

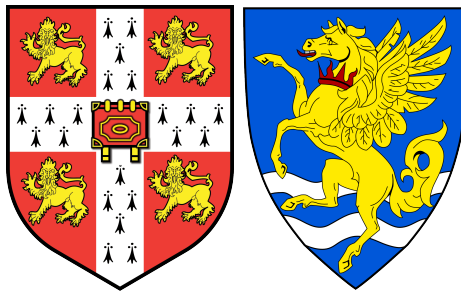
LIAM COATMAN

BLACK HOLE MASSES AND OUTFLOWS IN HIGH
REDSHIFT QUASARS

A DISSERTATION SUBMITTED TO THE
UNIVERSITY OF CAMBRIDGE FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

BLACK HOLE MASSES AND OUTFLOWS IN HIGH REDSHIFT
QUASARS

LIAM COATMAN
UNIVERSITY OF CAMBRIDGE
INSTITUTE OF ASTRONOMY
ROBINSON COLLEGE
SUBMITTED TO THE BOARD OF GRADUATE STUDIES APRIL 2017
UNDER THE SUPERVISION OF
PROF. PAUL C. HEWETT
DR. MANDA BANERJI



Liam Coatman: *Black hole masses and outflows in high redshift quasars*,
An Homage to The Elements of Typographic Style, © September 2015

Ohana means family.
Family means nobody gets left behind, or forgotten.
— Lilo & Stitch

Dedicated to the loving memory of Rudolf Miede.
1939–2005

ABSTRACT

Short summary of the contents in English...a great guide by Kent Beck how to write good abstracts can be found here:

<https://plg.uwaterloo.ca/~migod/research/beck00PSLA.html>

ZUSAMMENFASSUNG

Kurze Zusammenfassung des Inhaltes in deutscher Sprache...

PUBLICATIONS

This might come in handy for PhD theses: some ideas and figures have appeared previously in the following publications:

Attention: This requires a separate run of bibtex for your refsection, e.g., `ClassicThesis1-b1x` for this file. You might also use biber as the backend for biblatex. See also <http://tex.stackexchange.com/questions/128196/problem-with-refsection>.

*This is just an early
– and currently
ugly – test!*

ACKNOWLEDGMENTS

Put your acknowledgments here.

CONTENTS

1	INTRODUCTION	1
1.1	The AGN-Host Galaxy Connection	1
1.2	Measuring Black Hole Masses	1
1.2.1	Reverberation Mapping	2
1.2.2	Single-Epoch Virial Estimates	2
1.3	Orientation-Based Unification Models	3
1.4	The Torus	5
1.5	Evolutionary Models	5
1.6	Interesting Sub-Populations	7
1.6.1	Red and Reddened Quasars	7
1.6.2	Broad Absorption Line Quasars	7
1.6.3	Hot-Dust-Poor Quasars	8
1.6.4	Type II Quasars	8
1.7	Spectral Energy Distributions	9
2	A NEAR-INFRARED SPECTROSCOPIC DATABASE OF HIGH-REDSHIFT QUASARS	13
2.1	Coatman et al. (2016) Quasars	16
2.2	Shen & Liu (2012) and Shen (2016) Quasars	17
2.3	Quasar Pairs	18
2.4	VLT SINFONI Quasars	18
2.5	ESO NTT SOFI Quasars	19
2.6	Hale TripleSpec Quasars	20
3	BLACK HOLE MASSES	23
3.1	Abstract	23
3.2	Introduction	24
3.3	C IV blueshifts in the quasar population	27
3.3.1	Quasar systemic redshifts	27
3.3.2	C IV emission line blueshift measurements	29
3.3.3	Sample selection - C IV properties	31
3.3.4	Relation to virial BH mass estimates	31
3.4	Quasar Sample	35
3.4.1	Optical observations	36
3.5	Spectral Measurements	37
3.5.1	C IV	39
3.5.2	H α	43
3.5.3	H β and [O III]	44
3.5.4	Fitting procedure	44
3.5.5	Spectra removed from sample	45
3.5.6	Emission-line parameter uncertainties	48
3.5.7	Contemporaneity of spectra	49

3.5.8	Quasar monochromatic luminosity	50
3.5.9	Characterising the emission-line profiles	51
3.6	An empirical correction to C iv-based virial BH-mass estimates	54
3.6.1	H α /H β FWHM comparison	54
3.6.2	Measuring the quasar systemic redshift	59
3.6.3	Balmer/C iv line widths as a function of C iv-blueshift	60
3.6.4	C iv-derived BH masses at low C iv blueshift	64
3.6.5	C iv based virial BH mass estimates	67
3.7	Practical application of the C iv-based correction to virial BH-mass estimates	71
3.7.1	Recipe for unbiased C iv based BH masses	71
3.7.2	Systematic trends in residuals	72
3.7.3	Effectiveness of the C iv blueshift based correction to BH masses	74
3.7.4	Comparison to previous prescriptions	78
3.7.5	Population trends with C iv blueshift	80
3.8	Discussion	81
3.8.1	Biases in single-epoch C iv-based BH-mass estimates	81
3.8.2	Possible systematic trends in H α BH-mass estimates	83
3.8.3	Accretion-rate trends in the quasar population	85
3.8.4	The BAL parent population	86
3.8.5	The frequency of quasars with high accretion rates	86
3.9	Conclusions	87
4	NARROW LINE REGION PROPERTIES	91
4.1	Introduction	91
4.2	Quasar Sample	92
4.3	Parameteric Model Fits	93
4.3.1	Derived Parameters	95
4.3.2	Other flags	103
4.4	ICA Component Fits	103
4.5	Measuring the quasar systemic redshift	103
4.5.1	H α	106
4.5.2	ICA	106
4.5.3	Parameter uncertainties and upper limits	107
4.5.4	Absolute flux calibration of spectra and continuum luminosities	107
4.6	Results from ICA fits	108
4.7	Luminosity/redshift-evolution of [O III] properties	109
4.8	Equivalent width	109
4.8.1	OIII outflows	111
4.8.2	[O III] and C iv outflows are linked	114

4.9	Eigenvector one correlations	115
4.10	Mapping EV ₁ to CIV blueshift and EQW	116
4.11	Signal to noise tests	116
4.12	Broad Absorption Line Quasars	116
4.13	Discussion	117
4.13.1	Type II quasars	117
5	SED PROPERTIES	119
5.1	Data	119
5.1.1	The Sloan Digital Sky Survey	119
5.1.2	UKIDSS Large Area Survey	120
5.1.3	WISE All-WISE Survey	120
5.1.4	Completeness of Photometry	120
5.1.5	Final Sample	121
5.2	SED Model	121
5.2.1	Accretion Disk	122
5.2.2	Hot Dust	122
5.2.3	Emission Lines	123
5.2.4	Host Galaxy	123
5.2.5	Lyman- α Forest Absorption	125
5.2.6	Lyman-Limit Systems	126
5.2.7	Dust Extinction	126
5.3	The ‘Standard’ SED Model	127
5.4	Discussion of Fit	129
5.4.1	Flux Correction	132
5.5	Hot Dust	132
5.5.1	Parameterising the hot dust emission	133
5.5.2	Sample	133
5.5.3	Diversity in hot dust properties	135
5.6	Fitting procedure	138
5.7	Results	139
5.7.1	Correlations with quasar properties	139
5.7.2	Spectral properties	139
5.7.3	BALs and radio-loud/radio-quiet	143
5.8	Other works	143
5.8.1	Eddington ratio	144
6	CONCLUSIONS / FUTURE WORK	145
6.1	Future: Red quasars	145

LIST OF FIGURES

Figure 1.1	Illustration of the physical structure of an AGN in a simple orientation-based unification model. From Urry and Padovani, (1995). 3
Figure 1.2	Median SEDs for radio-loud and radio-quiet quasars from Shang et al., (2011). 9
Figure 2.1	The ranges in redshift and luminosity covered by our sample, relative to the redshift-luminosity distribution of the SDSS DR7 quasar catalogue. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator. For the SDSS sample we use Hewett and Wild, (2010) redshifts and bolometric luminosities measured by Shen et al., (2011). For the quasars in this paper the redshift is defined using the peak of the $H\alpha/H\beta$ emission and the luminosity is measured in the continuum at 1350\AA and converted to a bolometric quantity using the same conversion factor employed by Shen et al., (2011). Eight objects are missing because we do not have enough information to calculate the bolometric luminosity. 14
Figure 3.1	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 (orange). 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 (2016, in preparation) respectively, and C IV line measurements described in Sec. 3.3.2. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator. The three rectangles in panel (b) show the regions of parameter space used to generate the composite spectra shown in Fig. 3.2. 28

- Figure 3.2 Composite spectra of the C iv-emission line as a function of C iv blueshift for SDSS DR7 quasars. The quasars contributing to each composite are indicated in Fig. 3.18b. 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. 32
- Figure 3.3 2nd-order Gauss-Hermite (GH) polynomial and 3-component Gaussian fit to the C iv-emission line of SDSSJ0858+0152, which is the most blueshifted in our sample. We derive line parameters from the GH polynomial fit; using the Gaussian model changes the FWHM, line dispersion, and blueshift by -250, -150, and 500 km s⁻¹ respectively. For the C iv lines of all other quasars in our sample the GH polynomial and Gaussian models provide equally good fits. 41
- Figure 3.4 Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of (a) the Hβ FWHM with Shen, (2016), (b) the Hβ FWHM with Shen and Liu, (2012), and (c) the C iv FWHM with Shen et al., (2011). 42
- Figure 3.5 Model fits to continuum-subtracted Hα, Hβ, and C iv emission in six quasars, chosen to represent the range of S/N (indicated in the figure and given per 150km 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. 47
- Figure 3.6 The redshift and luminosity distributions of the spectra removed from our Hα/C iv (a, b) and Hβ/C iv (c, d) samples. 48

- Figure 3.7 The FWHM, dispersion (σ) and shape (FWHM/ σ) of C IV and H α as a function of the C IV blueshift. The vertical line demarcates the ‘high’ and ‘low’ C IV blueshift regimes discussed in the text. At high blueshift it is clear that BH masses estimated from the C IV FWHM (as is typically done at the redshifts considered) will be significantly larger than those estimated from the H α FWHM. 52
- Figure 3.8 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. 55
- Figure 3.9 Comparison of H α and H β 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. 56
- Figure 3.10 One- and two-dimensional projections of the MCMC sampling of the posterior distribution from the fit in Fig. 3.9. α is the power-law index, 10^β is the normalisation, and σ_1 is the intrinsic scatter. In the two-dimensional projections, 1- and 2- σ contours are shown. 58
- Figure 3.11 The H α and H β emission line regions in the median composite spectrum, shown as function of the velocity shift from the respective predicted line peak wavelengths. The line fluxes have been scaled in order for the profile shapes to be readily compared. The H α and H β line profiles are very similar, which suggests a tight correlation between the H α and H β line widths. Quasar narrow-line emission from O III $\lambda 5008.2$ is visible but, overall, the O III $\lambda 4960, 5008$ emission is relatively weak in these spectra. 60

