1

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 black holes (BHs) and the masses of host-galaxy spheroids has led to much work focussing on the properties of quasars and active galactic nuclei (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 Hβ 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 the Sloan Digital Sky Survey (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 (~ 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 Hβ (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 Mg II, the 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 C IV emission correlates with a number of properties of quasar spectral energy distributions (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).

Excess emission-line flux in the blue wing of the C IV emission increases commonly employed measures of the line-width, notably the full-width at half maximum (FWHM) and the line dispersion (σ). In general, researchers studying quasar demographics at high-redshift adopt estimates of BH masses based on the width of C IV-emission, without reference to the blueshift of the C IV-emission (e.g. Vestergaard, 2004; Kollmeier et al., 2006; Gavignaud et al., 2008; Vestergaard et al., 2008; Vestergaard and Osmer, 2009; Kelly et al., 2010; Kelly and Shen, 2013). Figure 1.1 shows the shape of the C IV-emission in composite spectra constructed from SDSS DR7 quasars as a function of C IV blueshift. The profiles show how, at large values of blueshift ($\gtrsim$ 2000 km s^{-1}) the C IV-profile is displaced to the blue by amounts comparable to the FWHM of the profile. This indicates that non-virial motions, very likely due to outflows, are having a significant effect on the observed C IV emission
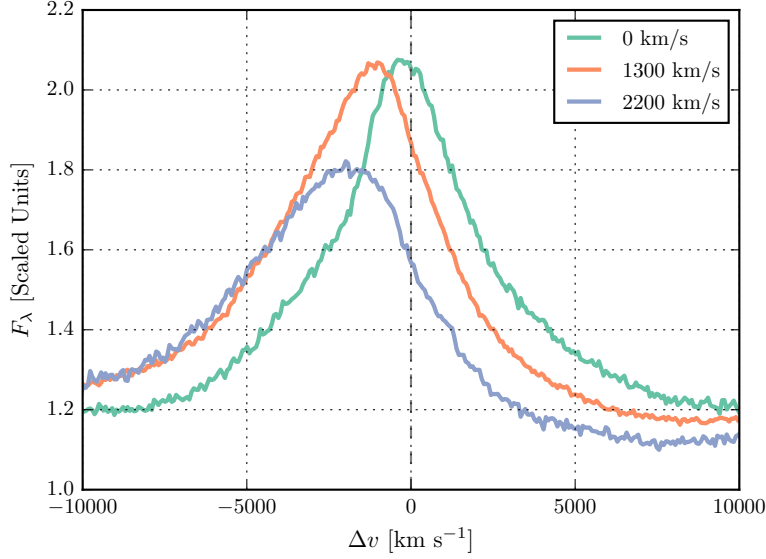
Figure 1.1: Composite spectra of the C IV emission-line as a function of C IV blueshift for SDSS DR7 quasars. Quasars classified as BALs, or possessing strong associated absorbers have been excluded, and the composite-spectra shown are derived using an arithmetic mean of a minimum of 200 spectra at each 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^{-1} and EW 33.1, 31.6, and 28.8 Å.

velocity profile (e.g. Gaskell, 1982; Baskin and Laor, 2005; Sulentic et al., 2007; Richards et al., 2011; Wang et al., 2013). At fixed emission-line EW, 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. 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 C IV emission-line in quasars exhibiting large blueshifts. 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; Coatman et al., 2016).

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 equivalent widths 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 EWs 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 IVλ1400 in Runnoe et al. 2013a 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 Hβ and Hα 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 Hα and Hβ 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 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 Hα line measurements and 144 Hβ 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 Hα and Hβ. The sample is considerably larger than previous studies of the rest-frame optical spectra of high-z quasars

Table 1.1: The numbers of quasars with reliable Hα and Hβ line measurements, and the spectrographs and telescopes used to obtain the near-infrared spectra

Spectrograph	Telescope	Hα Sample	Hβ 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

(e.g. Shen and Liu, 2012). As we demonstrate in Section 1.5.3, the quasars have C IV blueshifts of up to $\sim 5000 \, \mathrm{km \, s^{-1}}$, and span the full range observed in the population.

1.2.1 *Spectroscopic data*

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 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 Hα line measurements (the 'Hα sample') and quasars with reliable Hβ measurements (the 'Hβ 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 Hβ time-lags and velocity widths, but the line-widths of Hα and Mg $\textsc{ii}\lambda 2800$ 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 Hα and Hβ 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 $\textsc{iv}$ emission-line 'blueshift' provides the basis for the corrected C $\textsc{iv}$ velocity-width measurements, and hence BH masses. The effectiveness of the scheme is validated via a direct comparison of the C $\textsc{iv}$ 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, λ_{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 $\textsc{iii}$]$\lambda\lambda 4960,5008$ doublet for the subset of spectra where both are available (Section 1.4.2). Second, the blueshift of the C $\textsc{iv}$ emission-line is determined. Again, we adopt the line centroid to provide a robust measure of the C $\textsc{iv}$ emission blueshift. The blueshift (in km s^{-1}) is defined as c$\times$(1549.48-λ_{half})/1549.48 where c is the velocity of light and 1549.48 Å is the rest-frame wavelength for the C $\textsc{iv}$ doublet[1]. 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

---

[1] The adopted C $\textsc{iv}$ 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 ~ 80 km s^{-1}.

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 λ_{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; Denney 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, α, 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 $\pm 10\,000\,\mathrm{km\,s^{-1}}$ 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
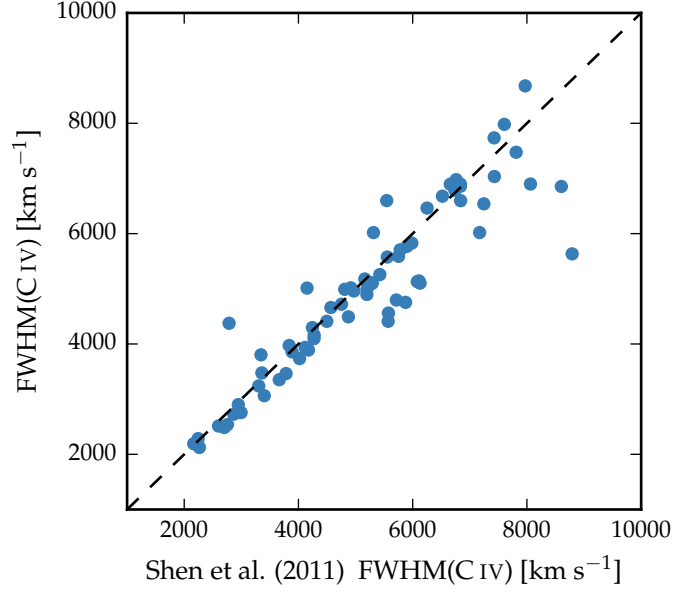
7

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).

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 Fig. 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 ($200 \, \mathrm{km \, s^{-1}}$).

1.3.2 *Hα*

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 com-

ponent of Hα is fit using one or two Gaussians, constrained to have a minimum FWHM of 1200 km s^{-1}. When two Gaussians are used, the velocity centroids are constrained to be the same.

The emission-line profiles of both Hβ and Hα 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 Hα as well as [N II]λλ6548,6584 and [S II]λλ6717,6731. This resulted in a better fit to the observed flux in 50 per cent of cases. We impose a 1200 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 [N II] components is also fixed at the expected value of 2.96. In 70 per cent of the spectra the [O III]λλ4960,5008 doublet is detected at moderate S/N in the Hβ region. In these cases the peak of the [O III] is used to fix the velocity offsets and the FWHMs of the narrow line components in the Hα region. For spectra where the [O III] doublet does not constrain the velocity and FWHM accurately, the narrow emission in the Hα and Hβ 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 Hα and Hβ emission regions simultaneously and we fix the centroids of the Gaussian components used to fit the broad emission. In Fig. 1.3 we plot our Hα 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 Hβ and [O III]

Emission from optical Fe II is generally strong in the vicinity of Hβ. 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
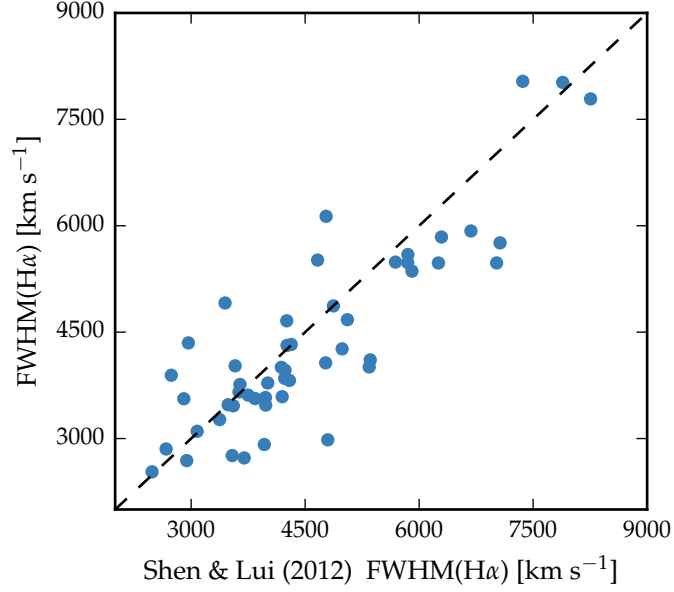
Figure 1.3: Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the Hα FWHM with Shen and Liu, (2012).

fit. We use the same model to fit the broad and narrow components of Hβ as was used with Hα. Each line in the [O III] doublet is fit with two Gaussians, to model both the systemic and any outflow contributions. The peak flux ratio of the [O 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 Hα, the width and velocity offsets of all the narrow components are set to be equal, and an upper limit of $1200\,\mathrm{km\,s^{-1}}$ is placed on the FWHM.

The parametric model we fit to the Hβ/[O III] emission region was very similar to the model employed by Shen, (2016). In Fig. 1.4 we plot our Hβ 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.
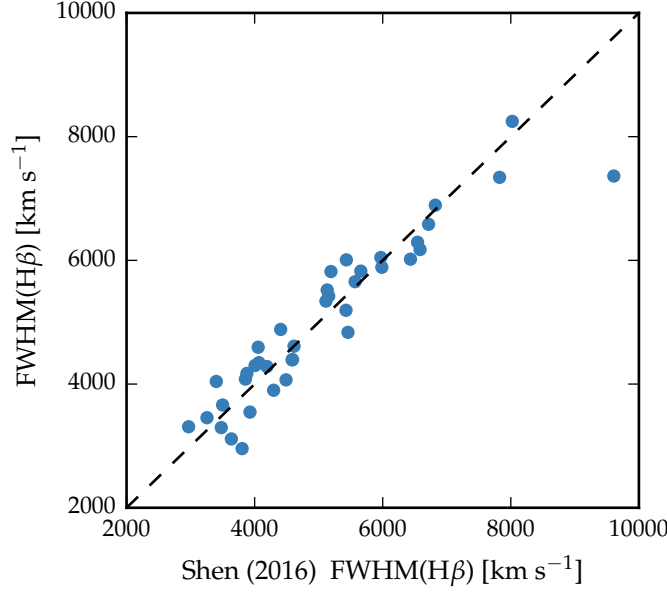
Figure 1.4: Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of the Hβ FWHM with Shen, (2016).

