when the blueshift is small. This can be seen in Figure 1.12b, where for almost every quasar with a C IV blueshift  $< 1200 \text{ km s}^{-1}$ , the C IV-derived BH mass is smaller than the corresponding H $\alpha$ -derived mass.

Denney, (2012) used multiple-epoch spectra from reverberation mapping campaigns to isolate the part of the C IV profile which is varying. In some profiles they identified a non-varying core component, possibly originating from gas at larger radii.

In single-epoch spectra, both parts of the line - varying and non-varying - are measured. The gas at larger radii enhances the profile at lower-velocity and leads to smaller FWHM or dispersion values. This picture is therefore consistent with the low C IV/H $\alpha$  FWHM ratio observed in our sample of quasars with modest C IV blueshifts.

## 1.5 PRACTICAL APPLICATION OF THE C IV-BASED CORRECTION TO VIRIAL BH MASS ESTIMATES

### 1.5.1 *Recipe for unbiased C IV based BH masses*

#### 1.5.1.1 *Measuring the systemic redshift*

Equations 1.4 and 1.6 together provide an un-biased estimate of the virial BH mass given the FWHM and blueshift of C IV, together with the continuum luminosity at 1350 Å. The FWHM is readily obtained, either directly from the data, or, via the fitting of a parametric model to the C IV emission-line. The blueshift – defined as the bisector of the cumulative line flux – is also straightforward to measure and our preferred procedure is described in Section ???. The only potential complication arises in establishing the quasar systemic redshift and hence defining the zero-point for the C IV-blueshift measurement, since both the blueshift and the systemic redshift cannot be determined from C IV alone. In practice, when rest-frame optical lines are accessible, as is the case for the quasar sample here, an accurate systemic redshift can be obtained. The [O III] doublet and the Balmer lines all have velocity centroids very close to systemic, and the same is true for the broad Mg II doublet. For quasars at very high redshifts,  $z \sim 6$ , systemic redshifts can also be derived using the [C II] 158  $\mu\text{m}$  emission in the sub-millimetre band (e.g. Venemans et al., 2016). However, in general, for example in determining the BH masses of quasars at redshifts  $z > 2$ , if only the rest-frame ultraviolet region is available determining a reliable systemic redshift is non-trivial.

The SDSS DR7 pipeline redshifts are not sufficiently reliable to measure the C IV blueshift accurately because, in part, the C IV emission-line itself contributes to the determination of the quasar redshifts. This is demonstrated in Figure 1.16a, in which we plot the C IV-blueshift versus C IV-emission EQW using the SDSS pipeline redshifts and the blueshifts calculated by Shen et al., (2011). A strong trend in the blueshift values as a function

Table 1.3: The fractional error on the corrected BH mass as a function of C IV blueshift for different uncertainties in the quasar systemic redshift.

$\delta v$ (km s <sup>-1</sup> )	C IV blueshift (km s <sup>-1</sup> )			
	0	1000	2000	4000
250	0.33	0.20	0.14	0.09
500	0.65	0.39	0.28	0.18
1000	1.30	0.79	0.57	0.36

of line EQW is not evident in Figure 1.16a; structure in the parameter space is being masked because the C IV emission-line is itself being used in the determination of the quasar redshifts.

The redshift-determination scheme of Hewett and Wild, (2010) provided much improved redshifts, not least because the redshift estimates for the majority of quasars were derived using emission-lines other than the C IV-line itself. Figure 1.16b shows SDSS DR7 quasars in the same C IV parameter space as Figure 1.16a, but now using Hewett and Wild, (2010) redshifts. The improved redshift estimates are predominantly responsible for the differences seen in Figure. 1.16a and 1.16b; the appearance in Figure 1.16b of the extension to high blueshift for quasars with low C IV EQW is particularly evident.

Shen et al., (2016) and our own work shows that there is an intrinsic variation of  $\sigma \simeq 220$  km s<sup>-1</sup> in the velocity centroids of the BLR emission-lines relative to a systemic-frame defined by emission-lines in the quasar NLR. The redshifts for quasars in the SDSS DR10 and DR12 catalogues (Pâris et al., 2014; Pâris et al., 2017) possess errors of  $\simeq 500 - 750$  km s<sup>-1</sup> (Pâris et al., 2012; Font-Ribera et al., 2013). The impact of low spectrum S/N for fainter quasars in all the SDSS data releases increases the uncertainty further. Table 1.3 includes the values for the fractional error in the corrected BH mass that result from a given error in the determination of the systemic rest-frame. For example, the fractional error in the corrected BH mass is 0.39 for a quasar with a 1000 km s<sup>-1</sup> C IV blueshift when there is a 500 km s<sup>-1</sup> uncertainty in the quasar systemic redshift.

Of potentially more significance for studies of BH masses as a function of quasar and host-galaxy properties are redshift errors that depend on the form of the quasar ultraviolet SED. The large systematic variation in the C IV emission-line profile within the population is evident from figures 11 and 12 of Richards et al., (2011). The plots and analysis in Richards et

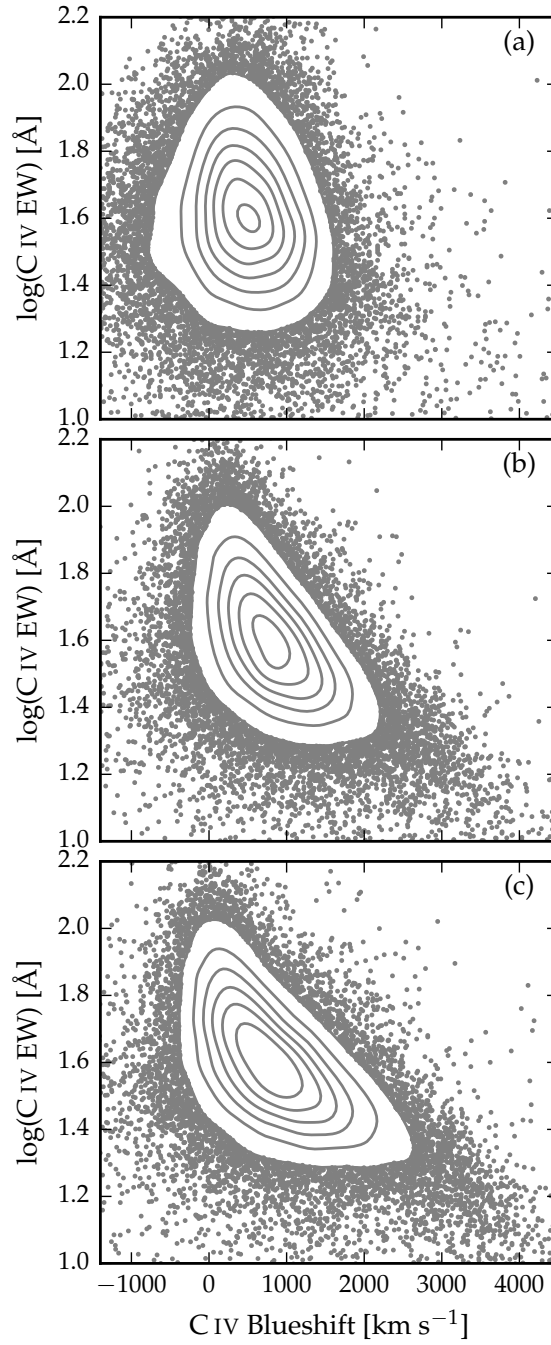


Figure 1.16: Rest-frame EQW versus blueshift of the broad C IV emission-line for 32,157 SDSS DR7 quasars at  $1.6 < z < 3.0$ . Panel (a) uses C IV line parameters from Shen et al., (2011) and SDSS pipeline systemic redshifts. Panels (b) and (c) use systemic redshifts from Hewett and Wild, (2010) and Allen & Hewett (2017, in preparation) respectively, and C IV line measurements described in Section 1.5.1.2.