- Figure 3.12 C iv FWHM relative to H α FWHM (a), and C iv based BH mass (BHM) compared to H α 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 H α FWHM have been scaled to match the H β FWHM using Eq. 3.3. 62
- Figure 3.13 C iv FWHM relative to H β FWHM (a), and C iv based BH mass (BHM) compared to H β based mass (b), both as a function of the C iv blueshift. 63
- Figure 3.14 One- and two-dimensional projections of the MCMC sample of the posterior distribution for a linear fit to the FWHM C iv/H α 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/H β ratio, which we do not show, has a very similar appearance. 65
- Figure 3.15 The distribution of the orthogonal displacement of each data point from the best-fitting linear relationship in the fit to FWHM(C iv)/FWHM(H α) 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. 66
- Figure 3.16 Comparison of virial BH mass estimates based on the C iv FWHM (*top*) and dispersion σ (*bottom*) and H α FWHM as a function of the C iv blueshift. The horizontal line indicates agreement between C iv and H α BH masses, and the vertical line demarcates the ‘high’ and ‘low’ C iv blueshift regimes discussed in the text. The BH masses of quasars with moderate C iv blueshifts are underestimated when using the C iv FWHM, while the masses of quasars with large blueshifts are severely overestimated. This situation cannot be corrected by changing the exponent on the FWHM (e.g. Rafiee and Hall, 2011; Park et al., 2013) or the overall scaling in standard virial BH mass relations. 68

- Figure 3.17 Same as Fig. ??a, with the marker colour representing the $H\alpha$ FWHM. At fixed C iv blueshift, there is a clear $H\alpha$ FWHM dependent systematic in the model residuals. 74
- Figure 3.18 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. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator. 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. 75
- Figure 3.19 Comparison of the C iv- and $H\alpha$ -based BH masses before (a) and after (b) applying the C iv blueshift-based correction to the C iv FWHM. The density of the plotted points (estimated using a Gaussian kernel density estimator) is represented by the colour. The correction to the C iv BH masses decreases the scatter by from 0.4 to 0.2 dex. 76
- Figure 3.20 Comparison of the C iv and $H\alpha$ 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 $H\alpha$ measurements by a factor of 1.4 on average, which is consistent with reverberation mapping measurements (Vestergaard and Peterson, 2006). 77

- Figure 3.21 Comparison of BH mass estimates derived from C IV and H α 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 paper). 79
- Figure 3.22 H α -derived Eddington ratio versus C IV blueshift. At blueshift $\gtrsim 2000\text{km s}^{-1}$ all quasars have high accretion rates ($L/L_{\text{Edd}} \simeq 1$). This is in agreement with Kratzer and Richards, (2015), but in contrast to what one would derive from naive use of C IV-based BH mass scaling relations. 80
- Figure 3.23 Comparison of BH mass estimates derived from the FWHM of C IV and H α as a function of the C IV blueshift (*black circles*), and after applying corrections to the C IV-derived mass based on the line shape (*blue triangles*; Denney, 2012) and the peak flux ratio of the Si IV+O IV blend relative to C IV (*red squares*; Runnoe et al., 2013a). While the shape-based correction improves the consistency between BH mass estimates in the low-blueshift population, only the Si IV+O IV/C IV peak flux-based correction is effective at high blueshifts (although a weak systematic remains). 84
- Figure 4.1 The ranges in redshift and luminosity covered by our sample, relative to the redshift-luminosity distribution of the SDSS DR7 quasar catalogue with measured [O III] line properties (Shen et al., 2011). The gaps in our sample coverage at $z \sim 1.8$ and $z \sim 3$ are due to the near-infrared transmission windows. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator. 92

- Figure 4.2 Multi-component Gaussian fits to the continuum-subtracted $H\beta/[O\text{III}]$ emission in 16 quasars, chosen to be the representative of the wide range of $[O\text{III}]$ line widths we measure in our sample. The data is shown in grey, the best-fitting parametric model in black, and the individual model components in orange. The broad $H\beta$ centroid is used to measure the systemic redshift, and Δv is the velocity shift from the line rest-frame transition wavelength for $H\beta$. Below each fit we plot the data minus model residuals, scaled by the errors on the fluxes. **Resample model at higher resolution.** 96
- Figure 4.3 The relative weight in each of the six positive ICA components for the high-luminosity (blue) and low luminosity samples (grey). In the high-luminosity sample Fe II emission is stronger (component w_1). The core $[O\text{III}]$ emission is weaker (components w_4, w_5) but the strength of the blueshifted wing is the same (w_6). 97
- Figure 4.4 The relative weight in the three ICA components corresponding to $[O\text{III}]$ emission (*left*) and the relative weight of the component most closely related to blueshifted $[O\text{III}]$ emission relative to all three $[O\text{III}]$ components (*right*). $[O\text{III}]$ emission is weaker in the high-luminosity sample, but the relative contribution but the fractional contribution from the blueshifted component to the total $[O\text{III}]$ emission is higher. Hence $[O\text{III}]$ is weaker, broader, and more asymmetric in the high-luminosity sample. See Zakamska discussion. 98
- Figure 4.5 Comparison of $[O\text{III}]$ velocity-widths in the high and low luminosity samples using the ICA component fits. If keep this need to explain in text how w_{80} is calculated from ICA component fits. 98
- Figure 4.6 Black solid line is the median from the ICA fits to the high-luminosity sample. Black dashed line shows the median from the low-luminosity sample. The six positive ICA components are also shown. 99

- Figure 4.7 [O III] strength decreases as the C IV blueshift increases, I run in to problems comparing the C IV blueshift to the [O III] blueshift / velocity-width. See similar thing if I use [O III] EQW instead. Need to fix y ticks. Only showing the core components here. The C IV blueshift is now measured relative to the NIR ICA redshift. I think this trend is mostly being driven by the Eigenvector 1 correlations: as the blueshift increases the Fe II strength increases and the [O III] strength decreases. Doesn't appear to be driven by the luminosity. Is this tighter than EV1 trend shown with Fe/OIII strength by other authors? Is the AGN NLR absent in objects where outflows have reached kiloparsec scales, sweeping up the low-density material responsible for the [OIII]-emission? 99
- Figure 4.8 I think there is a trend here but at high blueshifts the OIII is undetected / very low S/N. Need to determine when we can believe OIII parameters. Why at low CIV blueshift is there a much bigger dynamic range than in [O III] blueshifts in Fig. 13. Is it just because we have more objects? 100
- Figure 4.9 Comparison of median [O III] profiles from ICA fits to low- and high-luminosity samples. 101
- Figure 4.10 ICA median weights as a function of the CIV blueshift. 102
- Figure 4.11 Example where poorly subtracting the iron can be confused with [O III]. 102
- Figure 4.12 104
- Figure 4.13 Redshift comparisons. Lots have been excluded from Ha/Hb so need to look at flags greater than one. What is the big peak? Gaussian fit to the first one has failed. Find out why these plots look different to ones in paper. 105
- Figure 4.14 The [O III] velocity-width, characterised by w_{80} , as a function of the [O III] luminosity and the quasar redshift. The color of each hexagon denotes the mean w_{80} for the objects in that luminosity-redshift bin. We have supplemented our sample with low- z objects from Zakamska and Greene, (2014) and medium ($z \sim 1.5$) redshift objects from Harrison et al., (2016). If I keep this plot make sure it's clear which points belong to which sample. 108

- Figure 4.15 The [O III] EW as a function of the quasar bolometric luminosity for the sample presented in this paper (blue circles) and the low- z SDSS sample (grey points and contours). Upper limits are denoted by the downward arrows. 110
- Figure 4.16 The distributions of and correlations between a subset of the non-parameteric measures we made of the best-fitting [O III] models. 112
- Figure 4.17 The relation between the blueshifts of C IV and [O III]. Equivalent to Fig 8. We use the H β peak in this figure, which I think is responsible for some of the trend. However, we do see a correlation (albeit noisier) using the NIR ICA redshifts. Not a sensible to use the [O III] redshifts, since these become much more unreliable at the high C IV blueshift end (when [O III] is weaker: figure 7. Note that we are using v_{10} for the [O III] position and v_{50} for the C IV position. We can't use v_{50} for [O III] because sometimes we are using a single Gaussian, especially if the [O III] is weaker and we miss the broad component. Need to remake this plot / don't use at all because I don't believe some of the Gaussian fits to [O III], especially at high C IV blueshifts when [O III] is weak and Fe II is strong.) Only objects where fit with two components. 113
- Figure 4.18 114
- Figure 4.19 The [O III] EQW as a function of the H β FWHM and the optical Fe II strength (EQW Fe II/ EQW H β). 115
- Figure 4.20 The distribution of w_{80} and v_{10} for the 19 BALs are compared to the distribution for the non-BALs. These look rubbish. Cumulative distributions instead? Try doing gaussian kernel density estimator 116
- Figure 5.1 Model spectrum at $z = 1$, showing the contributions to the total flux from the blue power-law slope, red power-law slope, blackbody and host galaxy. The locations of the most prominent emission lines in the spectrum are also indicated. 121

- Figure 5.2 Model spectrum at three different redshifts (each arbitrarily scaled), and throughput functions for SDSS, UKIDSS and WISE band-passes (scaled so that the peak transmission is equal to one.) The dashed line indicates the slope of the AB magnitude system zero point. 128
- Figure 5.3 Colours of median SED (*black circles*), individual objects (*grey points*), best-fitting model (*black line*) as a function of redshift. 130
- Figure 5.4 Colours of median SED (*black circles*), individual objects (*grey points*), best-fitting model (*black line*) as a function of redshift. 131
- Figure 5.5 Residuals from fit to DR7Q-matched catalogue as a function of rest-frame wavelength. 132
- Figure 5.6 Ratio of NIR to UV luminosity ($R_{\text{NIR/UV}}$) against temperature (T_{BB}) for low- z sample. The density of points is shown in more dense regions of the space, and individual objects in less dense regions. 134
- Figure 5.7 $i - K$ vs z . Demonstrates how sample was defined. The grey points show, as a function of redshift, the $i - K$ colours of all DR7Q quasars which are not classified as broad-absorption line quasars by Shen et al. and i magnitude > 19.1 . The black line shows the $i - K$ colour of our standard, unreddened SED model as a function of redshift. The red and blue lines show the $i - K$ colours of our SED model with dust reddening $E(B-V) = 0.075$ and $E(B-V) = -0.075$ respectively. A significant amount of this reddening can be attributed to intrinsic variations in the UV power-law slopes of the individual quasars, which is why we allow a negative reddening. However, there is a clear ‘red tail’ to the colour distribution which can be explained by dust reddening at the redshift of the quasar. We defined two samples, at low ($0.5 < z < 1.5$) and high ($2 < z < 2.7$) redshift, which are shown in the figure. 136
- Figure 5.8 $W1 - W2$ colours of DR7 sample as a function of redshift. Above a certain density threshold points are represented by a density plot. On top we plot the colours of our standard SED model, with a fixed temperature and a varying NIR ($1 - 3 \mu\text{m}$) to UV ratio. 137