In Fig. 1.5 we present our parametric fits to the C IV, Hα and Hβ 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 Hα emission-line centroid is at $0\,\mathrm{km\,s^{-1}}$. The y-axes of the data-minus-model residual plots have been scaled by the spectrum flux errors. The mean reduced-χ^2 values in our Hα, Hβ and C IV fits are 1.69, 1.62, and 1.77 respectively and, in general, there are no strong features observable in the spectrum minus 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.

Table 1.2 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 format of the table containing the emission-line properties from our parametric model fits. The table is available in machine-readable form in the online version of Coatman et al., (2017).

	Units	Description
NAME		Catalogue name
FWHM_BROAD_HA	$\mathrm{km\,s^{-1}}$	FWHM of broad Hα line
FWHM_BROAD_HA_ERR	$\mathrm{km\,s^{-1}}$	
SIGMA_BROAD_HA	$\mathrm{km\,s^{-1}}$	Dispersion of broad Hα line
SIGMA_BROAD_HA_ERR	$\mathrm{km\,s^{-1}}$	
Z_BROAD_HA		Redshift from broad Hα line
FWHM_BROAD_HB	$\mathrm{km\,s^{-1}}$	FWHM of broad Hβ line
FWHM_BROAD_HB_ERR	$\mathrm{km\,s^{-1}}$	
SIGMA_BROAD_HB	$\mathrm{km\,s^{-1}}$	Dispersion of broad Hβ line
SIGMA_BROAD_HB_ERR	$\mathrm{km\,s^{-1}}$	
Z_BROAD_HB		Redshift from broad Hβ line
FWHM_CIV	$\mathrm{km\,s^{-1}}$	FWHM of C IV doublet
FWHM_CIV_ERR	$\mathrm{km\,s^{-1}}$	
SIGMA_CIV	$\mathrm{km\,s^{-1}}$	Dispersion of C IV doublet
SIGMA_CIV_ERR	$\mathrm{km\,s^{-1}}$	
BLUESHIFT_CIV_HA	$\mathrm{km\,s^{-1}}$	Blueshift of C IV relative to Hα
BLUESHIFT_CIV_HA_ERR	$\mathrm{km\,s^{-1}}$	
BLUESHIFT_CIV_HB	$\mathrm{km\,s^{-1}}$	Blueshift of C IV relative to Hβ
BLUESHIFT_CIV_HB_ERR	$\mathrm{km\,s^{-1}}$	
LOGL5100	$\mathrm{erg\,s^{-1}}$	Luminosity at 5100 Å
LOGL1350	$\mathrm{erg\,s^{-1}}$	Luminosity at 1350 Å

Table 1.3: The number of spectra removed from our sample by the cuts described in Section 1.3.5.

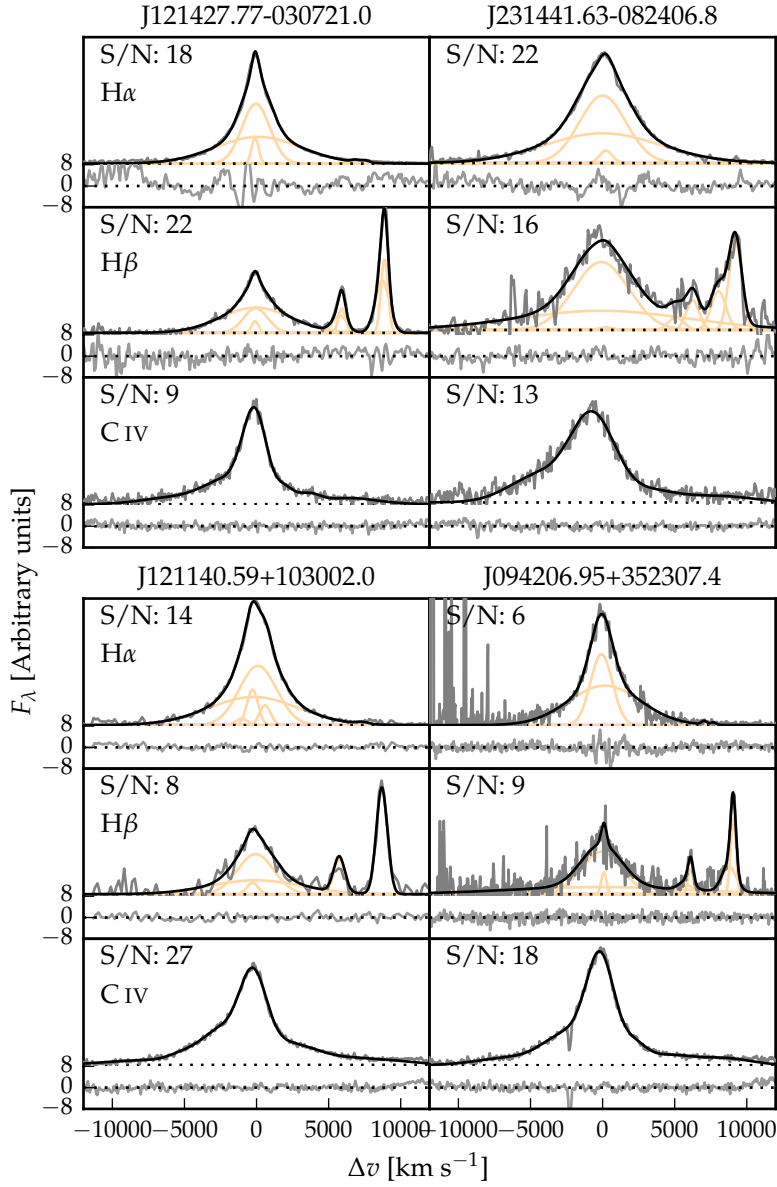|		Hα sample	Hβ sample
Total		194	279
Hα/Hβ	Wavelength	6	27
	S/N	8	83
C IV	Wavelength	6	5
	S/N	4	12
	Absorption	6	8
Total remaining		164	144

Figure 1.5: Model fits to continuum-subtracted Hα, Hβ, and C IV emission in four quasars, chosen to represent the range of S/N (indicated in the figure and given per $150\,\mathrm{km\,s^{-1}}$ 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 Hα emission is used to set the redshift, and Δv 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 1.3 and the distribution in redshift and luminosity is shown in Fig. 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 Hα emission falls at the edge of the K-band. Hβ is much weaker than Hα, and the Hβ spectra are generally of lower S/N. As a result, the fraction of Hβ spectra that are flagged – 39 per cent – is particularly high.

1.3.6   *Emission-line parameter uncertainties*

The 1σ 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[2]. We therefore chose to measure the spread in the parameter distribution by

---

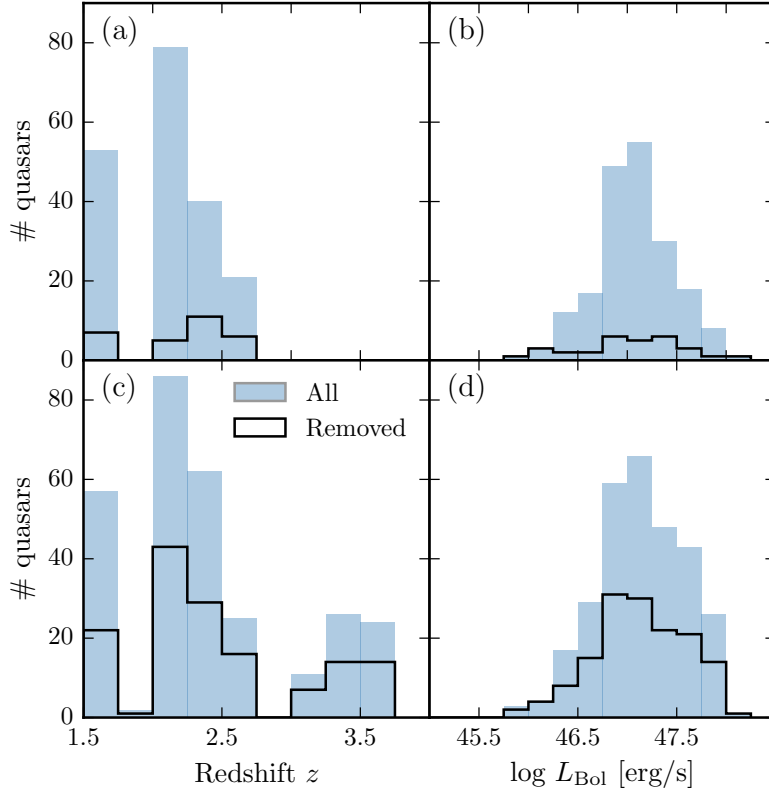2 In the analysis of the real spectra such fits are identified via visual inspection.

Figure 1.6: The redshift and luminosity distributions of the spectra removed from our Hα/C ɪᴠ (a, b) and Hβ/C ɪᴠ (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 ~14 years after the SDSS spectra, and the VLT SINFONI spec-

tra 20 years or more after the Hamburg/ESO observations[3]. 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 ∼ 8 years. The root-mean-square difference in the C IV FWHM measured from the BOSS and SDSS spectra is a modest $\simeq 500 \, \mathrm{km \, s^{-1}}$. 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 $500 \, \mathrm{km \, s^{-1}}$. 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 Å for C IV and 5100 Å for Hα and Hβ. The calculation of these luminosities is described in Chapter **??**.

As described in Chapter **??**, we estimate the uncertainties on the monochromatic luminosities to be ∼ 0.3 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, σ, 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

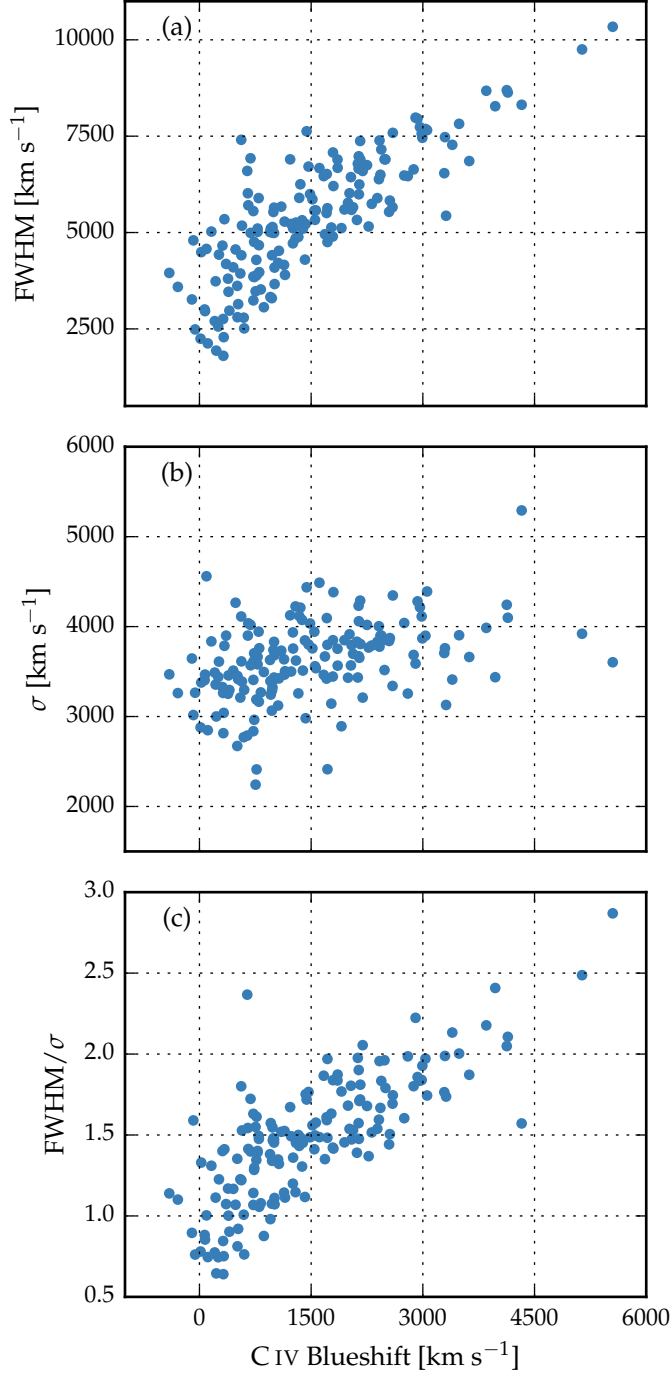3 Time differences in the quasar rest-frame are reduced by a factor of $(1 + z)$.