al., (2011) employ the quasar redshifts from Hewett and Wild, (2010) but, as is evident from the figures, the systematic variation in the C iv shape is correlated with changes in the quasar SEDs, including the strengths of the Si III  $\lambda$ 1892 and C III  $\lambda$ 1908 emission-lines in the rest-frame ultraviolet. As a consequence, the redshifts from Hewett and Wild, (2010) still suffer from systematic errors that are correlated with the shape, and particularly the blueshift, of the C iv emission-line. For the Hewett and Wild, (2010) redshifts, and ultraviolet emission-line based redshifts in general, quasars with large C iv EQW and modest blueshifts have relatively small ( $\simeq 300 \text{ km s}^{-1}$ ) SED-dependent redshift errors. Redshift uncertainties as large as  $\simeq 1000 \text{ km s}^{-1}$  for such quasars are unusual and the large relative error in the corrected C iv BH mass given in Table 1.3 is pessimistic.

Conversely, systematic redshift errors are greatest for quasars with large blueshifts, reaching  $\sim 750 \text{ km s}^{-1}$  in the extreme for the Hewett and Wild, (2010) values. The associated error in the corrected C iv BH masses is, however, mitigated somewhat due to the smaller gradient of the  $\text{MBH}(\text{C iv})/\text{MBH}(\text{Balmer})$  relation at large C iv blueshift (see Figures 1.12b and 1.13b). A definitive quantification of any systematic SED-dependent errors present in the quasar redshifts contained in the SDSS DR12 catalogue is not yet available but the principal component analysis (PCA) based redshift estimates are expected to be largely free of SED-dependent systematics.

Using published redshift estimates, notably those from Hewett and Wild, (2010) for the SDSS DR7 quasars and the BOSS PCA-based redshifts from Pâris et al., (2017) for SDSS DR12, the correction formula given in Section 1.4.3 produces significant improvements to C iv-based BH mass estimates. In a forthcoming work, Allen & Hewett (2017, in preparation) will present a new redshift-estimation algorithm that produces redshifts independent of the C iv blueshift and other variations in the ultraviolet SEDs of luminous quasars. The low-ionization emission-lines visible in the rest-frame ultraviolet (over wavelengths from Mg II  $\lambda$ 2796,2803 down to the O I  $\lambda$ 1304+Si II  $\lambda$ 1307 blend) using the new redshift-algorithm are located at rest-frame wavelengths in excellent agreement with the systemic redshift defined using the rest-frame narrow-line optical [O III] doublet and broad-line H $\beta$  and H $\alpha$ . SED-dependent systematic errors are below the apparent inherent dispersion of  $\simeq 220 \text{ km s}^{-1}$  associated with broad emission-line redshifts (Shen et al., 2016).



Figure 1.16c shows the C IV emission-line parameters calculated using the Allen & Hewett redshift-estimation algorithm. The systematic trends seen in Figure 1.16b, in particular the extension to high blueshift at low C IV EQW, become more apparent in Figure 1.16c, as expected from consideration of the known SED-related errors in the redshifts from Hewett and Wild, (2010). A population of quasars with only modest blueshifts and low EQW is also apparently still present.

#### 1.5.1.2 C IV emission-line blueshift measurements

The differences in the distribution of C IV emission-line properties seen in the three panels of Figure 1.16 are due primarily to the change in the systemic redshift estimates. It is also necessary, however, to obtain a measure of the C IV emission-line ‘location’ in order to calculate the blueshifts. When working with moderately-sized samples, parametric fits to the emission-line profile may be undertaken using careful mask-definition to minimise the effect of absorption features on the profiles used for the parametrization, and this is the approach we followed in Section 1.3.

Effective analysis of the tens of thousands of spectra from SDSS DR7, and now DR12, however, requires a more robust scheme to determine a C IV-blueshift estimate that is not very sensitive to the range of S/N among the spectra or the presence of narrow absorption systems within the C IV-emission profile. Shen et al., (2011) provide a discussion (their section 3) of the factors that effect the measurement of broad emission-lines in quasar spectra of modest S/N. Their careful analysis of the C IV emission properties employed the results of parametric fits of three Gaussians to the spectra. Our own experiments in quantifying the C IV emission properties of SDSS spectra showed that a simple non-parametric measure of the C IV emission location reduced the number of outliers significantly. Visual inspection of spectra demonstrated that the improvement is due primarily to the identification of, and interpolation over, associated and outflow absorption systems, which forms part of the non-parametric measurement scheme.

We therefore chose to use a non-parametric scheme to measure the blueshift of the C IV line, which we will now describe. A continuum is first defined as a power-law of wavelength,

$f(\lambda) \propto \lambda^{-\alpha}$ , with the slope,  $\alpha$ , determined using the median<sup>4</sup> values of the flux in two continuum windows at 1445-1465 and 1700-1705 Å (the same wavelengths as adopted by Shen et al. 2011). The C IV emission-line is taken to lie within the wavelength interval 1500-1600 Å, a recipe that is commonly adopted (e.g. Shen et al., 2011; Denney et al., 2013). To reduce the impact of narrow absorption systems on the emission-line profile a ‘pseudo continuum’ is defined by applying a 41-pixel median filter to the quasar spectrum. Pixels within the C IV profile that lie more than  $2\sigma$  below the pseudo-continuum are deemed to be affected by absorption and added to an ‘absorber’-mask. Two pixels on either side of each such pixel are also included in the mask. For each masked pixel, the flux values in the spectrum are replaced by values from the pseudo-continuum. The wavelength that bisects the cumulative total line flux is recorded and the blueshift is defined in exactly the same way as in Section 1.3.

#### 1.5.2 Systematic trends in residuals

The scatter about the best-fitting line in the C IV/H $\alpha$  FWHM versus C IV-blueshift relation is  $\sim 0.1$  dex, an order of magnitude smaller than the size of the C IV-blueshift dependent systematic but, nevertheless, still significant. With a view to reducing the scatter further, we searched for measurable parameters which correlate with the scatter at fixed C IV blueshift, including the luminosity, redshift, [O III] EQW, and Fe II EQW. The only significant correlation we find is with the H $\alpha$  FWHM (Figure 1.17). Quasars with broad H $\alpha$  lines tend to lie below the relation while quasars with narrow H $\alpha$  tend to lie above it. One possibility is that this correlation is simply due to random scatter (either intrinsic or measurement error) in the H $\alpha$  FWHM which, with the other quasar properties fixed, would naturally produce a correlation between FWHM(C IV)/FWHM(H $\alpha$ ) and FWHM(H $\alpha$ ). However, the fact that we see no such correlation between the model residuals and the C IV FWHM suggests that the H $\alpha$  FWHM correlation could be revealing something more fundamental. The H $\alpha$ /H $\beta$  FWHM is part of ‘eigenvector 1’ (EV1), the first eigenvector in a principal component analysis which originated from the work of Boroson and Green, (1992). Figure 1.17 suggests that part of the scatter between the Balmer and C IV velocity widths might be attributed to differences in

<sup>4</sup> The median is used to improve the robustness of the continuum estimate from the relatively small wavelength intervals.

the spectral properties which are correlated with EV1 (Marziani et al., 2013).