Figure 5.9	Ratio of NIR to UV luminosity ($R_{\text{NIR/UV}}$) against temperature (T_{BB}). The grey contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator on our data sample. The black points are for our mock data. 138
Figure 5.10	Best-fit black-body temperature against UV luminosity (left), black-hole mass (center) and Eddington ratio (right) for $1 < z < 1.5$ sample (black) and $2 < z < 2.7$ sample (black). In region of high-density we represent the density with contours generated using a Gaussian kernel density estimation. Needs re-making with new BH masses. 139
Figure 5.11	Composite SDSS spectra for objects at $z \sim 0.7$. We have divided sample into objects with objects best-fit by small (red line) and large (red line) values of β . Change this to select by $R_{\text{NIR/UV}} / T_{\text{BB}}$. Label prominent emission lines. 140
Figure 5.12	Rest-frame equivalent width and blueshift of the C IV line for 7,115 SDSS DR7 quasars. The colours of the hexagons denote the median hot dust ($T \simeq 1200$ K) abundance for all quasars at a given equivalent width and blueshift. Quasars with the most extreme outflow signatures are predominantly hot-dust rich. Only bins containing a minimum of two objects are plotted. 141
Figure 5.13	142
Figure 5.14	142

LIST OF TABLES

Table 2.1	The numbers of quasars, the spectrographs and telescopes used to obtain the near-infrared spectra, and the instrumental configurations. 15
Table 2.2	The format of the table containing basic information on our sample. The full catalogue is available to download at the following URL. 21
Table 3.1	The numbers of quasars with reliable $H\alpha$ and $H\beta$ line measurements, and the spectrographs and telescopes used to obtain the near-infrared spectra 36

Table 3.2	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 journal. 46
Table 3.3	The number of spectra removed from our sample by the cuts described in Section 3.5.5. 46
Table 3.4	The fractional error on the corrected BH mass as a function of C IV blueshift for different uncertainties in the quasar systemic redshift. 72
Table 4.1	The numbers of quasars with [O III] line measurements and the spectrographs and telescopes used to obtain the near-infrared spectra. Further details on the instrumental configurations are given in paper I. Check numbers. 94
Table 4.2	Approximate physical origin of the ICA components. 95
Table 5.1	Best-fitting parameters from fit to DR7Q-matched sample. Only give best-fit values after correction. 129

LISTINGS

ACRONYMS

INTRODUCTION

1.1 THE AGN-HOST GALAXY CONNECTION

Super-massive black holes (BHs) are found at the centres of most nearby massive galaxies and the BH mass and mass of the host galaxy spheroid are strongly correlated (Ferrarese and Merritt, 2000; Gebhardt et al., 2000; Kormendy and Ho, 2013). Although any underlying causal mechanism(s) responsible for the correlation is yet to be conclusively identified, there is considerable observational and theoretical support for models that involve BH-fuelling, outflows and a ‘feedback’ relationship (e.g. King and Pounds, 2015). The number density of quasars, which evolves strongly with redshift, peaks at redshifts $2 \lesssim z \lesssim 3$ (e.g. Brandt and Hasinger, 2005; Richards et al., 2006b) and the most massive ($M_{\text{BH}} \gtrsim 10^9 M_{\odot}$) present-day BHs experienced much of their growth during this epoch. The star formation rate, which closely follows the cosmological evolution of the quasar luminosity function, also peaks during this epoch (e.g. Boyle and Terlevich, 1998). Quantifying the growth-rate of massive BHs at $2 \lesssim z \lesssim 3$ would therefore help significantly in understanding the role quasars play in galaxy evolution.

There is now considerable observational and theoretical support for models of galaxy formation that involve black hole-fuelling, outflows and a ‘feedback’ relationship between active black holes and star formation in the host galaxy. Super-massive black holes accreted most of their mass and galaxies formed most of their stars at redshifts $z \gtrsim 2$ (e.g. Madau & Dickinson 2014 for star formation; find quasar reference.) During this key cosmological epoch star formation is believed to be suppressed by the energy output from the quasar, establishing the tight relationship between BH mass and host galaxy spheroid mass observed in the local Universe (e.g. Kormendy & Ho 2013).

1.2 MEASURING BLACK HOLE MASSES

The goal of better understanding the relationship between super-massive BH accretion and star formation has led to much work focussing on the properties of quasars and active galactic nuclei at these redshifts. Accurate BH mass estimates for quasars are essential in these studies. Furthermore, as one of just two fundamental quantities describing a black hole on astrophysical scales, the mass is of crucial importance

to virtually all areas of quasar science, including the evolution and phenomenology of quasars, and accretion physics.

1.2.1 Reverberation Mapping

Reliable estimates of BH masses are a prerequisite for investigating the relationship between BHs and their host galaxies. If the line-emitting clouds in the broad line region (BLR) are assumed to be virialized and moving in a potential dominated by the central BH, then the BH mass is simply a product of the BLR size and the square of the virial velocity. The reverberation-mapping technique uses the time lag between variations in the continuum emission and correlated variations in the broad line emission to measure the typical size of the BLR (Peterson, 1993; Peterson, 2014). The full width at half maximum (FWHM) or dispersion (σ ; derived from the second moment) velocity of the prominent broad emission line of $H\beta$ (4862.7\AA)¹ is used as an indicator of the virial velocity, with extensions to other low-ionization emission lines such as $H\alpha$ (6564.6\AA) and $Mg\text{ II}\lambda 2796.4, 2803.5$ (e.g. Vestergaard, 2002; McLure and Jarvis, 2002; Wu et al., 2004; Kollmeier et al., 2006; Onken and Kollmeier, 2008; Wang et al., 2009; Rafiee and Hall, 2011). Extensive reverberation mapping campaigns have provided accurate BH masses for ~ 50 active galactic nuclei (AGN) at relatively low redshifts and of modest luminosity (e.g. Kaspi et al., 2000; Kaspi et al., 2007; Peterson et al., 2004; Bentz et al., 2009; Denney et al., 2010). [See galaxies talk for a few more details]

1.2.2 Single-Epoch Virial Estimates

Reverberation mapping campaigns have also revealed a tight relationship between the radius of the BLR and the quasar optical (or ultra-violet) luminosity (the $R-L$ relation; e.g. Kaspi et al., 2000; Kaspi et al., 2007). This relation provides a much less expensive method of measuring the BLR radius, and large-scale studies of AGN and quasar demographics have thus become possible through the calibration of single-epoch virial-mass estimators using the reverberation-derived BH masses (e.g. Greene and Ho, 2005; Vestergaard and Peterson, 2006; Vestergaard and Osmer, 2009; Shen et al., 2011; Shen and Liu, 2012; Trakhtenbrot and Netzer, 2012). The uncertainties in reverberation mapped BH masses are estimated to be ~ 0.4 dex (e.g. Peterson, 2010), and the uncertainties in virial masses are similar (e.g. Vestergaard and Peterson, 2006). Since the structure and geometry of the BLR is unknown, a virial coefficient f is introduced to transform the observed line-of-sight velocity inferred from the line width in to a virial velocity. This simplification accounts for a significant part of the uncertainty in virial BH masses (in addition to, for example, describ-

¹ Vacuum wavelengths are employed throughout the thesis.

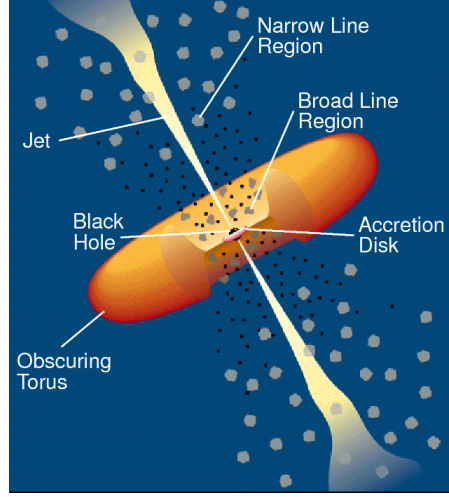


Figure 1.1: Illustration of the physical structure of an AGN in a simple orientation-based unification model. From Urry and Padovani, (1995).

ing the BLR with a single radius R and scatter in the $R - L$ relation; Shen, 2013). Furthermore, if the BLR is anisotropic (for example, in a flattened disk; e.g. Jarvis and McLure, 2006) then the line width will be orientation-dependent (e.g. Runnoe et al., 2013b; Shen and Ho, 2014; Brotherton et al., 2015).

For example, single epoch estimates have been used to calculate black hole masses in the highest redshift quasars to study the growth of SMBHs. This figure shows a compilation of SE mass estimates for quasars over a wide redshift range from different studies. These studies show that massive, 10^9 BHs are probably already in place by $z \sim 7$, when the age of the Universe is less than 1 Gyr. This places strong constraints on BH growth models. Single epoch masses have also been used to study the distribution of quasars in the BH mass-luminosity plane, which conveys important information about the accretion process of these active black holes (e.g. Kollmeier et al. 2016). Reshift evolution of BH-bulge scaling relations (e.g. Bennert et al. 2011). Clustering (Shen & Ho 2014; Timins et al.).

An *active galactic nucleus*, or AGN, is an energetic, non-stellar phenomena in the central region of a galaxy. AGN are powered by the accretion of gas, primarily through an accretion disk, onto a central *super-massive black hole* (SMBH) of mass $10^6 - 9 M_\odot$ (Lynden-Bell, 1969). With bolometric luminosities in the range $10^{44-48} \text{ ergs}^{-1}$, they are the most luminous persistent sources of radiation in the Universe.