Figure 1.7: The FWHM, dispersion (σ) and shape (FWHM/σ) of C IV as a function of the C IV blueshift.

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 Fig. 1.7 we show how the C IV FWHM, line dispersion, σ, and line shape, FWHM/σ, 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 $1200 \, \mathrm{km \, s^{-1}}$ 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

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

To emphasise this point, in Fig. 1.8 we overlay the C IV line profiles of SDSSJ1236+1129 and SDSSJ1525+2928, whose dispersions are indistinguishable (4168 ± 271 and $4303 \pm 128 \, \mathrm{km \, s^{-1}}$ 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 Shen and Liu, 2012; Sulentic et al., 2007), 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 C IV-BASED VIRIAL BH-MASS ESTIMATES

1.4.1   *Hα/Hβ FWHM comparison*

BH-mass calibrations which use the width of the broad Hβ emission-line as a proxy for the virial velocity are widely regarded as the most reliable, since most reverberation mapping employs the Hβ line and the $R - L$ relation has been established using Hβ. When Hβ is not available, Hα 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 Hα and Hβ lines. The 99 objects include 21
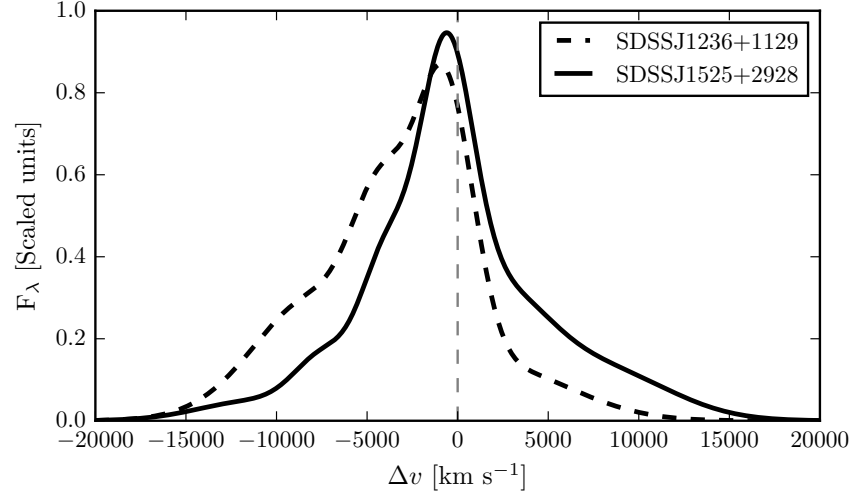
Figure 1.8: Comparison of the C IV line profiles of SDSSJ1236+1129 and SDSSJ1525+0426. Notwithstanding the essentially identical dispersion values, the emission-line velocity fields differ dramatically and, therefore, the dispersion values cannot be measuring accurately the virial-induced velocity spread of the C IV emission in both quasars.
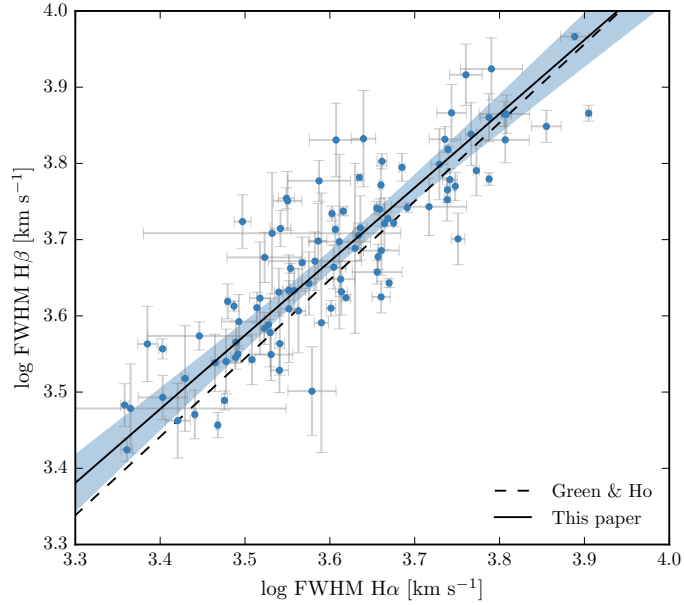

Figure 1.9: Comparison of Hα and Hβ FWHM measurements for 99 quasars. The solid line is our best-fitting power-law model, and the blue-shaded region shows the 2-σ uncertainties on the model parameters. The dashed line is the relation found by Greene and Ho, (2005) using a sample of $z < 0.35$ SDSS AGN.

quasars which were excluded from the main 308-object cata-
logue because the C IV FWHM and/or blueshift could not be
measured reliably. The line widths are compared in Fig. 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 Hα and Hβ FWHMs

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

(1.1)

The relation is shown as the dashed line in Fig. 1.9. The root-
mean-square scatter about this relation is 0.07 dex, compared
to the ~ 0.1 dex found by Greene and Ho, (2005). However, we
find a systematic offset, in the sense that the Hβ line-widths
we measure are on average larger by $270 \ \mathrm{km \ s^{-1}}$ 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 Hα and
Hβ FWHMs.

We assume a relation of the same form used by Greene and
Ho, (2005), i.e. a simple power-law, and infer the model param-
eters by fitting a linear model (with slope α and intercept β) 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 Σ_i^2, with a Gaussian of variance V representing the
intrinsic variance in the data. The log-likelihood is then given
by

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

(1.2)

where Δ_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 sug-
gestion by Hogg, Bovy, and Lang, (2010), the linear model was

parametrized in terms of $(\theta, b_\perp)$, where θ 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).

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

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

and the shaded region shows the 2σ uncertainties on the model parameters.

As discussed above, our relation is displaced to slightly higher Hβ FWHM than the Greene and Ho, (2005) relation – the offset is $210\,\mathrm{km\,s^{-1}}$ for a quasar with Hα FWHM $4500\,\mathrm{km\,s^{-1}}$. 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, σ_I, we infer from the fit is $0.04\,\mathrm{dex}$. This is smaller than the total scatter seen in Fig. 1.9 ($0.06\,\mathrm{dex}$), which suggests that measurement errors make a significant contribution to the total scatter in the relation.

We constructed composite spectra of the Hα and Hβ 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 $1\,\mathrm{\mathring{A}}$ resolution. The spectra were scaled by the mean flux in the interval 4700-5100 $\mathrm{\mathring{A}}$ (Hβ) and 6400-6800 $\mathrm{\mathring{A}}$ (Hα). The composite was then defined as the median flux from all of the normalised spectra in each wavelength bin. The Hα and Hβ lines in the composite spectra are shown in Fig. 1.11. The cores of the two lines are very similar, but Hβ has more flux in the wings of the line.

For 19 of the 99 quasars with Hβ and Hα emission profiles, one of the two Gaussians used to reproduce the Hβ profiles has a FWHM greater than $20\,000\,\mathrm{km\,s^{-1}}$ and a fractional contribution to the total Hβ broad line flux of > 0.3 (Marziani et al., 2009; Marziani et al., 2013). Such a broad component is not seen in the Hα profiles and the very broad Hβ-component may be
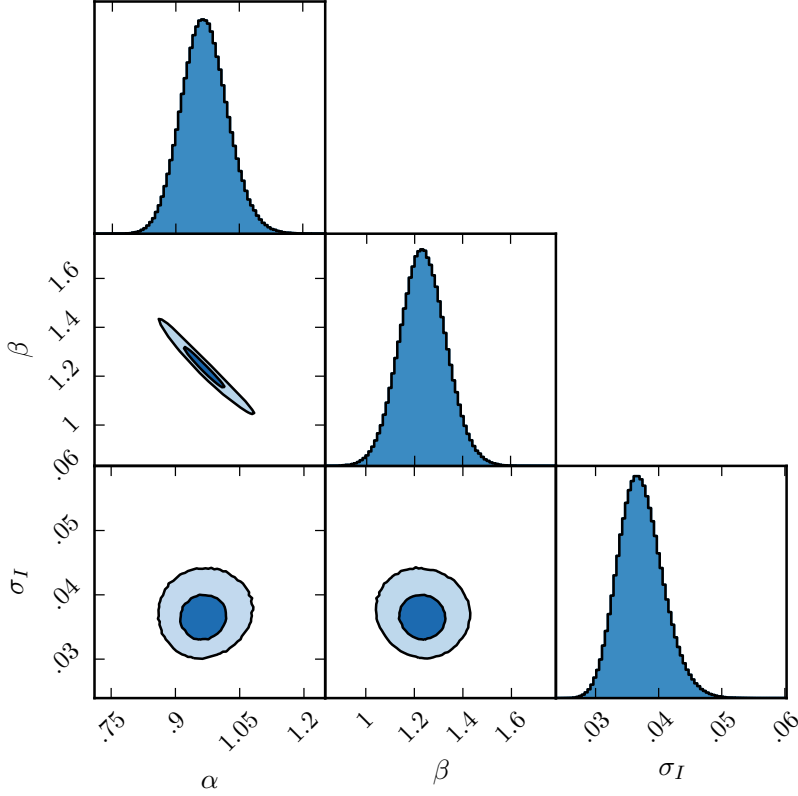
Figure 1.10: One- and two-dimensional projections of the MCMC sampling of the posterior distribution from the fit in Fig. 1.9. α is the power-law index, 10^β is the normalisation, and σ_I is the intrinsic scatter. In the two-dimensional projections, 1- and 2-σ contours are shown.

an artifact of the fitting scheme. A particular issue for Hβ is the presence of Fe II emission, often at a significant level. Furthermore, additional lines could be contributing to the underlying continuum (e.g. the He I$\lambda\lambda$4922,5017 doublet; Véron, Gonçalves, and Véron-Cetty, 2002; Zamfir et al., 2010).