The shape of an emission line can be characterised by the ratio  $\text{FWHM}/\sigma$ .  $\text{FWHM}/\sigma \simeq 2.35$  for a Gaussian profile, while  $\text{FWHM}/\sigma \simeq 1$  for a peakier Lorentzian profile<sup>5</sup>. In our sample, we find the model residuals and the  $\text{H}\alpha$  FWHM correlate with the shape of the line. The narrow lines are, on average, ‘peakier’ (with  $\text{FWHM}/\sigma \simeq 1$ ) than the broader lines (with  $\text{FWHM}/\sigma \simeq 2$ ). This suggests that the BLR structure (e.g. the balance between rotation, turbulence and radial motions) is changing with the emission-line FWHM (e.g. Collin et al., 2006; Kollatschny and Zetzl, 2011; Kollatschny and Zetzl, 2013). This raises the question of whether the Balmer-line FWHM is a reliable proxy for the virial-induced velocity dispersion for the full range of Balmer line shapes we have in our sample.

When calibrating the virial-product to masses derived independently using the  $M_{\text{BH}} - \sigma$  relation, Collin et al., (2006) find that the scaling factor,  $f$ , is a factor  $\sim 2$  larger for their Population ‘1’ sources (with  $\text{FWHM}/\sigma < 2.35$  and essentially equivalent to population A of Sulentic et al. 2000) than for their Population 2 (with  $\text{FWHM}/\sigma > 2.35$ ). For single-epoch BH mass estimates, assuming a constant value of  $f$ , as is normally done (e.g. Vestergaard and Peterson, 2006), means that Population 1 masses will be underestimated and Population 2 will be overestimated. This result could account for some of the remaining scatter between the C IV- and Balmer-based BH masses (Figure 1.17).

Shen and Ho, (2014) argue that a large part of the scatter observed in the  $\text{H}\beta$  FWHM relates not to a spread in BH masses, but rather to the orientation of the BLR relative to the line-of-sight of the observer. This would be the case if the BLR had a flattened disc-like geometry. In this case, the observed line width would increase with the inclination of the disc relative to the line of sight. At radio wavelengths, the morphology of the radio structure, parametrized in terms of ‘core dominance’ is believed, at least in a statistical sense, to be a proxy for the orientation of the accretion disk (e.g. Jackson and Browne, 1991). Twenty core-dominated quasars and six lobe-dominated quasars were identified in our sample, but no statistically sig-

<sup>5</sup> Strictly  $\text{FWHM}/\sigma \rightarrow 0$  for a Lorentzian profile, but values close to unity are typical when the dispersion is calculated over a velocity range,  $\simeq \pm 10\,000 \text{ km s}^{-1}$ , used to parametrize broad emission-lines in quasar spectra.

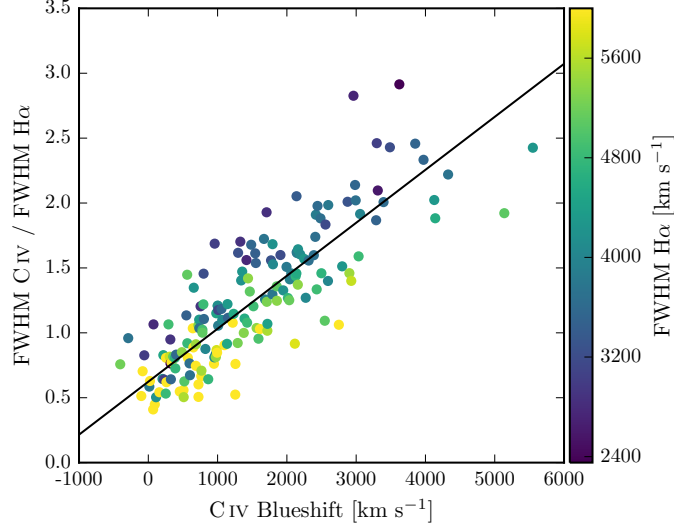


Figure 1.17: Same as Figure 1.12a, with the marker colour representing the  $H\alpha$  FWHM. At fixed C IV blueshift, there is a clear  $H\alpha$  FWHM dependent systematic in the model residuals.

nificant differences in the  $H\alpha$  line-widths of the two samples were found. It should be noted that the sub-sample of radio-detected quasars is small and the effectiveness of the test is further compromised by the lack of radio-detected quasars at large blueshifts (see figure 14 of Richards et al., 2011, for example).

As mentioned in Section ?? and discussed in Richards et al., (2011), quasars with current reverberation mapping measurements have a restricted range of C IV-line shapes. In particular, there are currently very few reverberation-mapping measurements of quasars with large C IV blueshifts. Looking to the future, the results of the large on-going statistical reverberation mapping projects (e.g. Shen et al., 2015; King et al., 2015) for luminous quasars at high-redshift will shed new light on the Balmer line emitting region of the BLR for quasars with a range of C IV blueshifts and lead to a greater understanding of the relation between the Balmer line profile and the BH mass.

### 1.5.3 Effectiveness of the C IV blueshift based correction to BH masses

Figure 1.18 demonstrates that our sample has an excellent coverage of the C IV EQW-blueshift parameter space in relation to SDSS DR7 quasars at redshifts  $1.6 < z < 3.0$ . The system-

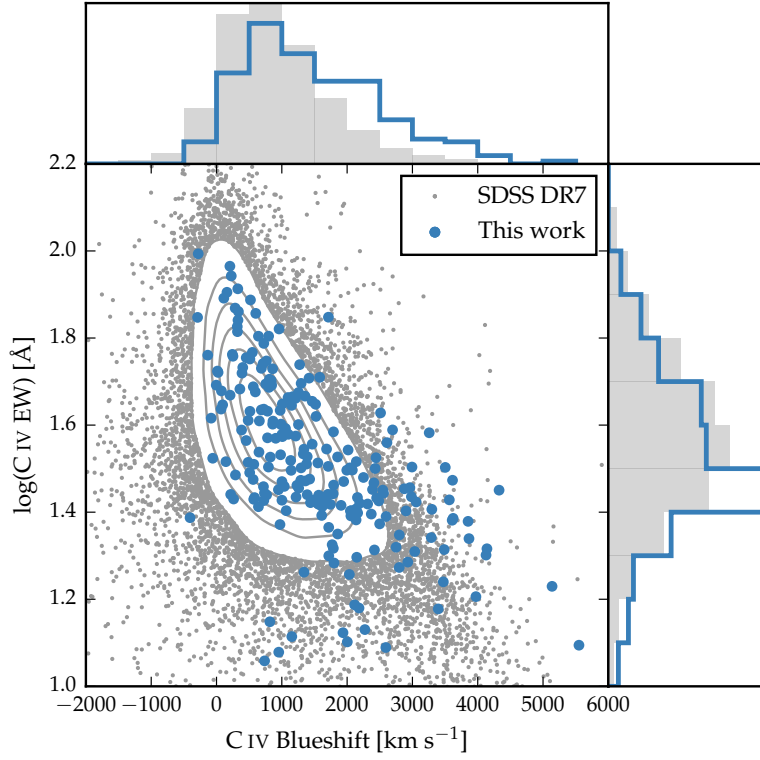


Figure 1.18: Rest-frame EQW versus blueshift of the broad C IV emission-line for SDSS DR7 quasars and our sample (see Figure 1.16b for details). Our sample has very good coverage of the parameter space; the shift to high blueshifts is a result of the high luminosity of our sample in relation to the SDSS sample and the correlation between luminosity and blueshift.

atic offset to higher C iv blueshifts for our catalogue relative to the SDSS quasars as a whole is a result of the higher mean luminosity relative to the SDSS sample (Figure ??). Our sample includes 21 quasars with C iv blueshifts  $> 3000 \text{ km s}^{-1}$ , and extends to  $\sim 5000 \text{ km s}^{-1}$ , i.e. at the very extreme of what is observed in this redshift and luminosity range. Our investigation thus demonstrates that the C iv-blueshift based correction derived in this Chapter is applicable to very high blueshifts. Conversely, there are no quasars in our catalogue with C iv blueshifts  $\lesssim 0 \text{ km s}^{-1}$  and we caution against extrapolating the correction formula to negative blueshifts.