1.3 ORIENTATION-BASED UNIFICATION MODELS

AGNs are divided into numerous classes and sub-classes based on their observational properties. AGN unification models (Antonucci,

1993; Urry and Padovani, 1995) attempt to explain the diversity in their observational properties using as few physical parameters as possible. In many unification models the optical and radio luminosities are considered to be intrinsic parameters. Variations in the radio luminosity explains the difference between the 15 – 20% of AGN which are *radio-loud* (i.e. have radio to optical flux ratios $\gtrsim 10$) and the remainder which are *radio-quiet*. The optical luminosity explains, for example, the difference between low-luminosity *Seyfert Galaxies* and high-luminosity *quasars*². In a typical quasar the optical emission may be brighter than the combined emission from all of the stars in the host galaxy by a factor of 100 or more. The large brightness contrast between the nucleus and the host galaxy makes the host galaxy difficult to detect, and the quasar is observed as an unresolved stellar-like object³.

Unification models attempt to explain all further observational differences as being apparent differences due to *orientation* effects. The basic physical structure of an AGN in this model is illustrated in Figure 1.1. Material is pulled towards the SMBH at the centre and sheds angular momentum through viscous and turbulent processes in an accretion disk, which radiates primarily at ultraviolet (UV) to soft-X-ray wavelengths. Strong optical and UV emission lines are produced in photo-ionised gas clouds moving rapidly in close proximity to the SMBH. The doppler-broadened emission line widths imply gas cloud velocities of thousands of km s^{-1} in this *broad-emission line*. Further out are dusty, molecular clouds, the geometry of which is often modelled as a torus co-planar with the accretion disk. Along some lines of sight from the observer to the accretion disk / broad line region the dusty torus obscures the UV/optical radiation. In this case, an observer would see a weak UV/optical continua and no broad emission lines and classify the AGN as being *Type II*. On the other hand, if the line of sight is unobscured by the dusty torus then a broad emission line component would be observable in the spectrum and the AGN would be classified as being *Type I*. Further away from the central black hole and beyond the dusty torus are slower moving clouds of gas which are photo-ionised by the continuum emission from the accretion disk and produce forbidden emission lines of narrower widths (typically hundreds of km s^{-1}). Outflows of energetic particles occur along the poles of the accretion disk and form collimated radio-emitting jets and in some cases giant radio-emitting lobes. A strong, relativistically beamed component with large variations in brightness on very short timescales (e.g. $\Delta m \gtrsim 0.1$ and $\Delta t \lesssim 1$

² The term ‘quasar’ is sometimes reserved for radio-loud objects and ‘quasi-stellar object’, or ‘QSO’, for radio-quiet objects. Here, ‘quasar’ is used to refer to all luminous AGN.

³ The name ‘quasar’ is shortened from ‘quasi-stellar radio source’, since quasars were originally discovered as optical point-source counterparts to a newly discovered population of radio sources.

day) is observed and a source with these properties is classified as a *blazar*.

While an orientation-based unification scheme such as this is somewhat successful at explaining many of the observational properties of AGNs, other factors such as the host galaxy morphology and gas/dust content may also be important (Peterson, 1997). It is also doubtful whether the geometry of the dusty torus is the same in all AGNs, and the fraction of obscured quasars has been shown to decrease with increasing nuclear luminosity (Lawrence, 1991). As we will now discuss, quasars might play an important role in a broader cosmological context, affecting the formation and evolution of the galaxies, groups, and clusters in which they reside. In this scenario of galaxy/quasar co-evolution the quasar is expected to transition from a highly active obscured phase to an unobscured phase as it clears out the dust surrounding it. If this picture is true then we should expect to find variations in the observational properties of the quasar and host galaxy as the system transitions through the different stages of its evolution.

1.4 THE TORUS

Elitzur & Shlosman (2006):

Recent high-resolution IR observations indicate that the torus size might be no more than a few parsecs (Elitzur 2005 and references therein); in particular, VLTI observations of NGC 1068 show that the FWHM size of the 12 mm emission is only 4 pc (Jaffe et al. 2004). The compact sizes place the torus inside the region where the SBH gravity dominates over the galactic bulge.

Two approaches have been taken for the torus dynamic origin. A hydrostatic scenario depicts the torus as a doughnut-like structure populated by molecular clouds accreted from the galaxy (Krolik & Begelman 1988). However, the origin of vertical motions capable of sustaining the clouds in a hydrostatic structure with $H \sim R$ was recognized from the start as problematic and has eluded solution thus far (e.g., Davies et al. 2006). The other scenario, based on the seminal work by Blandford & Payne (1982), involves the outflow of clouds embedded in a hydromagnetic disk wind (Emmering et al. 1992, hereafter EBS92; Konigl & Kartje 1994; Kartje & Konigl 1996; Bottorff et al. 1997, 2000; Kartje et al. 1999; Everett 2005). In this approach the torus is merely a region in the wind that happens to provide the required toroidal obscuration; i.e., it is that region wherein the clouds are dusty and optically thick.

1.5 EVOLUTIONARY MODELS

A number of observations link the growth and evolution of quasars to the growth and evolution of galaxies. These include the following:

1. SMBHs appear to be a ubiquitous feature at the centres of all massive galaxies (e.g. Kormendy and Ho, 2013).
2. SMBH masses are proportional to the mass/velocity dispersion of their host spheroid (the $M - \sigma$ relation; Ferrarese and Merritt, 2000; Gebhardt et al., 2000).
3. The cosmological evolution of the star formation rate and the quasar luminosity function are very similar (e.g. Wall et al., 2005).
4. Cosmological simulations of galaxy formation and evolution require feedback from SMBH growth in order to reproduce the galaxy luminosity function (Kauffmann and Haehnelt, 2000).

These observations suggest that all galaxies may have gone through a ‘quasar phase’ during which the SMBH accretes most of its mass and the stellar-bulge forms most of its stars. This evolutionary phase could be triggered by a major merger or by instabilities in the galactic disc or bulge. In a galaxy merger large amounts of gas can shed sufficient angular momentum to settle into dense clouds and form stars or be funnelled to the centre of the galaxy to grow the existing SMBH. The large amounts of gas and dust funnelled inward to the galactic nucleus is predicted to obscure the quasar until the dust is cleared out either by quasar-driven or stellar-driven processes. An unobscured quasar then emerges, and is active until all of the available material has been accreted (Hopkins et al., 2006; Narayanan et al., 2010). The feedback processes involved are also thought to be responsible for shutting down star formation in the galactic bulge (Silk and Rees, 1998) and establishing the $M - \sigma$ relation.

Such scenarios have been invoked to explain the presence of buried AGN seen in ultra-luminous infra-red galaxies (ULIRGs; Sanders et al., 1988), a high fraction of which also show evidence of merging and interaction. However, the full picture is likely to be more complicated. Although there is evidence that mergers dominate at high luminosities (Treister et al., 2012), stochastic accretion may be more important at low luminosities (e.g. Hopkins and Hernquist, 2006).

Luminous unreddened quasars show few signs of interaction (e.g. Dunlop et al., 2003) which, if the quasar-galaxy co-evolution model is true, suggests that indications of an interaction disappear during a transitional phase. Quasars in this transitional phase would be highly reddened, as the dust enshrouding the nucleus will not have been fully cleared, but not completely obscured. A population of quasars with these properties may therefore represent a link between ULIRGs and unobscured quasars.

1.6 INTERESTING SUB-POPULATIONS

1.6.1 *Red and Reddened Quasars*

Magnitude limited optical surveys of quasars are biased against selecting red and reddened quasars. Richards et al., (2003) studied a large sample of optically selected Sloan Digital Sky Survey (SDSS; York et al., 2000) quasars and showed the mean reddening to be $E(B - V) = 0.03$ at the redshift of the quasar. They estimated that $\sim 15\%$ of the population was missing from the survey due to dust extinction. The missing fraction, and its dependence on luminosity and redshift, could help to determine whether the reddened population is best explained in the context of orientation-based unification models with non-spherical geometry or as an evolutionary stage in a quasars lifetime.

Populations of heavily dust-reddened quasars have been identified using radio surveys (e.g. Glikman et al., 2012), by using the ‘K-band excess’ in the spectra of quasars relative to stars (Maddox et al., 2012), and using near-IR colour selection (Banerji et al., 2012; Banerji et al., 2013). Recently, Ross et al., (2014) identified a small sample of very red SDSS quasars based on their extreme IR to optical luminosity ratios. It is yet to be determined whether these extreme objects are simply the tail of a population dominated by less reddened quasars, or whether the distribution is bi-modal with reddening. A population of quasars with intermediate amounts of dust reddening ($0.1 \lesssim E(B - V) \lesssim 0.5$) would help to address this question.

1.6.2 *Broad Absorption Line Quasars*

Broad absorption line quasars (BALQSOs) are a sub-population of quasars exhibiting blue-shifted absorption troughs broader than 2000km s^{-1} (Weymann et al., 1991) which are unambiguously associated with AGN-driven out-flowing gas. As well as showing high rates of mergers, an anomalously large fraction of heavily reddened objects exhibit broad blue-shifted absorption troughs in their spectra (Urrutia et al., 2009; Glikman et al., 2012). This observation suggests that the BAL phenomenon may be related to a ‘blow-out’ phase of a quasars lifetime as it transitions from a dusty, obscured object to a luminous blue quasar, at the same time quenching star formation. Since outflows are believed to be fundamental to AGN feedback, a better understanding of their properties could shed light on the outflow phenomenon. Alternatively, whether a quasar is observed to have broad absorption lines could depend only on the orientation of the observer in relation to an intrinsically anisotropic system.

1.6.3 *Hot-Dust-Poor Quasars*

The near-IR emission from AGN is generally explained by thermal emission from dust grains at the edge of the dusty torus closest to the accretion disk. The dust is heated to its sublimation temperature (1300-2000K Barvainis, 1992) by emission from the accretion disc. However, Hao et al., (2010) reported that 6% (at $z \lesssim 2$) to 20% (at $2 \lesssim z \lesssim 3.5$) of the quasars in the X-ray selected XMM-COSMOS Type 1 AGN sample (Brusa et al., 2010) have an unusually small amount of hot dust emission, despite having normal accretion disc spectra. They infer a torus covering factor of $\sim 2\%$ to 30% for these ‘hot dust poor’ (HDP) quasars, well below the $\sim 75\%$ predicted by unified models (e.g. Krolik and Begelman, 1988). Hao et al., (2011) found that HDP quasars were just as common in the Richards et al., (2006a) Spitzer/S-DSS sample ($8.7\% \pm 2.2\%$) and the Elvis et al., (1994) Palomar-Green-quasar-dominated sample ($9.5\% \pm 5.0\%$). Either the hot dust is destroyed (dynamically or by radiation), or the dust is not centred on the SMBH, which could happen during a major merger (e.g. Blecha et al., 2011). Alternatively, misaligned accretion disks, which will result from discrete isotropic accretion events (Volonteri, Sikora, and Lasota, 2007), will lead to a wider range of covering factors (Lawrence and Elvis, 2010).