In Sec. 1.4.3 we use the whole of the Hβ profile to derive an un-biased BH mass. If, instead, the FWHM is calculated from the narrower of the two Gaussian components rather than the composite profile, then the Hβ FWHM decreases by $630\,\mathrm{km\,s^{-1}}$ on average. This effectively removes the average offset to broad Hβ profiles evident in Fig. 1.9. This will enhance the C IV FWHM relative to the Hα/Hβ FHWM by $\sim$ 15 per cent and increase the size of the correction which must be applied to the C IV-based BH masses by $\sim$ 30 per cent.
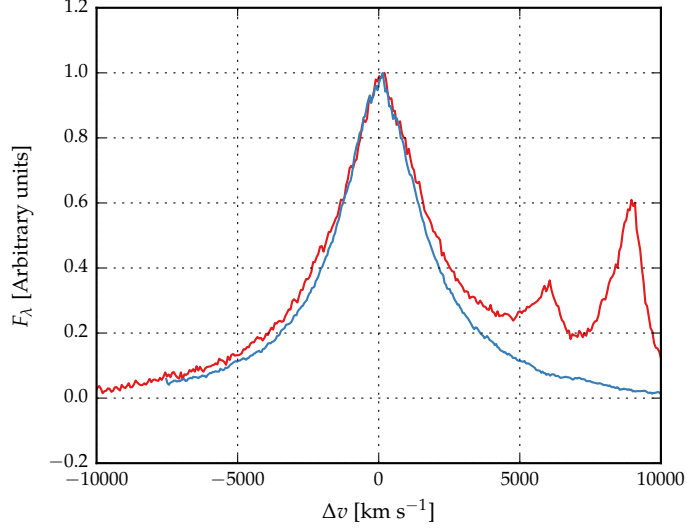
Figure 1.11: The Hα (blue) and Hβ (red) emission-line regions in the median composite spectrum, shown as function of the velocity shift from the respective predicted line peak wavelengths. The background continuum and optical Fe II emission has been modelled and subtracted. The line fluxes have been scaled in order for the profile shapes to be readily compared.

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 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 Hα and Hβ samples respectively narrow [O III] emission is also detected with sufficient S/N to measure the line properties. In the model fit to the Hβ region the velocity centroids of the broad Hβ-line and the core component of the [O III] emission were deliberately determined separately. We find the intrinsic difference in the velocity centroids of the Hα and Hβ emission and the narrow [O III] emission to have a dispersion of 300 and $400\,\mathrm{km\,s^{-1}}$, which is very similar to the value found by Shen et al., (2016). However, the median velocity centroid of the narrow component of the [O III] emission is blueshifted by $250\,\mathrm{km\,s^{-1}}$ 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 $L_{Bol} \sim 10^{44}\,erg\,s^{-1}$, the centroids of the broad component of Hβ 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 Hα and Hβ were modelled with up to two Gaussians, with identical velocity centroids. If there is any significant asymmetry in these lines then the emission will be poorly fit by our model and the redshift derived from the peak of the best-fitting model could be biased. To investigate this possibility we relaxed the requirement for the centroids of the two broad Gaussians to be the same, and measured the systemic redshift from the peak of the composite profile. With this new, more flexible model, the mean absolute difference between the centroids of the two Gaussian components used to model Hα and Hβ was 480 and 780 $km\,s^{-1}$ respectively. With these adjustments, we found the mean difference between the [O III]- and Hα(Hβ) based redshift estimates to be $-100(-120)\,km\,s^{-1}$ and the scatter to be $290(320)\,km\,s^{-1}$. Therefore, the shift between the Balmer and [O III] velocities is reduced, suggesting that there might be a $\sim 100\,km\,s^{-1}$ systematic bias in our measurements of the quasar systemic redshift. 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.

Later, in Section XX, we demonstrate how improvements in the estimation of systemic redshifts from ultraviolet quasar spectra means that it is now possible to quantify the distribution of C IV-blueshifts in the observed population as a whole. Clearly, this is a crucial development in making a blueshift-based correction viable.

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 Hα/Hβ line widths as a function of the C IV blueshift. Because virial BH mass estimates are generally based on the Hβ FWHM, we first convert our Hα FWHM measurements to equivalent Hβ FWHM using Eq. 1.3. In Figs. 1.12a and 1.13a we show the

C IV FWHM relative to both the (Hβ-scaled) Hα FWHM and the Hβ 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 Hα and Hβ 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 ∼ 0.7 data points) and so did not include such a component in our model.

In Fig. 1.14 we show the one- and two-dimensional projections of the posterior distribution from the linear fit to the FWHM C IV/Hα ratio. The projections from the FWHM C IV/Hβ fit, which we do not show, have very similar appearances. In Fig. 1.12a we plot the maximum likelihood model and the 2σ uncertainties on the model parameters. The maximum likelihood line is given by

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

(1.4)

for the C IV/Hα fit and

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

(1.5)

for the C IV/Hβ 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 ± 0.02 and 0.25 ± 0.02 for the Hα and Hβ fits respectively. The intrinsic scatter for the Hα fit is represented by the Normal probability density distribution shown in Fig. 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
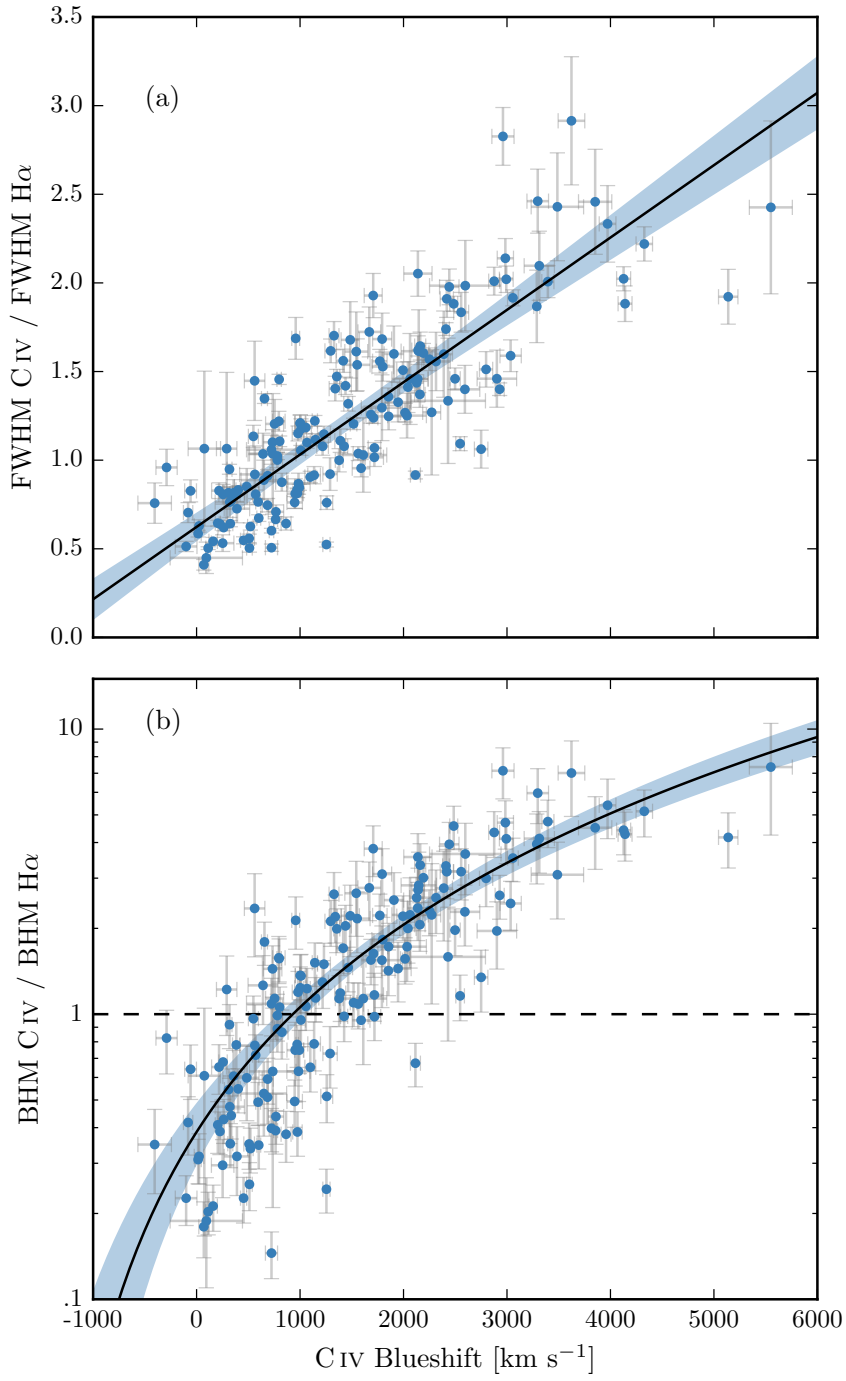
Figure 1.12: C IV FWHM relative to Hα FWHM (a), and C IV based BH mass (BHM) compared to Hα based mass (b), both as a function of the C IV blueshift. The black line is our best-fit linear model, and the shaded region shows the 2-σ uncertainties on the slope and intercept. The Hα FWHM have been scaled to match the Hβ FWHM using Eq. 1.3.
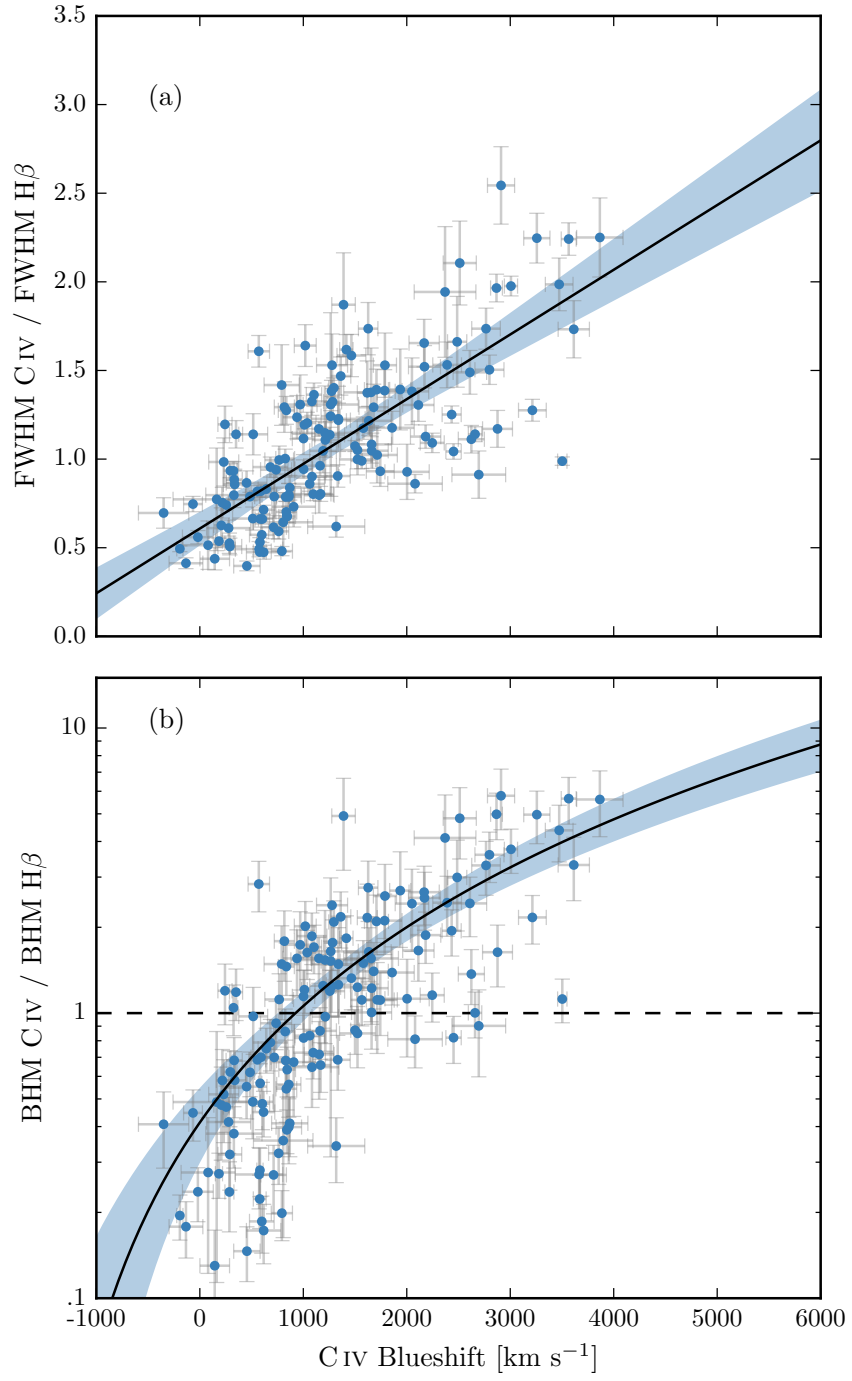
Figure 1.13: C IV FWHM relative to Hβ FWHM (a), and C IV based BH mass (BHM) compared to Hβ based mass (b), both as a function of the C IV blueshift.