Figure 1.19 compares the C iv- and H $\alpha$ -based BH masses before and after applying the blueshift-based correction to the C iv FWHM. Before the correction, the correlation between the C iv- and H $\alpha$ -based BH masses is very weak, and the scatter between the masses is 0.4 dex. After correcting the C iv FWHM for the non-virial contribution, the correlation improves dramatically. The scatter between the corrected C iv-based masses and the H $\alpha$ -based masses is reduced to 0.2 dex. The scatter is 0.24 dex at low C iv blueshifts ( $\sim 0 \text{ km s}^{-1}$ ) and 0.10 dex at high blueshifts ( $\sim 3000 \text{ km s}^{-1}$ ).

There has been a considerable amount of attention regarding the relative merits of using the FWHM or dispersion to characterise the velocity width (e.g. Denney et al., 2013). The existence of a trend in the C iv-dispersion values with C iv blueshift is evident from inspection of Figure 1.7b but the systematic trend relative to the spread at fixed blueshift is significantly smaller than when using C iv FWHM. Therefore, without the blueshift information, using the line dispersion would yield a more accurate BH mass than the FWHM (Figure 1.20).

The correlation between the H $\alpha$  and C iv line dispersion is, however, weak. The Pearson coefficient for the correlation is 0.36 (and just 0.15 when the H $\beta$  measurements are used in place of H $\alpha$ ). Furthermore, there is little dynamic range in the line dispersion: the scatter is just 480 and 460  $\text{km s}^{-1}$  for H $\alpha$  and C iv respectively. The observation suggests that the line dispersion does not fully trace the dynamic range in BH mass present in the quasar population. At least part of the reason is that the line dispersion is difficult to measure reliably in current survey-quality data, particularly because of the sensitivity to flux ascribed to the wings of the emission-line (e.g. Mejía-Restrepo et al., 2016). Figures 1.19 and 1.20 demonstrate that estimating a reliable BH mass from the C iv FWHM and blueshift line is

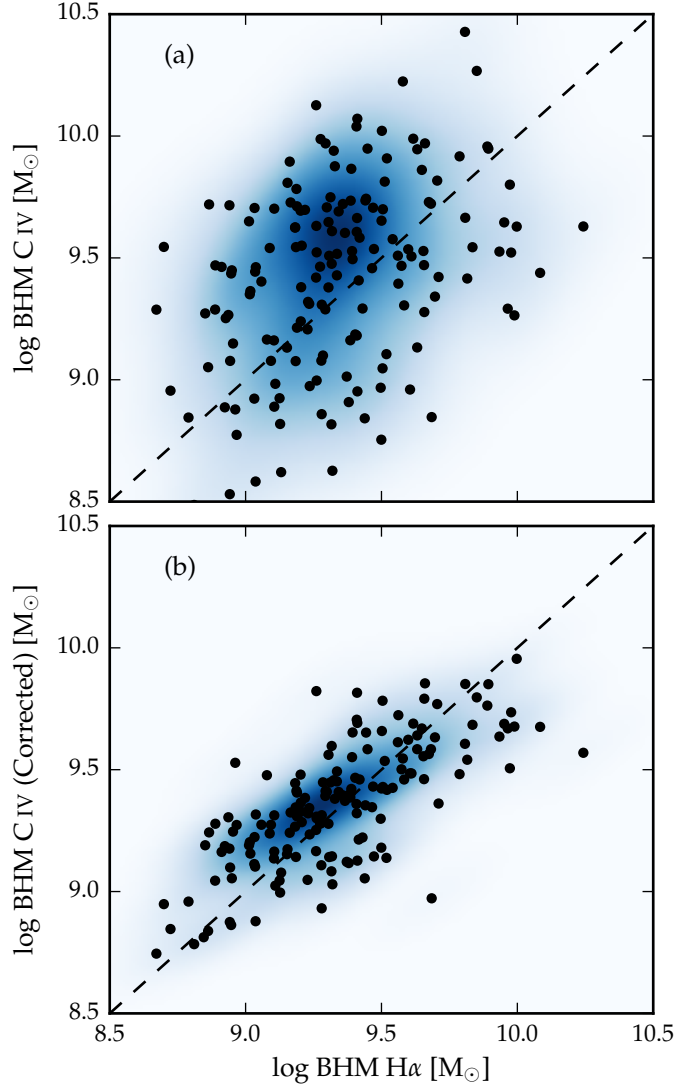


Figure 1.19: Comparison of the C IV- and H $\alpha$ -based BH masses before (a) and after (b) applying the C IV blueshift-based correction to the C IV FWHM. The density of the plotted points (estimated using a Gaussian kernel density estimator) is represented by the colour. The correction to the C IV BH masses decreases the scatter by from 0.4 to 0.2 dex.

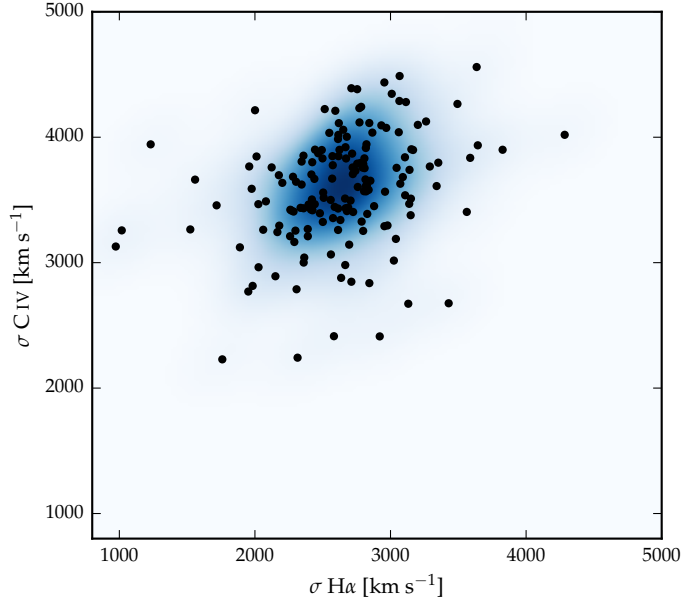


Figure 1.20: Comparison of the C IV and H $\alpha$  line dispersion,  $\sigma$ . The density of the plotted points (estimated using a Gaussian kernel density estimator) is represented by the colour. Estimating a reliable BH mass from the C IV FWHM and blueshift line is substantially more effective than using the C IV line dispersion with, or without, the line blueshift. The C IV dispersion values are larger than the corresponding H $\alpha$  measurements by a factor of 1.4 on average, which is consistent with reverberation mapping measurements (Vestergaard and Peterson, 2006).

substantially more effective than using the C IV line dispersion with, or without, the line blueshift.