At higher redshifts, Jiang et al., (2010) found two HDP quasars in a sample of 21 at $z \sim 6$. They find that at $z \sim 6$ the hot dust abundance is roughly proportional to the black hole mass, indicating that the two grow at about the same rate. The two HDP quasars also have the smallest SMBH masses, and may be too young to have formed a significant amount of hot dust.

1.6.4 *Type II Quasars*

As well as lacking a broad-line spectral component, Type II AGN tend to have high IR to optical light ratios, hard X-ray spectra, and be strongly polarised, consistent with dusty torus based unification schemes. The detection of unobscured continuum emission that is scattered and polarised by dust above the torus has confirmed the orientation-based unification of Type I and Type II Seyfert Galaxies. Their higher luminosity analogues, Type II quasars, have been much more difficult to detect and study. It is possible that the orientation-based Type I/II unification scheme may break down at high-luminosities, and that instead all quasars could pass through a Type II phase before the obscuring dust is cleared out by the quasar-driven outflows and a Type I quasar emerges.

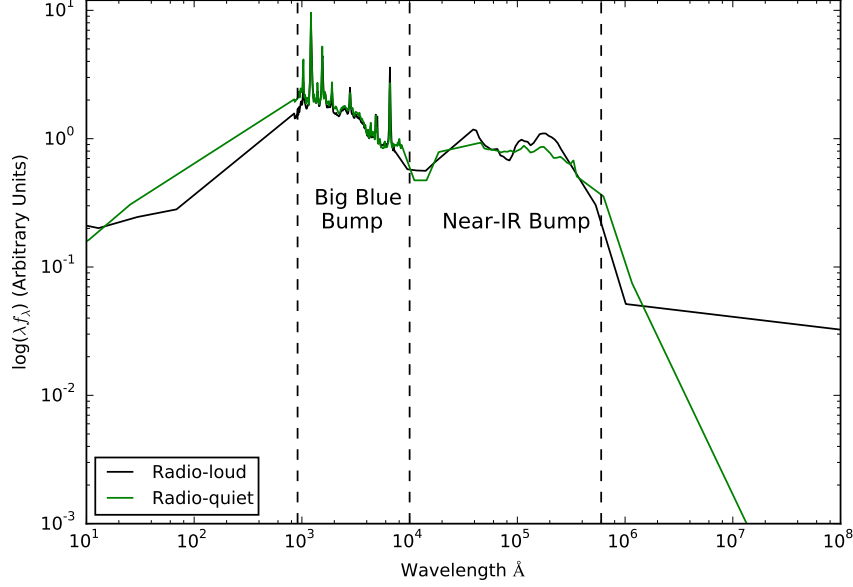


Figure 1.2: Median SEDs for radio-loud and radio-quiet quasars from Shang et al., (2011).

1.7 SPECTRAL ENERGY DISTRIBUTIONS

AGNs emit strongly over many decades in frequency of the electromagnetic spectrum and the energy emitted as a function of frequency is described by a *spectral energy distribution* (SED). As we will describe below, the broad features in the SED originate from processes which occur in different regions of the AGN. In the preceding sections, we have described how some interesting sub-populations of AGNs might relate to the broader population in the context of orientation-based unification schemes and evolutionary schemes. Comparing SEDs of different sub-populations can help to shed light on these relationships and the physical processes which drive them. Possible correlations between the SED shape and luminosity, redshift, and other properties of the AGN such as black hole mass and Eddington ratio can also constrain models of AGN structure and evolution.

Since the physical processes involved are generally understood only qualitatively, almost all AGN SED templates are empirical. The empirical template of Elvis et al., (1994), constructed using photometric observations from the radio to the hard X-rays of 29 radio-quiet and 18 radio-loud Type I quasars, is still the most commonly cited, despite many additions and updates (e.g. Polletta et al., 2000; Kuraszkiewicz et al., 2003; Risaliti and Elvis, 2004; Richards et al., 2006a; Polletta et al., 2007; Lusso et al., 2010; Shang et al., 2011; Marchese et al., 2012; Trichas et al., 2012). In Figure 1.2 we show the median spectrum from the radio-loud and radio-quiet samples of Shang et al. (2011). Short-

ward of the radio region, the radio-loud and radio-quiet spectra are almost indistinguishable.

A large amount of energy is emitted in the UV/optical region shortward of $\sim 4000\text{\AA}$: the *Big Blue Bump*. In the X-ray region, the *soft X-ray excess* may be the high-energy end of this feature. The Big Blue Bump is generally attributed to thermal emission from the accretion disk. In Type II AGNs, the continuum emission from the accretion disk is obscured, and so the Big Blue Bump in the SED of a Type II AGN would be less prominent than is seen in Figure 1.2.

The feature at wavelengths longward of $\sim 1\mu\text{m}$ is the *IR Bump*, and is generally attributed to thermal emission from dust at a wide range of temperatures ($\sim 50 - 1000\text{ K}$). The amount, geometry, ionisation and optical depth of absorbing dust and gas and its inclination determines the shape of the IR Bump and the absorption of the optical/UV continua. The relative strengths of the IR Bump and the Big Blue Bump are generally comparable, although they do vary from object to object. In particular, for the HDP objects we described above the IR Bump appears to be missing entirely. The minimum between the two peaks is at $\sim 1\mu\text{m}$, which reflects the sublimation of dust at $T \gtrsim 2000\text{ K}$ (Sanders et al., 1989).

Emission in the hard X-ray region of the spectrum is believed to be due to Compton up-scattering of accretion disk photons by hot electrons forming a corona in the vicinity of the disk (e.g. Sunyaev and Titarchuk, 1980). The radio emission, which originates from synchrotron emission in relativistic jets, contributes very little to the total energy output. However, the mechanical energy provided by the jets is an important component of AGN feedback models (e.g. Fabian, 2012).

Many parameters might be expected to affect the shape of the AGN SED (e.g. the black hole mass, the accretion rate, the physical properties of the accretion disk, the properties of the absorbing dust) and many of these properties might be expected to change as the quasar evolves (e.g. as dust is expelled from the nuclear regions). Given this, it is perhaps surprising that many authors have found no significant dependence of the mean SED on properties such as redshift, bolometric luminosity, SMBH mass, or accretion rate (e.g. Elvis et al., 2012; Hao et al., 2013) and that quasars up to redshift 7 have been shown to have similar UV spectra to low redshift quasars (e.g. Mortlock et al., 2011).

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

Everett et al. (2005):

A variety of observational signatures point to the importance of outflowing gas within many types of active galactic nuclei (AGNs).

Blueshifted absorption features (in broad absorption line quasars, or BALQSOs; see, e.g., Weymann et al. 1991) are seen in approximately 15% (Reichard et al. 2003) of radio-quiet quasars, with velocities up to $0.1c$. In addition, radio-loud quasars display relativistic, collimated outflows. There has also been observational evidence that suggests the mass outflow rate in AGNs is nearly equal to the mass inflow rate (see, e.g., Crenshaw et al. 2003; Chartas et al. 2003).

A NEAR-INFRARED SPECTROSCOPIC DATABASE OF HIGH-REDSHIFT QUASARS

Rest-frame optical quasar emission lines provide much information. This includes black hole masses and systemic redshifts. Rest-frame optical lines, including [O III], H β and H α , are redshifted to infrared wavelengths at redshifts $z \gtrsim 2$. However, the number of quasars at these redshifts with near-infrared spectra is limited. This is unfortunate because the peak epoch of galaxy evolution is at redshifts $2 \lesssim z \lesssim 4$.

In this chapter I will describe the construction of a database containing 553 high redshift quasars. This is the largest sample constructed, and is invaluable in a number of investigations. Because of this I intend to publically release the data in the near future.

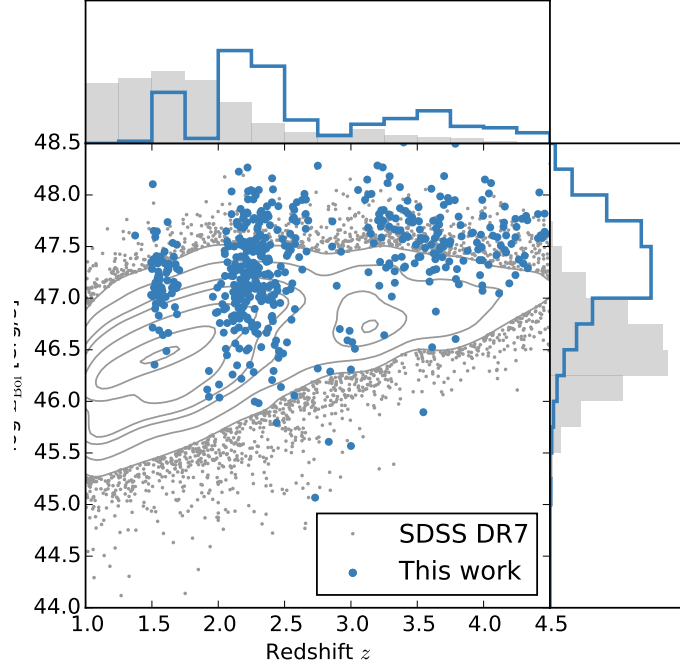


Figure 2.1: The ranges in redshift and luminosity covered by our sample, relative to the redshift-luminosity distribution of the SDSS DR7 quasar catalogue. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator. For the SDSS sample we use Hewett and Wild, (2010) redshifts and bolometric luminosities measured by Shen et al., (2011). For the quasars in this paper the redshift is defined using the peak of the $H\alpha/H\beta$ emission and the luminosity is measured in the continuum at 1350\AA and converted to a bolometric quantity using the same conversion factor employed by Shen et al., (2011). Eight objects are missing because we do not have enough information to calculate the bolometric luminosity.

Table 2.1: The numbers of quasars, the spectrographs and telescopes used to obtain the near-infrared spectra, and the instrumental configurations.

Spectrograph	Telescope	Resolving power $\lambda/\Delta\lambda$	Wavelength coverage μm	Slit width arcsec	Exposure times hr	Number
FIRE	MAGELLAN	6000	0.80-2.50	0.6	0.5-1.0	18
GNIRS	GEMINI-N	5400	0.85-2.50	0.3-0.45	0.3-1.3	22
ISAAC	VLT	5100	1.40-1.82	0.6	0.6-1.3	0
LIRIS	WHT	945	1.39-2.42	1.0	0.2-0.8	15
NIRI	GEMINI-N	520-825	1.43-1.96	0.47-0.75	0.5-2.7	0
NIRSPEC	Keck II					20
SINFONI	VLT	2000-3000	1.10-2.45 ^a		0.1-0.7	2
SOFI	NTT	1000-2000	1.53-2.52 ^b	0.6	0.5-1.8	47
TRIPLESPEC	ARC-3.5m	2500-3500	0.95-2.46	1.1-1.5	1.0-1.5	33
TRIPLESPEC	P200	2500-2700	1.00-2.40	1.0		23
XSHOOTER	VLT	4350-7450	0.30-2.50	0.5-1.6	0.2-0.8	4

^a J, H or K filters were employed to ensure coverage of the H β /[O III] spectral region.