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 Hβ (0.12 dex) than when compared to Hα (0.10 dex). This is likely due, at least in part, to the generally higher S/N of the Hα emission. In addition, contributions from the strong [O III] doublet in the vicinity of Hβ make de-blending the Hβ emission more uncertain. As a consequence, for quasars where Hα and Hβ are both measured, the mean uncertainty on the Hα FWHM is $130 \, \mathrm{km \, s^{-1}}$, compared to $340 \, \mathrm{km \, s^{-1}}$ for Hβ.

In the next Section, we use both the Hα and Hβ lines to calculate unbiased BH masses. However, we use the Hα measurements to derive an empirical C IV blueshift based correction to the C IV masses (Eq. 1.6) because of the issues related to the accurate modelling of the Hβ-profile just described. An extra advantage, which is evident in Figs. 1.12a and 1.13a, is that the Hα 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 Hβ 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 Vester-gaard and Peterson, (2006) C IV FWHM calibration uses the monochromatic continuum luminosity at 1350 Å to predict the BLR radius and corresponds to (a = 6.66, b = 2, c = 0.53) in Eq. **??**. For the Hβ calibration, Vestergaard and Peterson, (2006) use the monochromatic continuum luminosity at 5100 Å 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 1.2, and we convert our Hα emission-line velocity-width measures to predicted Hβ widths using Eq. 1.3.

In Figs. 1.12b and 1.13b the C IV-based estimates are compared to the Hα/Hβ 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
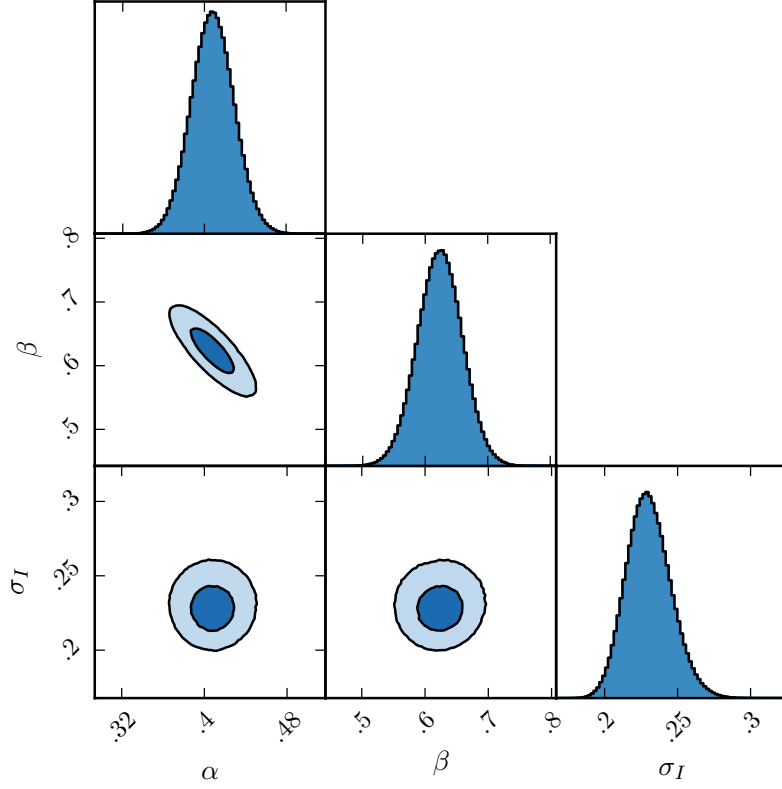
Figure 1.14: One- and two-dimensional projections of the MCMC sample of the posterior distribution for a linear fit to the FWHM C IV/Hα ratio as a function of the C IV blueshift. In the two-dimensional projections we show 1- and 2-σ contours. The posterior distribution for the linear fit to the FWHM C IV/Hβ ratio, which we do not show, has a very similar appearance.

emission-based BH-masses are in error by a factor of more than five at $3000 \, \mathrm{km \, s^{-1}}$ in C IV emission blueshift and the overestimate of the BH-masses reaches a factor of 10 for quasars exhibiting the most extreme blueshifts, $\gtrsim 5000 \, \mathrm{km \, s^{-1}}$.

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 Eq. 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 be-
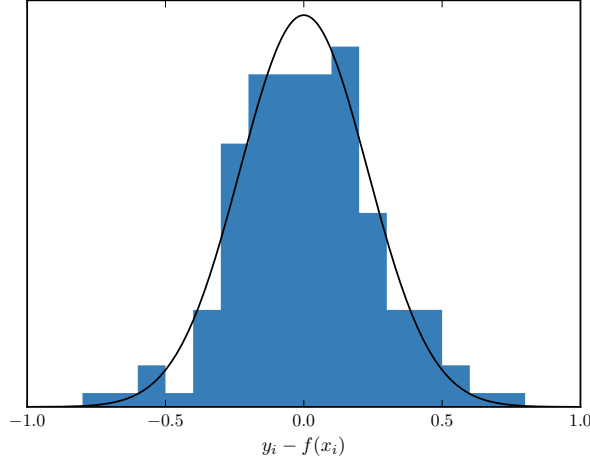
Figure 1.15: The distribution of the orthogonal displacement of each data point from the best-fitting linear relationship in the fit to FWHM(C IV)/FWHM(Hα) as a function of the C IV blueshift (blue histogram). The black curve is a Normal distribution with a width equal to the intrinsic scatter in the 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.

tween the virial products and the Hα-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

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

(1.6)

Given measured C IV emission-line FWHM and blueshift, equations 4 and 6 can then be used to provide an unbiased estimate of the quasar BH mass.

give equation references

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

Might work better later in discussion.

In this Section, we consider why the C IV based masses of quasars with modest C IV blueshifts ($\lesssim 1000\,\mathrm{km\,s^{-1}}$) are systematically underestimated relative to masses derived from the Balmer lines (Figs. 1.12b and 1.13b).

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 Hα and Hβ (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 Hβ/Hα emitting region. Given the $\Delta V \propto R_{BLR}^{-0.5}$ virial relation, this leads to the prediction that the C IV line widths should be $\simeq 1.4$ times broader than Hα for a given BH mass. More recently, Denney, (2012) found that there is a significant contribution from gas at larger radii to the C IV emission-line, enhancing the profile at lower-velocity and leading to smaller FWHM or dispersion values. The ratio of the line widths is therefore predicted to be lower than the factor of $\simeq 1.4$.

The Hα and C IV FWHM of the 77 quasars with C IV blueshifts $< 1200\,\mathrm{km\,s^{-1}}$ are linearly correlated, as expected if the dynamics of the BLR clouds are dominated by virial motions. The median C IV/Hα FWHM ratio is 0.97 with standard deviation 0.31. Thus, as predicted by considering the contribution from low-velocity gas at large radii, the FWHM-based comparison results in a systematically lower median C IV/Hα.

As a direct consequence of the empirically small C IV/Hα FWHM ratio, the C IV-derived BH mass estimates are systematically lower than the corresponding Hα-derived masses when the blueshift is small. This can be seen in Fig 1.12b, where for almost every quasar with a C IV blueshift $<1200\,\mathrm{km\,s^{-1}}$, the C IV-derived BH mass is smaller than the corresponding Hα-derived mass. The median fractional difference between the two estimates is 0.60.

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*

Measuring the systemic redshift

equation references wrong

Equations 4 and 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 1.5.1. 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 μm 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 Fig. 1.16a, in which we plot the C IV-blueshift versus C IV-emission equivalent width (EW) using the SDSS pipeline redshifts and the blueshifts calculated by Shen et al., (2011). A strong trend in the blueshift values as a function of line EW is not evident in Fig. 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 Figs. 1.16a and 1.16b; the appearance in Fig. 1.16b of the extension to high blueshift for quasars with low C IV EW is particularly evident.

Shen et al., (2016) and our own work shows that there is an intrinsic variation of $\sigma \simeq 220\,\mathrm{km\,s^{-1}}$ in the velocity centroids of the broad-line region relative to a systemic-frame defined by the quasar narrow-line regions. The redshifts for quasars in the SDSS DR10 and DR12 catalogues (Pâris et al., 2014; Pâris et al., 2017) possess errors of $\simeq 500 - 750\,\mathrm{km\,s^{-1}}$ (Pâris et al., 2012;
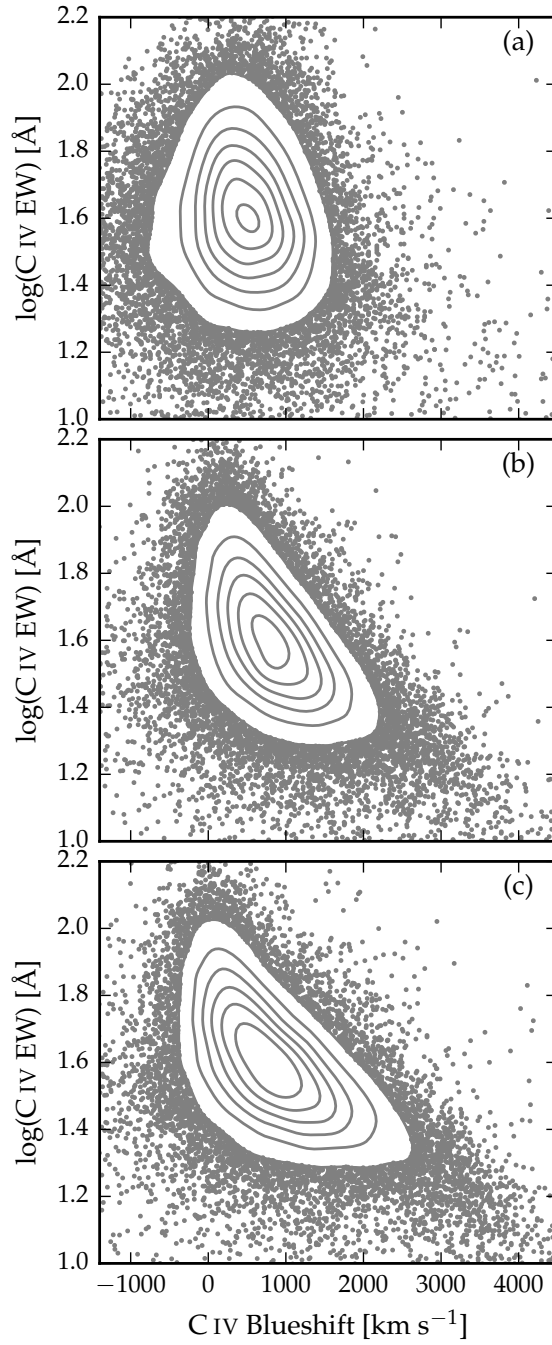
Figure 1.16: Rest-frame EW versus blueshift of the broad C IV emission-line for $32,157$ SDSS DR7 quasars at $1.6 < z < 3.0$. Panel (a) uses C IV line parameters from Shen et al., (2011) and SDSS pipeline systemic redshifts. Panels (b) and (c) use systemic redshifts from Hewett and Wild, (2010) and Allen & Hewett (2017, in preparation) respectively, and C IV line measurements described in Sec. 1.5.1.