#### 1.5.4 Comparison to previous prescriptions

In Figure 1.21 we compare various prescriptions which have been proposed in the literature to derive BH masses from the C IV line which are consistent with the masses derived from the Balmer lines. In each case, we compare the corrected C IV-based masses to the H $\alpha$ -based masses as a function of the C IV blueshift. The C IV blueshift-based correction presented in this Chapter is also tested in Figure 1.22. The correction proposed by Runnoe et al., (2013) is based on the spectral region at rest-frame wavelengths of  $\sim 1400 \text{ \AA}$  (see below). Therefore, our analysis is based on the 123 quasars with spectra covering this region.



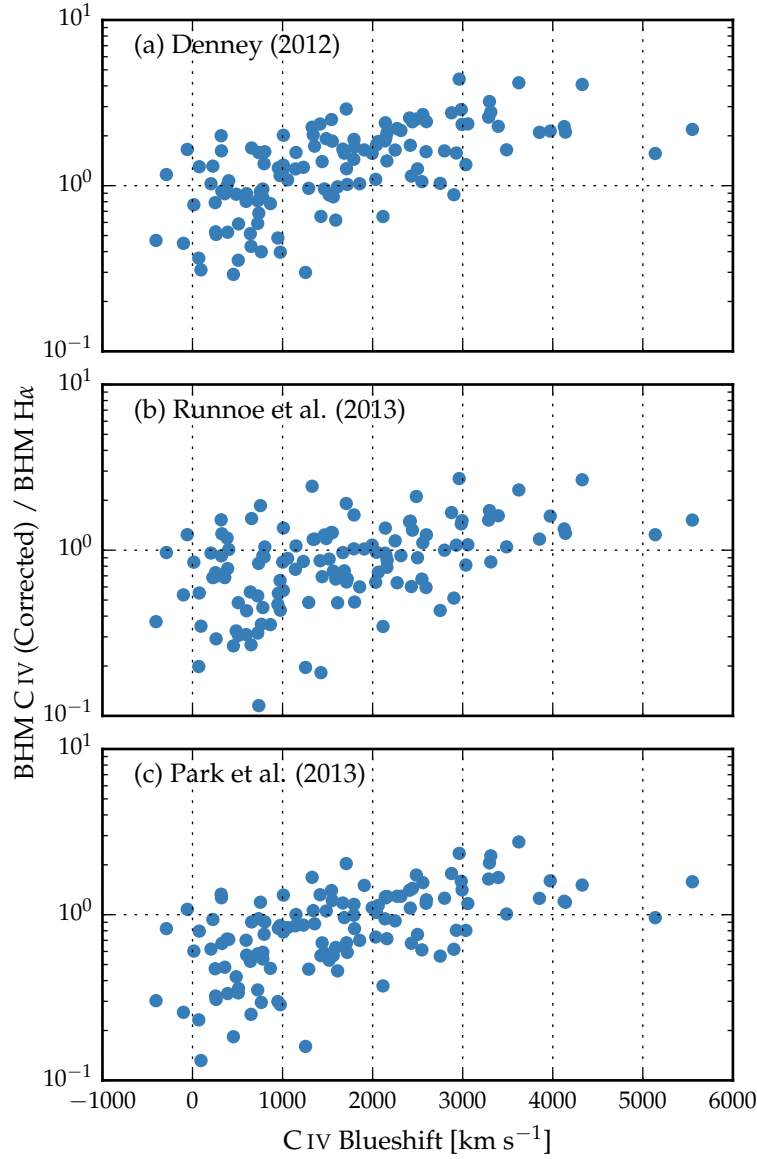


Figure 1.21: Comparison of BH mass estimates derived from C IV and H $\alpha$  as a function of the C IV blueshift. Corrections to the C IV-based masses have been applied based on the shape (FWHM/ $\sigma$ ) of the C IV emission-line (a; Denney, 2012), the peak flux ratio of the Si IV+O IV blend relative to C IV (b; Runnoe et al., 2013), and by significantly reducing the dependence of the derived BH mass on the C IV velocity-width (c; Park et al., 2013).

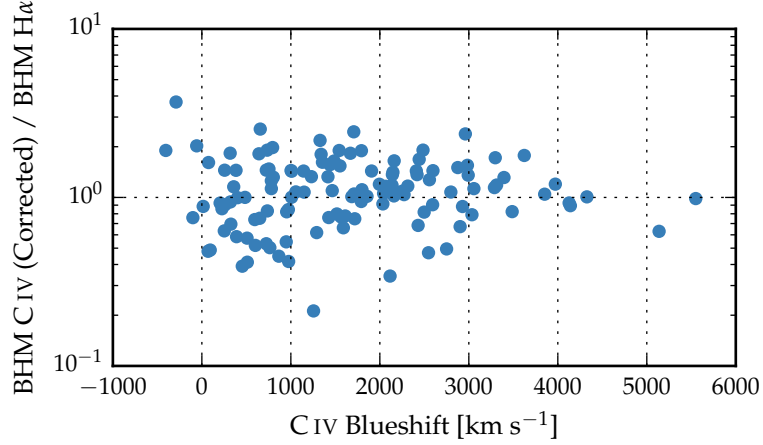


Figure 1.22: Comparison of BH mass estimates derived from C IV and H $\alpha$  as a function of the C IV blueshift. C IV-based masses have been corrected using the C IV blueshift-based prescription presented in this Chapter.

In Figure 1.21a the C IV BH masses have been corrected using the C IV shape (FWHM/ $\sigma$ ) based correction proposed by Denney, (2012). Denney, (2012) found the level of contamination in single-epoch spectra from non-reverberating gas to be correlated with the shape (FWHM/ $\sigma$ ) of the C IV profile. In our sample, we observe a strong correlation between the shape of the C IV line and its blueshift (Figure 1.7c); between the two extremes in the C IV blueshift distribution the line shape changes from FWHM/ $\sigma \sim 1 - 2.5$ . The investigation of Denney, (2012) was based on a sample of reverberation mapped quasars, which have a narrow range of C IV emission-line shapes, including the absence of any objects with large C IV blueshifts. The correction is not applicable at large C IV blueshifts. Therefore, while the consistency between the H $\alpha$ - and C IV-based masses at low C IV blueshifts is improved, at high C IV blueshifts the C IV-based masses remain seriously overestimated.

As explained above, reliably measuring the quasar systemic redshift from the UV region of the spectrum has proved difficult. However, the situation is improved dramatically by the new scheme developed by Allen & Hewett (2017, in preparation). Given the difficulty of measuring reliable C IV blueshifts without the Allen & Hewett scheme, Runnoe et al., (2013) opted instead to use the continuum-subtracted peak flux ratio of the ultraviolet emission-line blend of Si IV+O IV (at 1400 Å) to that of C IV to correct for non-virial contributions to the C IV veloc-

ity width. This parameter was chosen because it showed the strongest correlation with the FWHM C iv/H $\beta$  residuals, as well as with the strengths of optical [O III] and Fe II.

Following Runnoe et al., (2013), we measure the peak flux by fitting a model with four Gaussian components (two for each emission-line) to the continuum-subtracted flux. As is evident from Figure 1.18, a correlation exists between the blueshift and EQW of C iv: C iv emission which is strongly blueshifted is typically weak. The Si iv+O iv emission-line blend, however, shows significantly less systematic variation. Therefore, the Si iv+O iv-based correction is quite effective in practice: the systematic bias in the C iv BH masses at large C iv blueshifts is reduced to a factor of  $\sim 2$  (Figure 1.21b). However, the C iv based masses are still systematically overestimated at large C iv blueshifts.