^b Both the low resolution red grism and the medium resolution grism, with K and H filters, were employed.

2.1 COATMAN ET AL. (2016) QUASARS

We selected quasars from the spectroscopic quasar catalogue of the Sloan Digital Sky Survey (SDSS; York et al., 2000) Seventh Data Release (DR7; Schneider et al., 2010) with a wide range of C IV blueshifts. The sample was restricted to objects with redshifts $2.14 < z < 2.51$ (7,258 quasars), to ensure that the H β and H α emission lines fall within the H- and K-bands respectively, allowing us to observe both simultaneously with the appropriate grism configuration. Given the limited number of quasars for which near-infrared spectra could be obtained, the quasar sample was further restricted to objects that are radio-quiet (5,980 quasars), show no evidence of broad absorption lines (BALs) in their spectra (5,299 quasars), and are free from significant dust extinction. We removed radio-loud objects from our sample using the same radio-loud classification as Shen et al., (2011), and BAL quasars using the classifications of both Shen et al., (2011) and Allen et al., (2011). The removal of quasars with significant dust extinction was achieved by identifying quasars with $i - K$ colours redder than a parametric spectral energy distributions (SED) model + SMC-like extinction curve with $E(B - V) = 0.05$ (see Maddox et al., 2012). The K-magnitude was taken from the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al., 2007) Large Area Survey (ULAS). The requirement to be in the ULAS footprint and have reliable K band photometry reduced our sample of possible targets to 1,683, and the $E(B - V)$ cut left 1,204 in our sample. Finally, a flux-limit of $K < 18.5$ (AB) was applied to ensure that spectra of sufficient signal-to-noise ratio (S/N) could be obtained (412 quasars).

We were able to obtain new infra-red spectra for 19 quasars from this sample of 412 possible targets (Section ??). The quasars included in this sub-sample were selected to have C IV-emission shapes which span the full range observed in the population. Reliably quantifying the distribution of C IV-emission shapes has been made possible thanks to redshift-determination algorithms (Hewett and Wild, 2010, Allen & Hewett 2016, in preparation) which are independent of the C IV-emission shape. Calculation of the C IV emission line parameters is described in detail in the next section.

Near-infrared spectra were obtained with the Long-slit Intermediate Resolution Infrared Spectrograph (LIRIS) mounted on the 4.2m William Herschel Telescope (WHT) at the Observatorio del Roque de los Muchachos (La Palma, Spain). Observations took place over four non-contiguous nights from 2015 March 31 to April 4. Approximately one night was lost due to poor weather and a further half-night was affected by poor transparency due to cloud. A one arc-second slit-width was employed and the LIRIS H + K low-resolution grism was selected, which covers the spectral ranges $1.53\text{--}1.79\text{ }\mu\text{m}$ and $2.07\text{--}2.44\text{ }\mu\text{m}$ with a dispersion of $9.7\text{ }\text{\AA}/\text{pixel}$. The spatial scale of

the instrument is 0.25 arcsec/pixel. Observations were divided into 60 s sub-exposures and performed in an ABBA nodding pattern, with the object placed at two positions along the slit 12 arcsec apart. Bright A0-5V stars were observed at similar air-masses to the targets in order to provide both telluric absorption corrections and a flux calibration of the quasar spectra.

The raw LIRIS data frames incorporate a known ‘pixel shift’ which was first removed from all frames using the LIRIS data reduction package LIRISDR. Subsequent data reduction was undertaken with standard IRAF¹ procedures. The flat-field images, which were taken at the beginning of each night via illumination of the dome, were averaged and normalised to remove any wavelength-dependent signature. Each individual two-dimensional spectrum was then flat-field corrected. Consecutive AB and BA pairs of two-dimensional spectra were subtracted to remove the sky background. All the subtracted AB/BA-pairs for a target were then averaged to give the final two-dimensional spectrum.

The size of the one-dimensional spectrum extraction windows, in the slit direction, varied from 6-10 pixels. To increase the S/N, optimal variance-weighted extraction with sigma clipping was employed. For the fainter objects in our sample we were unable to trace the spectrum across the dispersion axis reliably and the trace from a telluric standard-star observation, observed at a similar air mass and time, was used instead. The wavelength calibration, using argon and xenon lamp exposures, resulted in root mean square errors in the range 1.01–1.71 Å, with a mean of 1.47 Å. The telluric standard star observations were reduced using the same steps described above. The stellar continuum was divided out of the standard star spectrum, which was then divided into the quasar spectrum to remove telluric absorption features. The spectral type and magnitude of the standard star were used to flux calibrate the quasar spectrum both in a relative and absolute sense. Variable atmospheric conditions combined with the narrow slit width resulted in a significant level of uncertainty in the absolute flux calibration for the quasar observations. The use of the UKIDSS broadband magnitudes (H and K) to normalise the spectra results in a significantly improved calibration.

2.2 SHEN & LIU (2012) AND SHEN (2016) QUASARS

Shen, (2016) and Shen and Liu, (2012) obtained near-infrared spectroscopy for a sample of 74 luminous, $1.5 < z < 3.5$ quasars selected from the SDSS DR7 quasar catalogue. Targets had to possess good optical spectra covering the C IV line and have redshifts $z \sim 1.5, 2.1$, and

¹ IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under a cooperative agreement with the National Science Foundation.

3.3 to ensure that the $H\beta$ -[O III] region was covered in one of the near-infrared JHK bands. Thirty-eight of the quasars were observed with TripleSpec (Wilson et al., 2004) on the Astrophysics Research Consortium (ARC) 3.5 m telescope, and 36 with the Folded-port InfraRed Echellette (FIRE; Simcoe et al., 2010) on the 6.5 m Magellan-Baade telescope. The reduction of the spectra is described in Shen, (2016) and Shen and Liu, (2012).

2.3 QUASAR PAIRS

Twenty per cent of our catalogue was observed as part of an ongoing effort to identify quasar pairs at very close projected separations (Quasars Probing Quasars² (QPQ); Hennawi et al., 2006a; Hennawi et al., 2010). The primary science driver of this work is to study the circum-galactic medium of the foreground quasars in absorption (Hennawi et al., 2006b). Very accurate systemic redshift measurements are a requirement and a large amount of effort has gone into obtaining near-infrared spectra which cover low-ionisation broad lines or features from the quasar narrow line region (Prochaska and Hennawi, 2009; Lau, Prochaska, and Hennawi, 2015; Hennawi et al., 2015). From the QPQ data set we identified 46 quasars with good-quality near-infrared spectra covering the $H\alpha$ and/or $H\beta$ lines and SDSS and/or BOSS spectra covering the C IV line. Twenty-two quasars were observed with the Gemini Near-Infrared Spectrograph (GNIRS; Elias et al., 2006) on the 8.1 m Gemini North telescope, 4 using the Infrared Spectrometer And Array Camera (ISAAC; Moorwood et al., 1998) on the European Southern Observatory (ESO) Very Large Telescope (VLT), 11 with the Near InfraRed Imager and Spectrometer (NIRI; Hodapp et al., 2003) also on Gemini North and 9 with XSHOOTER (Vernet et al., 2011), again, on the VLT. The broad wavelength coverage of XSHOOTER means that the spectra cover the region from C IV to $H\alpha$ at the redshifts targeted. The XSHOOTER spectra have higher S/N and resolution than the SDSS/BOSS spectra in the rest-frame ultraviolet and therefore the XSHOOTER spectra are used by default to measure the C IV emission.

The XSHOOTER spectra were reduced with a custom software package developed by George Becker (for details, see Lau, Prochaska, and Hennawi, 2015). The remaining data was processed with algorithms in the LowRedux³ package (see Prochaska and Hennawi, 2009).

2.4 VLT SINFONI QUASARS

We performed a search of the ESO archive for high- z quasars observed with the SINFONI integral field spectrograph (Eisenhauer et

² www.ucolick.org/~xavier/QPQ/Quasars_Probing_Quasars

³ www.ucolick.org/~xavier/LowRedux

al., 2003; Bonnet et al., 2004) at VLT/UT4. We found 79 quasars with redshifts $1.5 < z < 3.7$ which have H and/or K SINFONI spectroscopy, covering the H β and H α lines respectively. Seventy-two of the quasars are from a large programme led by L. Wisotzki (programme 083.B-0456(A)) to study the mass function and Eddington ratios of active BHs at redshifts $z \sim 2$ drawn from the Hamburg/ESO survey (Wisotzki et al., 2000). A further seven SINFONI spectra are from a programme led by J. D. Kurk (programme 090.B-0674(B)) to obtain reliable BH mass estimates from H α /H β for a sample of radio-loud/radio-quiet SDSS quasars.

The SINFONI spectra were reduced using the package EASYSINF⁴. The package, which is based on the ESO-SINFONI pipeline, is described in Williams et al., (2016).

2.5 ESO NTT SOFI QUASARS

Twelve per cent of the quasar catalogue derives from a large programme (programme 187.A-0645; PI: J. Hennawi) to combine near-infrared spectra from SOFI (Moorwood, Cuby, and Lidman, 1998) on the 3.6 m New Technology Telescope (NTT) with archival high-resolution optical spectra from the UV-Visual Echelle Spectrograph (UVES; Dekker et al., 2000) at VLT/UT2 and the High Resolution Echelle Spectrometer (HIRES; Vogt et al., 1994) at Keck to construct a legacy database of bright, high-redshift ($2 < z < 4$) quasars with both rest-frame optical spectra, covering the H β -[O III] complex, and high-resolution rest-frame ultraviolet spectra. The main science goal is to obtain precise systemic redshifts which are crucial for the study of absorption line systems. The SOFI spectra were reduced using a custom data reduction pipeline using algorithms in the LowRedux package.

Eighteen quasars have been targeted as part of the SDSS/BOSS spectroscopic quasar surveys. In addition, 13 reduced and fluxed UVES spectra have been made available to us by A. Dall’Aglia (a description of the reduction procedure is contained in Dall’Aglia, Wisotzki, and Worseck, (2008)). The sample is larger (~ 100 quasars) but reduced UVES spectra providing rest-frame ultra-violet coverage of C IV are not yet available for the remainder. The spectral resolution of the UVES observations is very high ($R \sim 40\,000$) and the S/N of the spectra re-binned to a resolution of $\simeq 2000$ is $S/N \simeq 300$.