Table 1.4: The fractional error on the corrected BH mass as a function of C IV blueshift for different uncertainties in the quasar systemic redshift.

δv (km s^{-1})	C IV blueshift (km s^{-1})			
	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

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 1.4 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^{-1} C IV blueshift when there is a 500 km s^{-1} 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 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 IIIλ1892 and C IIIλ1908 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 EW and modest blueshifts have relatively small ($\simeq 300$ km s^{-1}) SED-dependent redshift errors. Redshift uncertainties as large as $\simeq 1000$ km s^{-1} for such quasars are unusual and the large relative error in the corrected C IV BH-mass given in Table 1.4 is pessimistic.

Conversely, systematic redshift errors are greatest for quasars with large blueshifts, reaching ~ 750 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 MBH(C IV)/MBH(Balmer)

relation at large C IV blueshift (see Figs. 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 (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\lambda$2796,2803 down to the O Iλ1304+Si IIλ1307 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 Hβ and Hα. SED-dependent systematic errors are below the apparent inherent dispersion of $\simeq 220\,\text{km}\,\text{s}^{-1}$ 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 Fig. 1.16b, in particular the extension to high blueshift at low C IV EW, become more apparent in Fig. 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 EW is also apparently still present.

C IV emission-line blueshift measurements

The differences in the distribution of C IV emission-line properties seen in the three panels of Fig. 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, α, determined using the median[4] 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 2σ 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.

Allen & Hewett will publish improved redshifts for all quasars in the SDSS DR7 and DR12 catalogues. At the same time we will publish catalogues of unbiased BH masses for both

---

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

SDSS DR7 and DR12 based on the Allen & Hewett redshifts. The components from the mean-field independent component analysis (see Allen et al., 2013, for an application to astronomical spectra) used in the Allen & Hewett redshift algorithm will also be published. With these components, if a rest-frame ultraviolet spectrum is available, it will be straightforward to determine the systemic redshift, via a simple optimisation procedure, and hence calculate the C IV blueshift.

1.5.2 *Systematic trends in residuals*

The scatter about the best-fitting line in the C IV/Hα FWHM versus C IV-blueshift relation is ~ 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] equivalent width (EW), and Fe II EW. The only significant correlation we find is with the Hα FWHM (Fig. 1.17). Quasars with broad Hα lines tend to lie below the relation while quasars with narrow Hα tend to lie above it. One possibility is that this correlation is simply due to random scatter (either intrinsic or measurement error) in the Hα FWHM which, with the other quasar properties fixed, would naturally produce a correlation between FWHM(C IV)/FWHM(Hα) and FWHM(Hα). However, the fact that we see no such correlation between the model residuals and the C IV FWHM suggests that the Hα FWHM correlation could be revealing something more fundamental. The Hα/Hβ 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). While a number of parameters have been considered within the EV1 context (e.g. Brotherton and Francis, 1999), Fig. 1.17 suggests that part of the scatter between the Balmer and C IV velocity widths might be attributed to differences in the spectral properties which are correlated with EV1 (Marziani et al., 2013).

The shape of the line can be characterised by the ratio (FWHM/σ, where σ is the dispersion, derived from the second moment velocity; e.g. Kollatschny and Zetzl, 2011; Kollatschny and Zetzl, 2013). FWHM/$\sigma \simeq 2.35$ for a Gaussian pro-

file, while FWHM/$\sigma \simeq 1$ for a peakier Lorentzian profile[5]. In our sample, we find the residuals and the Hα FWHM correlate with the shape of the line. The narrow lines are, on average, 'peakier' (with FWHM/$\sigma \simeq 1$) than the broader lines (with FWHM/$\sigma \simeq 2$). The origin of the Balmer-line shape correlation is not clear but one possibility is an orientation-dependence of the Hα FWHM (e.g. Shen and Ho, 2014). In this scenario quasars with broader emission-lines are more likely to be in an edge-on orientation relative to our 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- quasars and six lobe-dominated quasars were identified in our sample, but no statistically significant differences in the Hα 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).

Maybe include: On the other hand, a concern is raised by the observation that the BLR structure changes with emission-line FWHM (Kollatschny & Zetzl 2011). They show that FWHM/σ is a strong, smooth function of FWHM. These changes in line profile shapes suggest that the balance between rotation, random velocities, and in- or outflow changes with line width. Hints of this were seen in earlier work (Collin et al. 2006; Marziani & Sulentic 2012).

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.

---

5 Strictly FWHM/$\sigma \rightarrow 0$ for a Lorentzian profile, but values close to unity are typical when the dispersion is calculated over a velocity range, $\simeq \pm 10\,000\,\mathrm{km\,s^{-1}}$, used to parametrize broad emission-lines in quasar spectra.
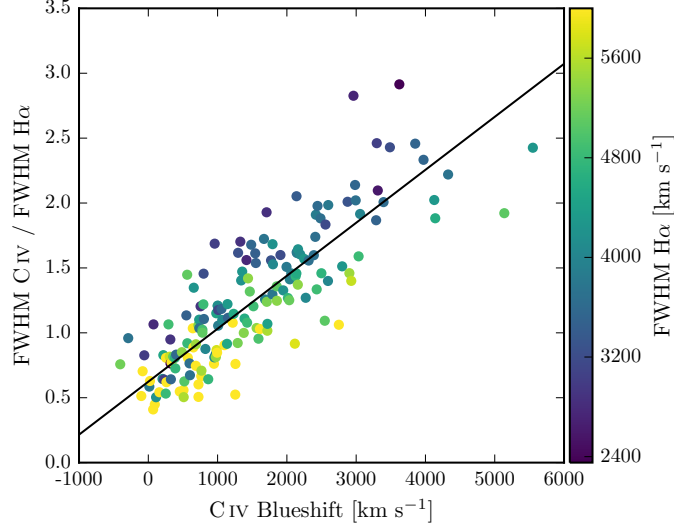
Figure 1.17: Same as Fig. 1.12a, with the marker colour representing the Hα FWHM. At fixed C ɪᴠ blueshift, there is a clear Hα FWHM dependent systematic in the model residuals.

1.5.3 *Effectiveness of the C ɪᴠ blueshift based correction to BH masses*

Figure 1.18 demonstrates that our sample has an excellent coverage of the EW-blueshift parameter space in relation to SDSS DR7 quasars at redshifts $1.6 < z < 3.0$. The systematic offset to higher C ɪᴠ 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 (Fig. **??**). Our sample includes 21 quasars with C ɪᴠ blueshifts $> 3000\,\mathrm{km\,s^{-1}}$, and extends to $\sim 5000\,\mathrm{km\,s^{-1}}$, i.e. at the very extreme of what is observed in this redshift and luminosity range. Our investigation thus demonstrates that the C ɪᴠ-blueshift based correction derived in this chapter is applicable to very high blueshifts. Conversely, there are no quasars in our catalogue with C ɪᴠ blueshifts $\lesssim 0\,\mathrm{km\,s^{-1}}$ and we caution against extrapolating the correction formula to negative blueshifts.

Figure 1.19 compares the C ɪᴠ- and Hα-based BH masses before and after applying the blueshift-based correction to the C ɪᴠ FWHM. Before the correction, the correlation between the C ɪᴠ- and Hα-based BH masses is very weak, and the scatter between the masses is 0.4 dex. After correcting the C ɪᴠ FWHM for the non-virial contribution, the correlation improves dramatically. The scatter between the corrected C ɪᴠ-based masses
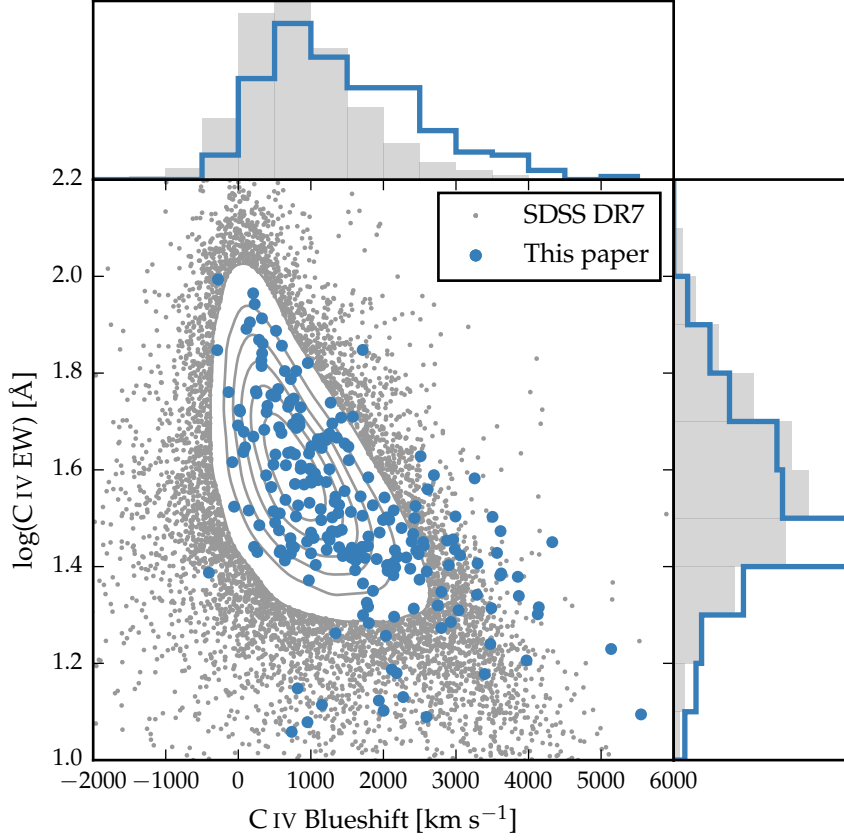
Figure 1.18: Rest-frame EW versus blueshift of the broad C IV emission-line for 32,157 SDSS DR7 quasars at 1.6 < z < 3.0 (*grey*) and our sample (*blue*). For the SDSS quasars, the systemic redshifts used to calculate the blueshifts are from Hewett and Wild, (2010) and C IV emission properties are described in Paper I. Our sample has very good coverage; 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.
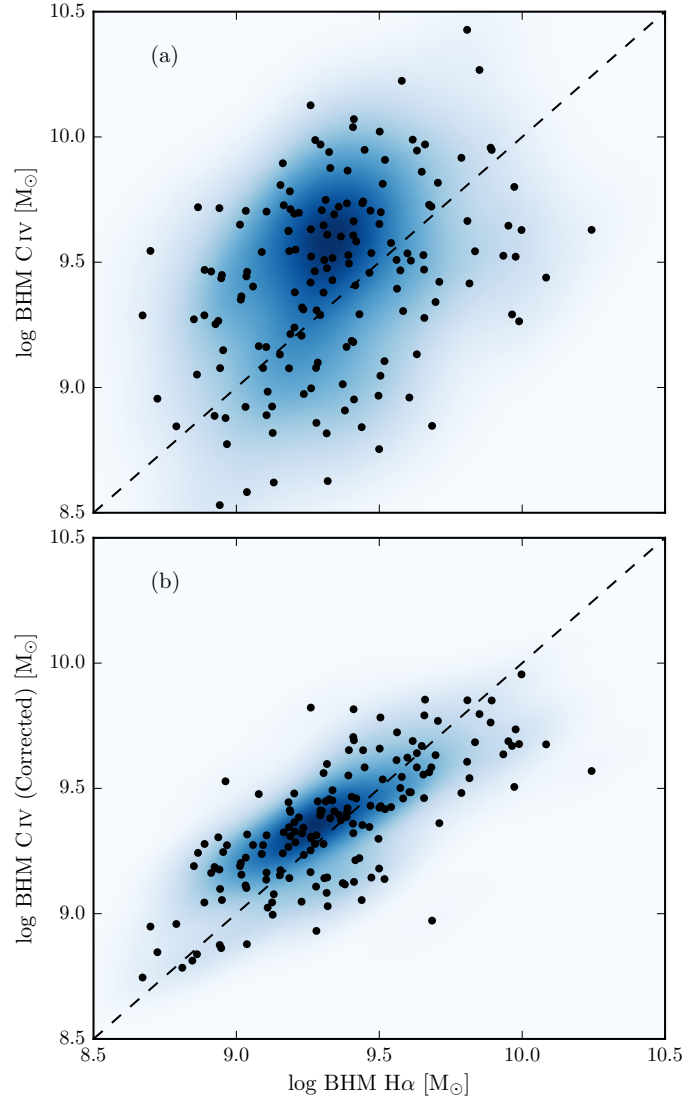
Figure 1.19: Comparison of the C IV- and Hα-based BH masses be-
fore (a) and after (b) applying the C IV blueshift-based
correction to the C IV FWHM. The density of the plot-
ted points (estimated using a Gaussian kernel density es-
timator) is represented by the colour. The correction to
the C IV BH masses decreases the scatter by from 0.4 to
0.2 dex. Should definitely include some empirical valida-
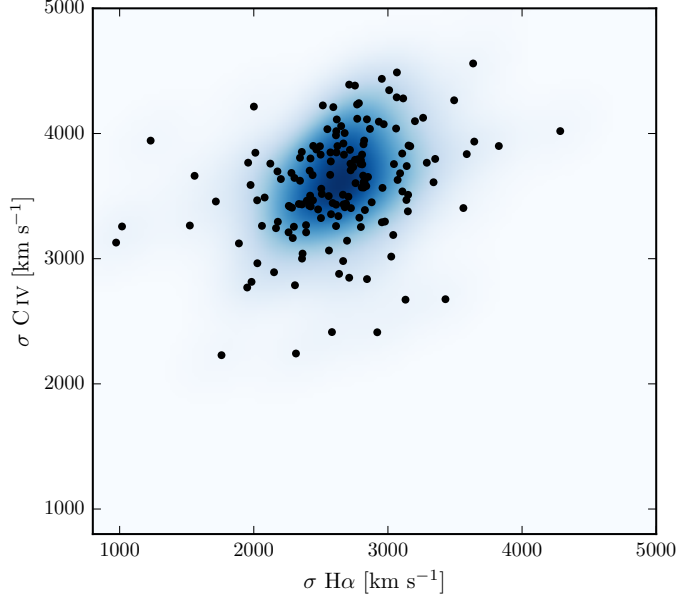tion of ICA redshifts since that is what we are telling
people to use.