In contrast to the widely-used Vestergaard and Peterson, (2006) C iv-based virial BH mass calibration, the more recent Park et al., (2013) calibration significantly reduces the dependence of the derived masses on the emission-line velocity width (from the  $V^2$  dependence predicted assuming a virialized BLR to just  $V^{0.56}$ ). As a consequence, the C iv based masses of the quasars with large C iv blueshifts are much reduced (Figure 1.21c). However, the systematic error in the C iv-based BH masses as a function of C iv blueshift remains.

For comparison, the C iv-based masses shown in Figure 1.22 have been corrected using to the C iv blueshift-based procedure presented in this Chapter. No systematic in the BH masses as a function of the C iv blueshift is evident.

## 1.6 POPULATION TRENDS WITH C IV BLUESHIFT

As shown in Figure 1.23, there are systematic variations in the H $\alpha$  line profile as a function of the C iv blueshift. At C iv-blueshift  $< 1200 \text{ km s}^{-1}$ , the H $\alpha$  FWHM range is  $\simeq 2000 - 8900 \text{ km s}^{-1}$ , with mean  $\simeq 4300 \text{ km s}^{-1}$ . However, amongst the quasars with C iv-blueshift  $> 2000 \text{ km s}^{-1}$ , the mean H $\alpha$  FWHM =  $3500 \text{ km s}^{-1}$ , with a scatter of just  $700 \text{ km s}^{-1}$ . The apparent trend of peakier H $\alpha$ -emission, with FWHM/ $\sigma$  close to unity, at large C iv-blueshift is enhanced by the modest increase in H $\alpha$  EQW with blueshift. Amongst the low-C iv-blueshift population there are in addition quasars with broader and more Gaussian-like H $\alpha$  line profiles, with FWHM/ $\sigma \simeq 2$ .

The change in the H $\alpha$  emission-line profiles as a function of C iv-blueshift means that the H $\alpha$ -FWHM derived BH masses

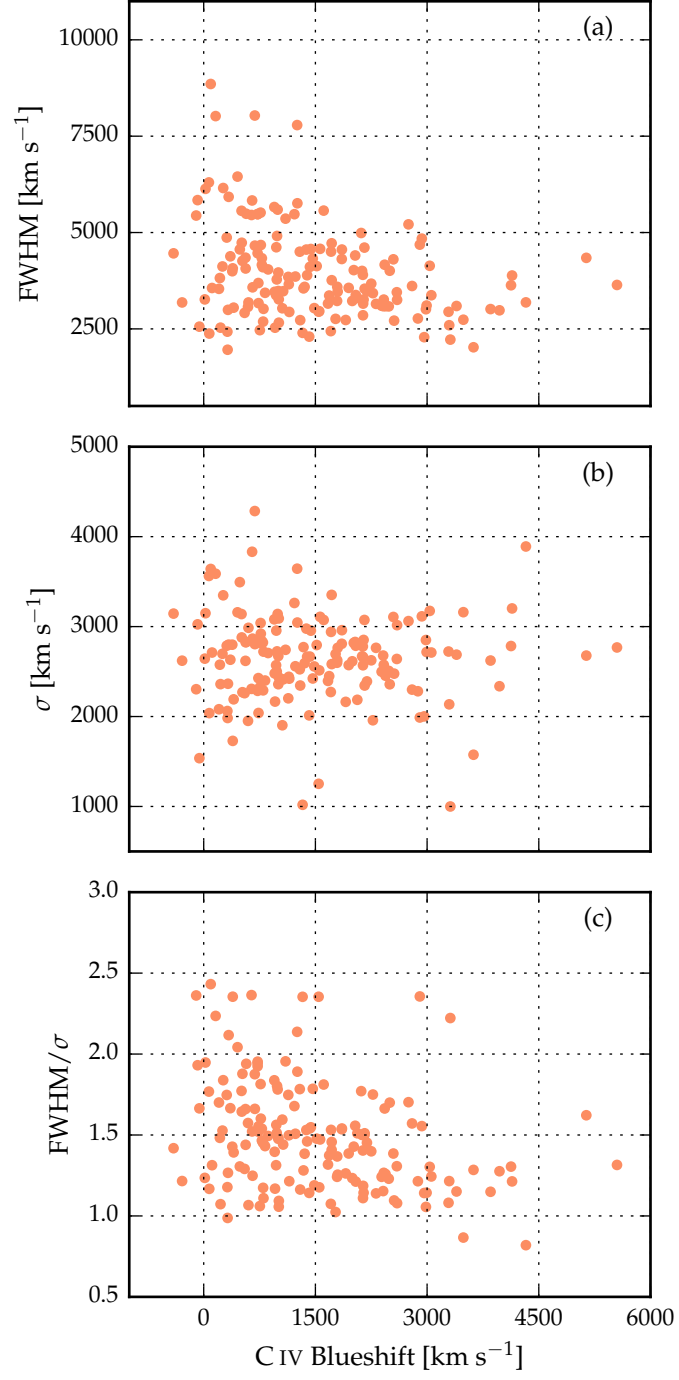


Figure 1.23: The FWHM, dispersion ( $\sigma$ ) and shape (FWHM/ $\sigma$ ) of H $\alpha$  as a function of the C IV blueshift.

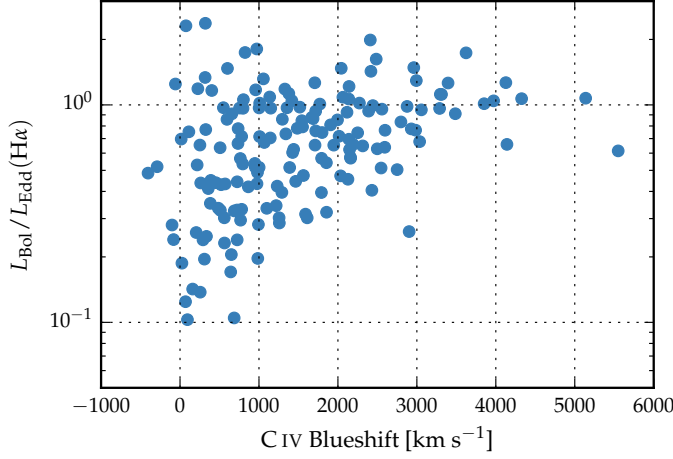


Figure 1.24:  $H\alpha$ -derived Eddington ratio versus C IV blueshift. At blueshift  $\gtrsim 2000 \text{ km s}^{-1}$  all quasars have high accretion rates ( $L/L_{\text{Edd}} \simeq 1$ ). This is in agreement with Kratzer and Richards, (2015), but in contrast to what one would derive from naive use of C IV-based BH mass scaling relations.

at high-blueshift are smaller than the sample mean. We transformed the observed luminosity into a mass-normalised accretion rate (Eddington ratio). To convert the monochromatic luminosity, which is observed, into a bolometric luminosity we use the bolometric correction factor given by Richards et al., (2006) ( $L_{\text{bol}} = 9.26L_{5100}$ ). Although there is evidence that the bolometric correction factor is a function of the luminosity, as well as of other parameters including the C IV blueshift (Krawczyk et al., 2013), the differences are small over the parameter range covered by our sample, and for simplicity we adopt a constant factor.