Over five nights from 2015 August 31 to September 4 we obtained near-infrared SOFI spectra for a further 26 quasars (programme 095.B-0644(A); PI: L. Coatman). These quasars were selected from the SDSS DR7 quasar catalogue using criteria very similar to those described in Paper I (see Section ??). In particular, we selected quasars with

⁴ www.mrao.cam.ac.uk/~rw480/easysinf

large C IV blueshifts to improve the statistics in this region of the C IV emission-line parameter space.

2.6 HALE TRIPLESPEC QUASARS

A further 36 quasars in our catalogue are bright SDSS quasars which were observed with the TRIPLESPEC spectrograph on the Palomar 200-inch Hale telescope (P200). The objects were observed with the same science goals as the SOFI NTT large programme. The spectra were reduced using a custom pipeline, again using algorithms in the LowRedux package.

Table 2.2: The format of the table containing basic information on our sample. The full catalogue is available to download at the following URL.

	Units	Description
NAME		Name of spectra
DR7_SPEC		Name of SDSS/DR7 spectra
DR12_SPEC		Name of BOSS/DR12 spectra
ID		Name of quasar (not unique)
DATE		Date near-infrared spectra acquired
RA		
DEC		
INSTR_BAND		Instrument configuration
INSTR		Instrument used
z_ICA		
z_Line		
2massMag_j		
2massMag_h		
2massMag_k		
2massMagErr_j		
2massMagErr_h		
2massMagErr_k		
psfMag_u		
psfMag_g		
psfMag_r		
psfMag_i		
psfMag_z		
psfMagErr_u		
psfMagErr_g		
psfMagErr_r		
psfMagErr_i		
psfMagErr_z		
WISE_W1MPRO		
WISE_W2MPRO		
WISE_W3MPRO		
WISE_W4MPRO		
WISE_W1SIGMPRO		
WISE_W2SIGMPRO		
WISE_W3SIGMPRO		
WISE_W4SIGMPRO		
RADIO_FLAG		
BAL_FLAG		

3.1 ABSTRACT

The C IV $\lambda\lambda 1548, 1550$ broad emission line is visible in optical spectra to redshifts exceeding $z \sim 5$. C IV has long been known to exhibit significant displacements to the blue and these ‘blueshifts’ almost certainly signal the presence of strong outflows. As a consequence, single-epoch virial black hole (BH) mass estimates derived from C IV velocity-widths are known to be systematically biased compared to masses from the hydrogen Balmer lines. Using a large sample of 230 high-luminosity ($L_{\text{Bol}} = 10^{45.5} - 10^{48} \text{ erg s}^{-1}$), redshift $1.5 < z < 4.0$ quasars with both C IV and Balmer line spectra, we have quantified the bias in C IV BH masses as a function of the C IV blueshift. C IV BH masses are shown to be a factor of five larger than the corresponding Balmer-line masses at C IV blueshifts of 3000 km s^{-1} and are over-estimated by almost an order of magnitude at the most extreme blueshifts, $\gtrsim 5000 \text{ km s}^{-1}$. Using the monotonically increasing relationship between the C IV blueshift and the mass ratio $\text{BH(C IV)}/\text{BH(H}\alpha\text{)}$ we derive an empirical correction to all C IV BH-masses. The scatter between the corrected C IV masses and the Balmer masses is 0.24 dex at low C IV blueshifts ($\sim 0 \text{ km s}^{-1}$) and just 0.10 dex at high blueshifts ($\sim 3000 \text{ km s}^{-1}$), compared to 0.40 dex before the correction. The correction depends only on the C IV line properties - i.e. full-width at half maximum and blueshift - and can therefore be applied to all quasars where C IV emission line properties have been measured, enabling the derivation of un-biased virial BH mass estimates for the majority of high-luminosity, high-redshift, spectroscopically confirmed quasars in the literature.

Black-hole masses are crucial to understanding the physics of the connection between quasars and their host galaxies and measuring cosmic black hole-growth. At high redshift, $z \gtrsim 2.1$, black hole masses are normally derived using the velocity-width of the C IV $\lambda\lambda 1548, 1550$ broad emission line, based on the assumption that the observed velocity-widths arise from virial-induced motions. In many quasars, the C IV-emission line exhibits significant blue asymmetries (‘blueshifts’) with the line centroid displaced by up to thousands of km s^{-1} to the blue. These blueshifts almost certainly signal the presence of strong outflows, most likely originating in a disc wind. We have obtained near-infrared spectra, including the H α $\lambda 6565$ emission line, for 19 luminous ($L_{\text{Bol}} = 46.5 - 47.5 \text{ erg s}^{-1}$) Sloan Digital Sky Survey quasars, at redshifts $2 < z < 2.7$, with C IV emission lines spanning the full-range

of blueshifts present in the population. A strong correlation between C iv-velocity width and blueshift is found and, at large blueshifts, $>2000 \text{ km s}^{-1}$, the velocity-widths appear to be dominated by non-virial motions. Black-hole masses, based on the full width at half maximum of the C iv-emission line, can be overestimated by a factor of five at large blueshifts. A larger sample of quasar spectra with both C iv and H β , or H α , emission lines will allow quantitative corrections to C iv-based black-hole masses as a function of blueshift to be derived. We find that quasars with large C iv blueshifts possess high Eddington luminosity ratios and that the fraction of high-blueshift quasars in a flux-limited sample is enhanced by a factor of approximately four relative to a sample limited by black hole mass.

3.2 INTRODUCTION

Super-massive black holes (BHs) are found at the centres of most nearby massive galaxies and the BH mass and mass of the host galaxy spheroid are strongly correlated (Ferrarese and Merritt, 2000; Gebhardt et al., 2000; Kormendy and Ho, 2013). Although any underlying causal mechanism(s) responsible for the correlation is yet to be conclusively identified, there is considerable observational and theoretical support for models that involve BH-fuelling, outflows and a ‘feedback’ relationship (e.g. King and Pounds, 2015). The number density of quasars, which evolves strongly with redshift, peaks at redshifts $2 \lesssim z \lesssim 3$ (e.g. Brandt and Hasinger, 2005; Richards et al., 2006b) and the most massive ($M_{\text{BH}} \gtrsim 10^9 M_{\odot}$) present-day BHs experienced much of their growth during this epoch. The star formation rate, which closely follows the cosmological evolution of the quasar luminosity function, also peaks during this epoch (e.g. Boyle and Terlevich, 1998). Quantifying the growth-rate of massive BHs at $2 \lesssim z \lesssim 3$ would therefore help significantly in understanding the role quasars play in galaxy evolution. Reliable estimates of BH masses are a prerequisite for such investigations.

If the line-emitting clouds in the broad line region (BLR) are assumed to be virialized and moving in a potential dominated by the central BH, then the BH mass is simply a product of the BLR size and the square of the virial velocity. The reverberation-mapping technique uses the time lag between variations in the continuum emission and correlated variations in the broad line emission to measure the typical size of the BLR (Peterson, 1993; Peterson, 2014). The full width at half maximum (FWHM) or dispersion (σ ; derived from the second moment) velocity of the prominent broad emission line of H β (4862.7 \AA)¹ is used as an indicator of the virial velocity. Extensive reverberation mapping campaigns have provided accurate BH masses for ~ 50 active galactic nuclei (AGN) at relatively low redshifts and of modest

¹ Vacuum wavelengths are employed throughout the chapter.

luminosity (e.g. Kaspi et al., 2000; Kaspi et al., 2007; Peterson et al., 2004; Bentz et al., 2009; Denney et al., 2010).

Reverberation mapping campaigns have also revealed a tight relationship between the radius of the BLR and the quasar optical (or ultraviolet) luminosity (the $R - L$ relation; e.g. Kaspi et al., 2000; Kaspi et al., 2007). This relation provides a much less expensive method of measuring the BLR radius, and large-scale studies of AGN and quasar demographics have thus become possible through the calibration of single-epoch virial-mass estimators using the reverberation-derived BH masses (e.g. Greene and Ho, 2005; Vestergaard and Peterson, 2006; Vestergaard and Osmer, 2009; Shen et al., 2011; Shen and Liu, 2012; Trakhtenbrot and Netzer, 2012). The uncertainties in reverberation mapped BH masses are estimated to be ~ 0.4 dex (e.g. Peterson, 2010), and the uncertainties in virial masses are similar (e.g. Vestergaard and Peterson, 2006). Since the structure and geometry of the BLR is unknown, a virial coefficient f is introduced to transform the observed line-of-sight velocity inferred from the line width in to a virial velocity. This simplification accounts for a significant part of the uncertainty in virial BH masses (in addition to, for example, describing the BLR with a single radius R and scatter in the $R - L$ relation; Shen, 2013). Furthermore, if the BLR is anisotropic (for example, in a flattened disk; e.g. Jarvis and McLure, 2006) then the line width will be orientation-dependent (e.g. Runnoe et al., 2013b; Shen and Ho, 2014; Brotherton et al., 2015).

Single-epoch virial BH mass estimates using $H\beta$ are possible up to redshifts $z \sim 0.7$, and the technique has been extended to redshifts $z \sim 1.9$ via the calibration of the broad $Mg\ II\lambda\lambda 2796, 2803$ emission line (McLure and Jarvis, 2002; Onken and Kollmeier, 2008; Wang et al., 2009; Rafiee and Hall, 2011). At redshifts $z \gtrsim 2$, however, ground-based statistical studies of the quasar population generally have no access to the rest-frame optical and near-ultraviolet spectral regions. The $C\ IV\lambda\lambda 1548, 1550$ emission doublet is both relatively strong in the majority of quasars and visible in modern optical spectra, such as those provided by the Sloan Digital Sky Surveys, 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.

The luminosities of quasars at redshifts $z \gtrsim 2$ are much greater than the majority of AGN at lower redshifts for which reverberation mapping results are available. Therefore, the reliability of the existing calibration involving $C\ IV$ FWHM velocity measurements and ultraviolet luminosity is not established definitively when extrapolating to high-redshifts and luminosities. While some authors have found good agreement between BH mass-estimates based on $C\ IV$ and $H\beta$ (e.g. Vestergaard and Peterson, 2006; Assef et al., 2011; Tilton and Shull,

2013), others have questioned the consistency (e.g. Baskin and Laor, 2005a; 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 displacements to the blue (Gaskell, 1982) and 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). 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, that should allow a BH-mass estimate to be derived. 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). Certainly, 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). As a consequence, BH-masses derived from C IV emission line velocity-widths are known to be systematically biased compared to masses from the Balmer lines (e.g. Shen et al., 2008; Shen and Liu, 2012).