Figure 1.20: Comparison of the C IV and Hα line dispersion, σ. 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 Hα measurements by a factor of 1.4 on average, which is consistent with reverberation mapping measurements (Vestergaard and Peterson, 2006).

and the Hα-based masses is reduced to 0.2 dex. The scatter is 0.24 dex at low C IV blueshifts ($\sim 0\,\mathrm{km\,s}^{-1}$) and 0.10 dex at high blueshifts ($\sim 3000\,\mathrm{km\,s}^{-1}$).

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 Fig. 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 (Fig. 1.20).

The correlation between the Hα and C IV line dispersion is, however, weak. The Pearson coefficient for the correlation is 0.36 (and just 0.15 when the Hβ measurements are used in place of Hα). Furthermore, there is little dynamic range in the

line dispersion: the scatter is just 480 and 460 km s^{-1} for Hα 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 substantially more effective than using the C IV line dispersion with, or without, the line blueshift.

1.5.4 *Comparison to previous prescriptions*

In Fig. 1.21 we compare the C IV blueshift-based correction presented in this chapter to 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 Hα-based masses as a function of the C IV blueshift. The correction proposed by Runnoe et al., (2013a) is based on the spectral region at rest-frame wavelengths of ~1400 Å (see below). Therefore, our analysis is based on the 123 quasars with spectra covering this region.

In Fig 1.21a the C IV BH masses have been corrected using the C IV shape (FWHM/σ) 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/σ) of the C IV profile. In our sample, we observe a strong correlation between the shape of the C IV line and its blueshift (Fig. 1.7c); between the two extremes in the C IV blueshift distribution the line shape changes from FWHM/σ ~ 1 to 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 Hα- 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
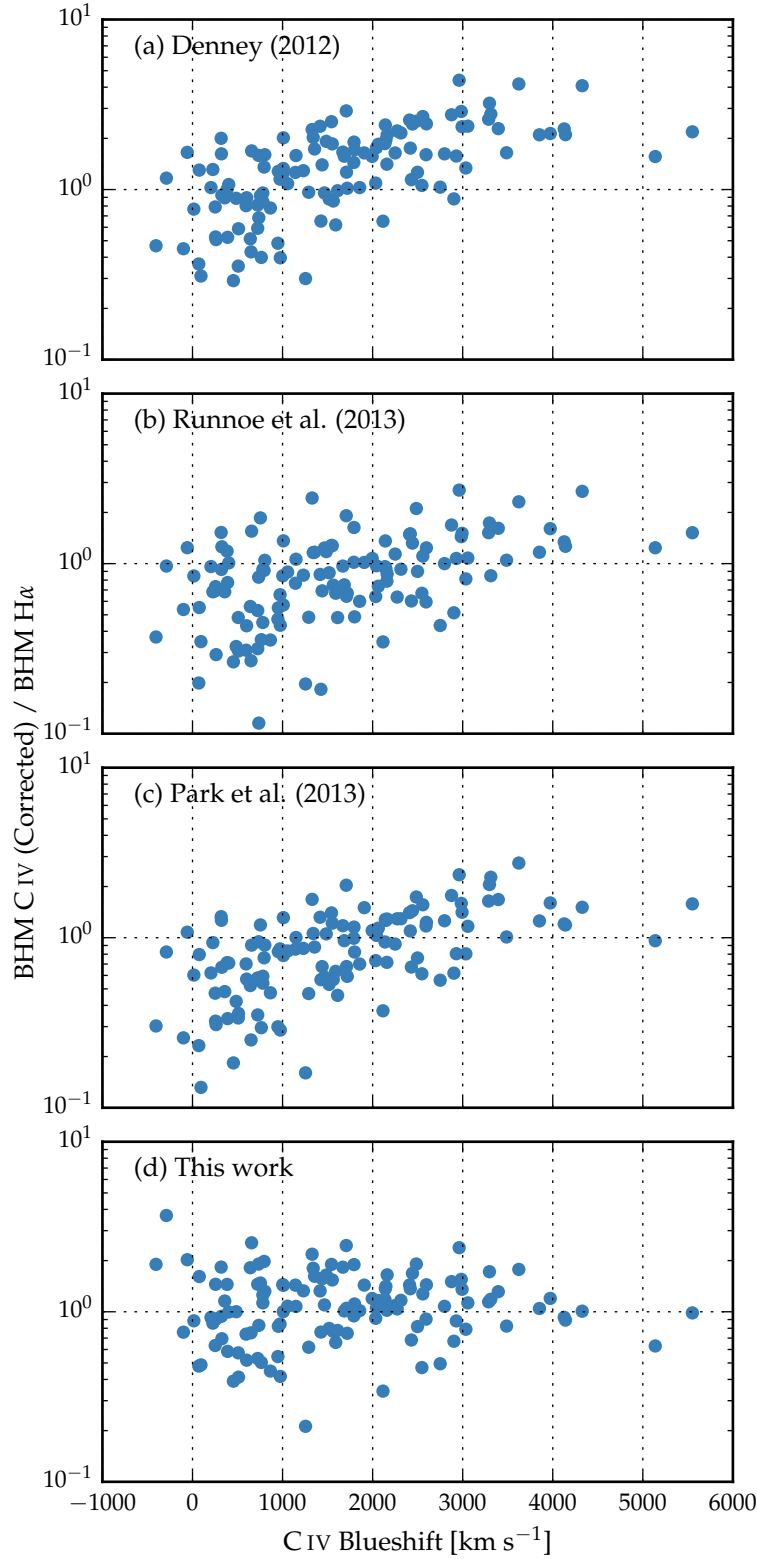
Figure 1.21: Comparison of BH mass estimates derived from C IV and Hα as a function of the C IV blueshift. Corrections to the C IV-based masses have been applied based on the shape (FWHM/σ) of the C IV emission-line (a; Denney, 2012), the peak flux ratio of the Si IV+O IV blend relative to C IV (b; Runnoe et al., 2013a), by significantly reducing the dependence of the derived BH mass on the C IV velocity-width (c; Park et al., 2013), and based on the C IV blueshift (d; this chapter).

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., (2013a) 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 velocity width. This parameter was chosen because it showed the strongest correlation with the FWHM C IV/Hβ residuals, as well as with the strengths of optical O III and Fe II.

Following Runnoe et al., (2013a), 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 Fig. 1.18, a correlation exists between the blueshift and equivalent width 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 ∼ 2 (Fig. 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 (Fig. 1.21c). However, the systematic error in the C IV-based BH masses as a function of C IV blueshift remains.