The results, shown in Figure 1.24, demonstrate that at large blueshifts quasars are accreting at around their Eddington limits (Figure 1.24). This finding is in accord with our interpretation that the blueshifting of C IV is evidence for strong outflows resulting from the presence of a radiation-driven accretion-disc wind. Richards et al., (2002) found that quasars with large C IV blueshifts have weak He II. This is evidence for weak soft X-ray continuum emission (Leighly, 2004), which would allow a strong line-driven wind to form. The strength of such a wind is predicted to be related to the quasar far-ultraviolet SED, which, in turn, could be related to the mass-accretion rate.

All of the objects in our sample which exhibit large C IV blueshifts would be classified as population A in the Sulentic et al., (2000) scheme based on the H $\alpha$  FWHM. Our results therefore support the idea of the Sulentic et al., (2000) A/B division being driven by the Eddington ratio, with population A sources possessing higher accretion rates. However, we also observe a number of quasars which have high Eddington ratios but do not have line profiles suggestive of strong outflows in the C IV BLR. This suggests that a high accretion rate is a necessary but not sufficient condition for the existence of outflows (Baskin and Laor, 2005).

The two-dimensional nature of the C IV emission-line parametrization and the apparent anti-correlation between C IV EQW and C IV blueshift suggests that the quasar population exhibits a continuum of properties. As such, more accurate C IV blueshift measurements for SDSS-quasars should allow an improved mapping between the C IV-emission properties and key physical parameters of the quasars. This includes improving our understanding of the origin of quasars with exceptionally weak, blueshifted C IV emission (weak emission-line quasars; Luo et al., 2015) which could be exotic versions of wind-dominated quasars (Plotkin et al., 2015).

As we described in Section 1.5.2, the shape of the Balmer lines (FWHM/ $\sigma$ ) depends strongly on the FWHM. This suggests that the BLR structure is changing with the emission-line FWHM. One possibility we discussed is that the Balmer line-width is orientation-dependent (e.g. Shen and Ho, 2014). This raises the question of whether the narrow H $\alpha$  emission-lines observed in the quasars with the largest C IV blueshifts could be an orientation effect. However, there is no evidence that the C IV blueshift is dependent on the orientation (inferred from the radio core-dominance; Richards et al., 2011; Runnoe et al., 2014). Furthermore, Leighly, (2004) showed that the He II  $\lambda$ 1640 emission-line properties of quasars with large C IV blueshifts are more consistent with differences in the SED rather than differences in the orientation. Overall, therefore, orientation does not appear to be the dominant effect in determining the C IV blueshift and correlated changes in the H $\alpha$  line profile.

#### 1.6.1 *The BAL parent population*

Classical high-ionization BAL (HiBAL) quasars are also predominantly Population A objects in the scheme of Sulentic

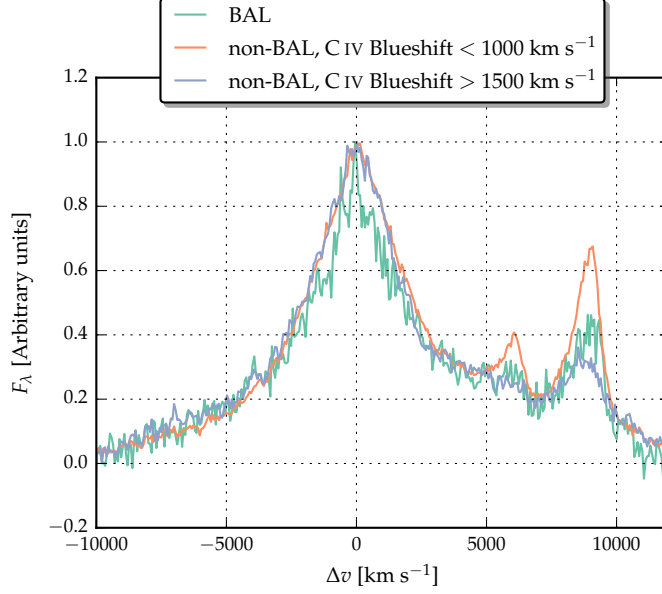


Figure 1.25

et al., (2000). There are no HiBAL quasars in our sample by design but it is generally accepted that quasars which show high-ionisation BALs are likely to be radiating with relatively high  $L/L_{\text{Edd}}$  (e.g. Zhang et al., 2014). We therefore propose that the subset of the quasar population that exhibits large C IV-emission blueshifts, with high-EQW and narrow-H $\alpha$  emission-lines, may be directly related to the HiBAL quasar population – perhaps even the ‘parent’ population (Richards, 2006). A prediction of such a linkage is that near-infrared observations of the rest-frame optical spectra of HiBAL quasars will show strong, relatively narrow, Balmer emission-lines, very similar to those of the quasars with high C IV-blueshifts presented in this Chapter.

To test this hypothesis, we selected 18 C IV BAL quasars from our catalogue with near-infrared spectra including H $\beta$ . Using the same method described in Section XX, we constructed a median composite spectrum from this sample. We also constructed composite spectra for quasars with modest and large C IV blueshifts. The results are shown in Figure 1.25. We find that [O III] is very weak and H $\beta$  is relatively narrow in the BAL quasars (e.g. Yuan and Wills, 2003). Therefore, as predicted, the optical emission-line properties of the BAL quasars are very similar to the properties of the non-BAL quasars with large C IV blueshifts.

### 1.6.2 *The frequency of quasars with high accretion rates*

Quantifying the frequency of quasars producing outflows as a function of key parameters, e.g. quasar luminosity, BH mass, redshift, etc. will be important to constrain models of quasar-galaxy evolution. At fixed BH mass, the intrinsic and the observed fraction of quasars exhibiting properties that depend on the Eddington ratio can differ significantly. As an illustration, we consider the implications for the intrinsic fraction of quasars possessing large C IV blueshifts given the observed numbers in the  $m_i < 19.1$  flux-limited sub-sample of the SDSS DR7 quasar catalogue. In order to estimate the size of the selection effect, we considered the detection probability for a much-simplified quasar population. We assume that all quasars with C IV blueshifts  $> 1200 \text{ km s}^{-1}$  have enhanced accretion rates relative to the ‘normal’ population (with C IV blueshifts  $< 1200 \text{ km s}^{-1}$ ). If the accretion rate of the high-blueshift population is double the rate of the low-blueshift population (which is true in an average sense – see Figure 1.24), then the high-blueshift population will be brighter by  $\simeq 0.75$  magnitude. Under the assumption that the BH mass distribution is independent of the C IV blueshift, the high-blueshift population will then be over-represented in a flux-limited sample. To estimate the size of the bias, we need to know how many more quasars, at redshifts  $2 < z < 2.5$ , there are with  $m_i < 19.1 + 0.75 = 19.85$  relative to  $m_i < 19.1$ . This is the fraction of the population which, as a consequence of having enhanced accretion rates, are boosted above the survey flux limit. The main colour-selected SDSS DR7 quasar catalogue extends only to  $m_i = 19.1$  and, assuming the luminosity function is continuous<sup>6</sup> we thus use the number counts at  $m_i < 19.1$  and  $m_i < 18.35$ , which differ by a factor of  $\simeq 4$ .

At redshifts  $2 < z < 2.5$ , there are 3,834 quasars with C IV blueshifts  $< 1200 \text{ km s}^{-1}$  and 2,484 with blueshifts  $> 1200 \text{ km s}^{-1}$  in the SDSS DR7  $m_i < 19.1$  quasar sample, a ratio of  $\sim 2 : 1$ . The above calculation, although much idealised, suggests that the intrinsic fraction of high-blueshift quasars is a factor of four smaller than in the flux-limited sample (i.e.  $\sim 15$  per cent of the ultraviolet-selected non-BAL quasar population).

<sup>6</sup> The luminosity function and number-counts vary only smoothly (e.g. Ross et al., 2013) for the magnitude and redshift range used here.



## 1.7 CONCLUSIONS

The main results of this Chapter are as follows:

- We have analysed the spectra of 230 high-luminosity ( $10^{45.5} - 10^{48} \text{ erg s}^{-1}$ ), redshift  $1.5 < z < 4.0$  quasars for which spectra of the Balmer emission-lines and the C iv emission-line exist. The large number of quasars in our spectroscopic catalogue and the wide range in C iv blueshifts the quasars possess has allowed us to directly investigate biases in C iv-based BH mass estimates which stem from non-virial contributions to the C iv emission as a function of the C iv blueshift, which, in turn, depends directly on the form of the quasar ultraviolet SEDs (Richards et al., 2011).
- The C iv emission-based BH masses are systematically in error by a factor of more than five at  $3000 \text{ km s}^{-1}$  in C iv emission blueshift and the overestimate of the BH masses reaches a factor of 10 for quasars exhibiting the most extreme blueshifts,  $\gtrsim 5000 \text{ km s}^{-1}$ .
- We have derived an empirical correction formula for BH mass estimates based on the C iv emission-line FWHM and blueshift. The correction may be applied using Equations 1.4 and 1.6 in Section 1.4.3. The large SED-dependent systematic error in C iv-based BH masses is removed using the correction formulae. The remaining scatter between the corrected C iv-based masses and the H $\alpha$ -based masses is 0.24 dex at low C iv blueshifts ( $\sim 0 \text{ km s}^{-1}$ ) and 0.10 dex at high blueshifts ( $\sim 3000 \text{ km s}^{-1}$ ). This is a significant improvement on the 0.40 dex scatter observed between the un-corrected C iv and H $\alpha$  BH masses. The correction depends only on the C iv line properties - i.e. the FWHM and blueshift - and allows single-epoch virial BH mass estimates to be made from optical spectra, such as those provided by the SDSS, out to redshifts exceeding  $z \sim 5$ .

## 1.8 CATALOGUE OF DERIVED PROPERTIES

Table 1.4 includes the line parameters from our emission line fits, and other derived properties used in this Chapter. The columns in Table 1.4 are as follows:

- 1 Catalogue name
- 2-3 Broad H $\alpha$  FWHM, and its error, in km s $^{-1}$ .
- 4-5 Broad H $\alpha$  line dispersion, and its error, in km s $^{-1}$ .
- 6-7 Broad H $\alpha$  redshift, and its error.
- 8-9 H $\alpha$ -FWHM-based BH mass using Vestergaard and Peterson, (2006) calibration, and its error, in  $M_{\odot}$ . H $\alpha$  FWHM is first converted into equivalent H $\beta$  FWHM using Equation 1.3.
- 10 H $\alpha$  S/N flag. When flag is 1 the S/N around H $\alpha$  is low, and the emission line is excluded from the main analysis in this Chapter.
- 11-12 Broad H $\beta$  FWHM, and its error, in km s $^{-1}$ .
- 13-14 Broad H $\beta$  line dispersion, and its error, in km s $^{-1}$ .
- 15-16 Broad H $\beta$  redshift, and its error.
- 17-18 H $\beta$ -FWHM-based BH mass using Vestergaard and Peterson, (2006) calibration, and its error, in  $M_{\odot}$ .
- 19 H $\beta$  S/N flag. When flag is 1 the S/N around H $\beta$  is low, and the emission line is excluded from the main analysis in this Chapter.
- 20-21 Broad C iv FWHM, and its error, in km s $^{-1}$ .
- 22-23 Broad C iv line dispersion, and its error, in km s $^{-1}$ .
- 24-25 Broad C iv EQW, and its error, in Å.
- 26-27 Si iv+O iv/C iv peak flux ratio, and its error.
- 28-29 C iv blueshift, relative to H $\alpha$ , and its error, in km s $^{-1}$ .
- 30-31 C iv blueshift, relative to H $\beta$ , and its error, in km s $^{-1}$ .
- 32-33 Uncorrected C iv-FWHM-based BH mass using Vestergaard and Peterson, (2006) calibration, and its error.
- 34-35 Corrected C iv-FWHM-based BH mass, and its error.

Table 1.4: The format of the table containing the emission-line properties from our parametric model fits. The table will be available in machine-readable form online.

Column	Name	Units	Description
1	UID		Catalogue name
2	FWHM_BROAD_HA	$\text{km s}^{-1}$	Broad H $\alpha$ FWHM
3	FWHM_BROAD_HA_ERR	$\text{km s}^{-1}$	
4	SIGMA_BROAD_HA	$\text{km s}^{-1}$	Broad H $\alpha$ $\sigma$
5	SIGMA_BROAD_HA_ERR	$\text{km s}^{-1}$	
6	Z_BROAD_HA		H $\alpha$ redshift
7	Z_BROAD_HA_ERR		
8	LogMBH_Ha	$M_{\odot}$	H $\alpha$ BH mass
9	LogMBH_Ha_Err	$M_{\odot}$	
10	SNR_FLAG_Ha		H $\alpha$ S/N flag
11	FWHM_BROAD_HB	$\text{km s}^{-1}$	Broad H $\beta$ FWHM
12	FWHM_BROAD_HB_ERR	$\text{km s}^{-1}$	
13	SIGMA_BROAD_HB	$\text{km s}^{-1}$	Broad H $\beta$ $\sigma$
14	SIGMA_BROAD_HB_ERR	$\text{km s}^{-1}$	
15	Z_BROAD_HB		H $\beta$ redshift
16	Z_BROAD_HB_ERR		
17	LogMBH_Hb	$M_{\odot}$	H $\alpha$ BH mass
18	LogMBH_Hb_Err	$M_{\odot}$	
19	SNR_FLAG_Hb		H $\beta$ S/N flag
20	FWHM_CIV	$\text{km s}^{-1}$	Broad C iv FWHM
21	FWHM_CIV_ERR	$\text{km s}^{-1}$	
22	SIGMA_CIV	$\text{km s}^{-1}$	Broad C iv $\sigma$
23	SIGMA_CIV_ERR	$\text{km s}^{-1}$	
24	EQW_CIV	$\text{\AA}$	Broad C iv EQW
25	EQW_CIV_Err	$\text{\AA}$	
26	1400_CIV		Si iv+O iv/C iv peak flux ratio
27	1400_CIV_ERR		
29	BLUESHIFT_CIV_HA	$\text{km s}^{-1}$	C iv blueshift, relative to H $\alpha$
30	BLUESHIFT_CIV_HA_ERR	$\text{km s}^{-1}$	
31	BLUESHIFT_CIV_HB	$\text{km s}^{-1}$	C iv blueshift, relative to H $\beta$
32	BLUESHIFT_CIV_HB_ERR	$\text{km s}^{-1}$	
33	LogMBH_CIV_VP06	$M_{\odot}$	Uncorrected C iv BH mass
34	LogMBH_CIV_VP06_Err	$M_{\odot}$	
35	LogMBH_CIV_C17	$M_{\odot}$	Corrected C iv BH mass
36	LogMBH_CIV_C17_Err	$M_{\odot}$	