As a comparison, the C IV-based masses shown in Fig 1.21d 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 Fig. 1.22, there are systematic variations in the Hα line profile as a function of the C IV blueshift. At C IV-blueshift $< 1200 \, \text{km s}^{-1}$, the Hα 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 Hα FWHM $= 3500 \, \text{km s}^{-1}$, with a scatter of just $700 \, \text{km s}^{-1}$. The
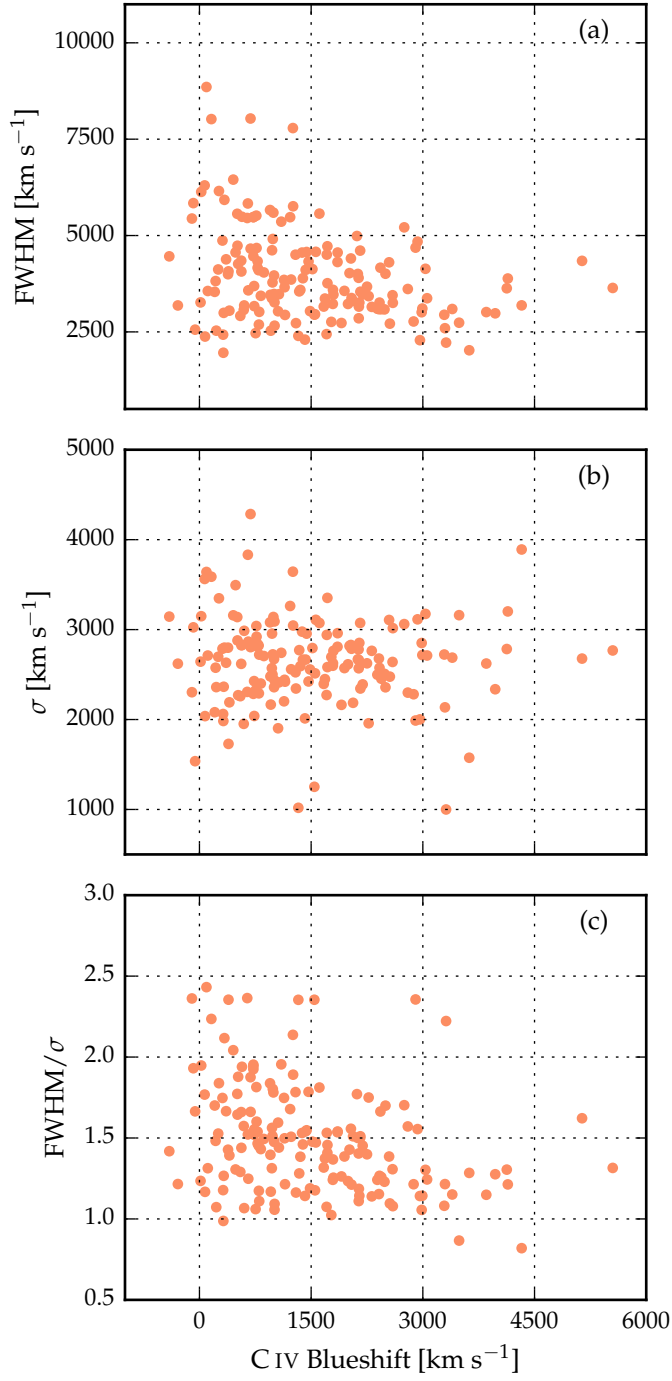
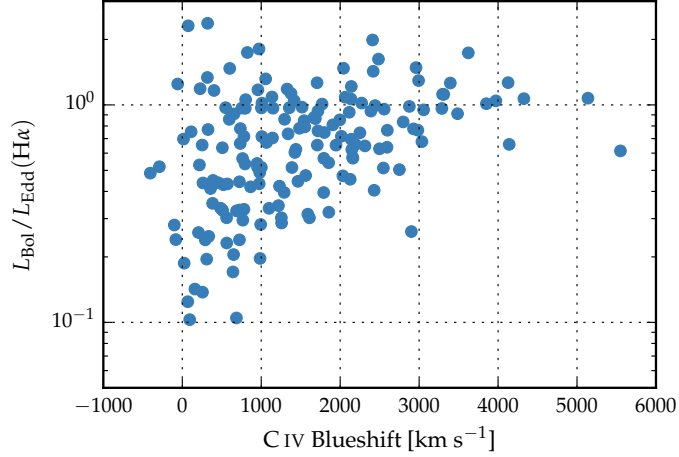Figure 1.22: The FWHM, dispersion (σ) and shape (FWHM/σ) of Hα as a function of the C IV blueshift.

Figure 1.23: Hα-derived Eddington ratio versus C IV blueshift. At blueshift $\gtrsim 2000$ km s^{-1} all quasars have high accretion rates ($L/L_{Edd} \simeq 1$). 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.

apparent trend of peakier Hα-emission, with FWHM/σ close to unity, at large C IV-blueshift is enhanced by the modest increase in Hα EW with blueshift. Amongst the low-C IV-blueshift population there are in addition quasars with broader and more Gaussian-like Hα line profiles, with FWHM/σ $\simeq 2$.

The change in the Hα emission-line profiles as a function of C IV-blueshift means that the Hα-FWHM derived BH masses 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) ($L_{bol} = 9.26 L_{5100}$). 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 Fig. 1.23, demonstrate that at large blueshifts quasars are accreting at around their Eddington limits (Fig. 1.23). 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 Hα 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 EW 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).

1.6.1 *Systematic biases in Balmer-based masses*

The interpretation described in the preceding Section requires some caution since the emission-line shape (characterized by the value of FWHM/σ) of Hα is also changing as a function of the C IV blueshift (Fig. 1.22c). At low C IV blueshifts there are a range of shapes, but all of the quasars exhibiting large C IV blueshifts have peaky Hα profiles with FWHM/σ $\simeq$ 1. This raises the question of whether the Hα FWHM is a reliable proxy for the virial-induced velocity dispersion for the full range of Hα line shapes we have in our sample.

When calibrating the virial-product to masses derived independently using the BH mass – stellar velocity dispersion ($M_\odot - \sigma$) relation, Collin et al., (2006) find that the scaling fac-

Can move most of this to earlier Section on residuals and then refer back to that.

tor, f, is a factor ~ 2 larger for their Population '1' sources (with FWHM/σ < 2.35 and essentially equivalent to population A of Sulentic and co-workers and to the high-blueshift quasars here) than for their Population 2 (with FWHM/σ > 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. In the context of this result from Collin et al., (2006), our high-blueshift objects all possess peaky Hα-lines and, while our quasar sample probes much higher luminosities and masses, the true BH-masses may also be underestimated. Adopting such an interpretation, the amplitude of the trend seen in Figs. 1.12 and 1.13 might not be so pronounced.

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. There are currently very few reverberation-mapping measurements of quasars with large C IV blueshifts but the results of the large on-going statistical reverberation mapping projects (e.g. Shen et al., 2015) for luminous quasars at high-redshift will go some way to establishing whether the quasar broad line regions producing Balmer emission look the same for objects with very different C IV-emission blueshifts.

Although the EV1-trends (Sulentic et al., 2000; Shen and Ho, 2014) are most likely driven by the accretion rate, orientation may also have a role to play in determining the observed properties of the BLR. Shen and Ho, (2014) argue that a large part of the scatter observed in the Hβ FWHM relates not to a spread in BH masses, but rather to the orientation of the BLR relative to the line-of-sight. For this to be true, the BLR would need to be in a flattened disc-like geometry, in which case the observed line width would increase with the inclination of the disc relative to the line of sight. Brotherton, Singh, and Runnoe, (2015) found that the core-dominance of radio-loud quasars, which is believed to be a reliable proxy for orientation, at least in a statistical sense, is significantly correlated with the Hβ FWHM and hence with the BH-mass estimates. This raises the question of whether the narrow Hα 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). Further-

more, Leighly, (2004) showed that the He IIλ1640 emission-line properties of quasars with large C IV blueshifts are more consistent with differences in the SED rather than differences in the orientation. Collin et al., (2006) showed that orientation effects were also sub-dominant to the Eddington ratio in determining the shape of the Hβ line and the Hα line shape trend we observe is consistent with the finding of Marziani et al., (2003) that the Hβ emission profiles of high/low Eddington ratio low-z quasars and type 1 Seyfert nuclei are well fit by Lorentzian and double Gaussian profiles respectively. Overall, therefore, orientation does not appear to be the dominant effect in determining the C IV blueshift and correlated changes in the Hα line profile.

1.6.2 *The BAL parent population*

Classical high-ionization BAL (HiBAL) quasars are also predominantly Population A objects in the scheme of Sulentic 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_{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-EW and narrow-Hα 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 (see Runnoe et al., 2013b, for such a study).

Do I actually say BALs are removed from sample?

1.6.3 *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,... 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

I must have a few BALs. Make a composite spectra of Balmer lines of BALs?

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 $> 1200 \, \mathrm{km \, s^{-1}}$ have enhanced accretion rates relative to the 'normal' population (with C IV blueshifts $< 1200 \, \mathrm{km \, s^{-1}}$). 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 Fig. 1.23), 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[6] 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 $< 1200 \, \mathrm{km \, s^{-1}}$ and 2,484 with blueshifts $> 1200 \, \mathrm{km \, s^{-1}}$ 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. ~ 15 per cent of the ultraviolet-selected non-BAL quasar population).

1.7 CONCLUSIONS

The main results of this chapter are as follows:

- We have analysed the spectra of 230 high-luminosity ($10^{45.5} - 10^{48} \, \mathrm{erg \, s^{-1}}$), 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

---

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

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 $3000 \, \mathrm{km \, s^{-1}}$ in C IV emission blueshift and the overestimate of the BH-masses reaches a factor of 10 for quasars exhibiting the most extreme blueshifts, $\gtrsim 5000 \, \mathrm{km \, s^{-1}}$.

- 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 4 and 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 Hα-based masses is 0.24 dex at low C IV blueshifts ($\sim 0 \, \mathrm{km \, s^{-1}}$) and 0.10 dex at high blueshifts ($\sim 3000 \, \mathrm{km \, s^{-1}}$). This is a significant improvement on the 0.40 dex scatter observed between the un-corrected C IV and Hα 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 $z \sim 5$.

equation references

1.8 FUTURE WORK

- Clustering

- Data-driven mapping

1.9 DESCRIPTION OF CATALOGUE

2 Unique ID: QSOXXX.

LogMBH_Ha, LogMBH_Ha_Err

Edd_Ratio_Ha, Edd_Ratio_Ha_Err

LogMBH_Hb, LogMBH_Hb_Err

Edd_Ratio_Hb, Edd_Ratio_Hb_Err

LogMBH_CIV_VP06, LogMBH_CIV_VP06_Err

Edd_Ratio_CIV_VP06, Edd_Ratio_CIV_VP06_Err

Add corrected BH mass

FWHM_Broad_Ha, FWHM_Broad_Ha_Err

Sigma_Broad_Ha, Sigma_Broad_Ha_Err

EQW_Broad_Ha, EQW_Broad_Ha_Err

LogL_Broad_Ha, LogL_Broad_Ha_Err

FWHM_Broad_Hb, FWHM_Broad_Hb_Err

Sigma_Broad_Hb, Sigma_Broad_Hb_Err

EQW_Broad_Hb, EQW_Broad_Hb_Err

LogL_Broad_Hb, LogL_Broad_Hb_Err

z_Broad_Ha. Repeated in chapter 4 with slightly different model.

z_Broad_Hb. Repeated in chapter 4 with slightly different model.

WARN_Ha (Remove 2, just flag 1 as low S/N)

RedChi_Ha

WARN_Hb (Remove 2, just flag 1 as low S/N)

RedChi_Hb

FWHM_CIV_BEST, FWHM_CIV_BEST_Err (remove best, make sure it's clear where it comes from) (Just give BEST values then note of which optical spectrum was used).

Sigma_CIV_BEST, Sigma_CIV_BEST_Err

Median_CIV_BEST, Median_CIV_BEST_Err

EQW_CIV_BEST, EQW_CIV_BEST_Err

LogL_CIV_BEST, LogL_CIV_BEST_Err

Max_CIV_BEST, Max_CIV_BEST_Err

Max_1400_BEST, Max_1400_BEST_Err

1400_CIV_BEST, 1400_CIV_BEST_ERR

RedChi_CIV_BEST

WARN_CIV_BEST (will have removed warn=2, so should just be warn=1 low S/N flag)

WARN_1400_BEST (remove 1400 max if warn=2)

Blueshift_CIV_Ha, Blueshift_CIV_Ha_Err

Blueshift_CIV_Hb, Blueshift_CIV_Hb_Err