In recent literature, attempts have been made to minimise the influence of the systematic non-virial contribution to the C IV emission on estimates of the BH mass. Strategies include (i) significantly reducing the dependence of the derived masses on the emission-line velocity width (e.g. from the V^2 dependence predicted assuming a virialized broad line region to just $V^{0.56}$ in Park et al. 2013; see also Shen and Liu 2012), (ii) adopting a measure of emission-line velocity-width that is relatively insensitive to changes in the core of the emission-line profile (e.g. Denney et al., 2013) and (iii) estimating the amplitude of the non-virial contribution to the C IV emission-line via comparison with other ultraviolet emission lines (e.g. Si IV+O IV $\lambda 1400$ in Runnoe et al. 2013a and Brotherton et al. 2015). The increased number of high-quality spectra of quasars where information on both the Balmer lines in the rest-frame optical and C IV in the ultraviolet is available enables a rather different approach. Specifically, to investigate whether, using the properties of the C IV emission line itself, it is possible to reduce, or even remove, the systematic bias in the BH-mass estimates.

In this chapter we analyse the spectra of 230 high-luminosity ($10^{45.5} - 10^{48}$ erg s $^{-1}$), redshift $1.5 < z < 4.0$ quasars for which spectra of the hydrogen Balmer emission lines and the C IV emission line exist. A direct comparison of the emission-line velocity widths is therefore

possible, allowing us to determine a highly effective empirical correction to the C IV emission line velocity width as a function of the C IV emission line blueshift.

3.3 C IV BLUESHIFTS IN THE QUASAR POPULATION

Recognition that the C IV emission line in quasars can exhibit significant asymmetric structure, with an excess of flux to the blue of the predicted rest-frame transition wavelength, extends back to Gaskell, (1982). Significant progress in understanding the relationship between changes in C IV-emission shape and quasar properties has come about through studies in which near-infrared spectra of the hydrogen Balmer lines have been obtained. Such studies typically involve samples of modest size and the location of the Balmer lines provides a reliable estimate of the quasar systemic redshifts; recent examples include Shen and Liu, (2012) and Marziani et al., (2016). In Section ?? we adopt the same approach to estimate systemic redshifts for the quasar sample presented here with near-infrared spectra. However, 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 quasar population as a whole.

3.3.1 *Quasar systemic redshifts*

Historically, the parametrisation of the C IV emission-line properties for quasars in large surveys has not proved straightforward because the C IV emission line has itself been used in the determination of the quasar redshifts. The SDSS provided the first catalogue of tens of thousands of redshift $z > 1.6$ quasars with spectra of adequate velocity resolution and S/N that effective statistical studies of the rest-frame ultraviolet emission-line properties, including line-shape, have proved possible.

The comprehensive compilation of quasar properties for the SDSS DR7 quasars by Shen et al., (2011) provides a natural starting point for population studies. In Fig. 3.18a 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). The grey points show all SDSS DR7 quasars for which measurements exist and the orange circles show the 19 quasars with near-infrared spectra presented in this paper. A strong trend in the blueshift values as a function of line EW is not evident in Fig. 3.18; 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

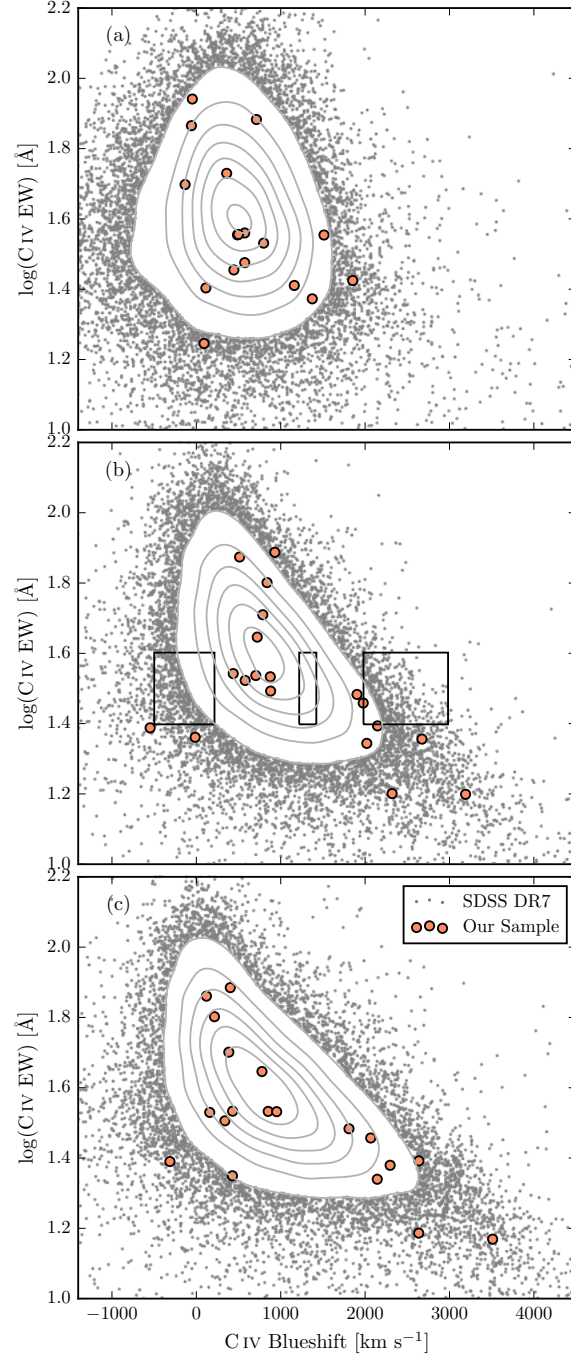


Figure 3.1: Rest-frame EW versus blueshift of the broad C IV-emission line for 32,157 SDSS DR7 quasars at $1.6 < z < 3.0$ (grey) and our sample (orange). Panel (a) uses C IV line parameters from Shen et al., (2011) and SDSS pipeline systemic redshifts. Panels (b) and (c) use systemic redshifts from Hewett and Wild, (2010) and Allen & Hewett (2016, in preparation) respectively, and C IV line measurements described in Sec. 3.3.2. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator. The three rectangles in panel (b) show the regions of parameter space used to generate the composite spectra shown in Fig. 3.2.

other than the C IV-line itself. Figure 3.18b shows SDSS DR7 quasars in the same C IV parameter space as Figure 3.18a, but now using Hewett and Wild, (2010) redshifts. The improved redshift estimates are predominantly responsible for the differences seen in Fig. 3.18a and b; the appearance in Fig. 3.18b of the extension to high blueshift for quasars with low C IV EW is particularly evident.

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 III λ 1892 and C III λ 1908 emission lines in the rest-frame ultraviolet. As a consequence, the redshifts from Hewett and Wild, (2010) still suffer from systematic errors that are correlated with the shape, and particularly the blueshift, of the C IV emission line. The nature of the systematic variations in the quasar ultraviolet SEDs are such that for quasars with close-to symmetric C IV profiles and line centroids close to the systemic redshift, the Hewett and Wild, (2010) redshifts result in C IV blueshifts that are overestimated by a few hundred km s^{-1} , whereas, for quasars with strong blue-asymmetric C IV profiles and line centroids displaced significantly to the blue of the systemic redshift, the C IV blueshifts are underestimated by, in the most extreme cases, up to 1200 km s^{-1} .

Figure 3.18c shows the C IV emission line parameters calculated using a new redshift-estimation algorithm (Allen & Hewett 2016, in preparation) that takes account of the quasar SED variations, producing redshifts independent of the large systematic shape changes seen in the C IV emission line. The low-ionization emission lines visible in the rest-frame ultraviolet (over wavelengths from Mg II λ 2796,2803 down to the O I λ 1304+Si II λ 1307 blend) using the new redshift-algorithm are located at rest-frame wavelengths in excellent agreement with the systemic redshift defined using the rest-frame narrow-line optical O III λ 4960,5008 and broad-line H β and H α .

The systematic trends seen in Fig. 3.18b, in particular the extension to high blueshift at low C IV EW, become more apparent in Fig. 3.18c, 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.

3.3.2 C IV emission line blueshift measurements

The differences in the distribution of C IV emission line properties seen in the three panels of Fig. 3.18 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 follow below in Section ??.

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 median² values of the flux in two continuum windows at 1445–1465 and 1700–1705Å (the same wavelengths as adopted by Shen et al., (2011)). The C iv emission line is taken to lie within the wavelength interval 1500–1600Å, 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, λ_{half} , is recorded and the blueshift (in km s^{-1}) defined as $c \times (1549.48 - \lambda_{\text{half}}) / 1549.48$ where c is the velocity of light and 1549.48Å is the rest-frame wavelength for the C iv doublet³. Positive blueshift values indicate an excess of emitting material moving towards the observer and hence out-

² The median is used to improve the robustness of the continuum estimate from the relatively small wavelength intervals.

³ The adopted C iv rest-frame wavelength assumes an optically thick BLR, in which case the contribution from each component is equal. Adopting a 2:1 ratio (appropriate for an optically thin BLR) changes the blueshifts by $\sim 80 \text{ km s}^{-1}$.

flowing from the quasar. Hewett and Wild, (2010) redshifts are used to define the quasar rest-frame.

3.3.3 Sample selection - C IV properties

The primary aim of the paper is to investigate the potential systematic effects on the C IV-emission based BH masses for quasars with large, $\gtrsim 1200 \text{ km s}^{-1}$, C IV blueshifts, using the properties of the H α emission line to provide BH-mass estimates for the objects unbiased by non-virial contributions to the emission-line profile. The orange symbols in Fig. 3.18 show the C IV parameters of our quasar targets for which near-infrared spectra of adequate S/N were obtained. These quasars were selected using our non-parametric blueshift measures (based on the Hewett and Wild, (2010) redshifts). The sample of 19 quasars spans the full dynamic range in C IV-parameters based on the Hewett and Wild, (2010) systemic redshifts and the coverage is in fact even more complete when using the forthcoming SED-independent redshifts from Allen & Hewett (2016, in preparation). As is evident from the sparsity of quasars with large C IV blueshifts when the SDSS pipeline systemic redshifts are used (Fig. 3.18a), improvements in the estimation of systemic redshifts from ultraviolet spectra have been a crucial factor in allowing us to reliably select a sample of quasars with a range of C IV blueshifts. In subsequent sections we re-derive the systemic redshifts and C IV blueshifts for this sample using parametric fits to the H α and C IV emission (the former from our near-infrared observations). Thus, while the systematic trends in BH masses inferred from measurements of the C IV emission line depend on the distribution of C IV emission line properties within the quasar population, the results of our analysis of the H α and C IV emission line properties are independent of the redshifts used to produce the panels in Fig. 3.18.

3.3.4 Relation to virial BH mass estimates

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). The systemic redshift is often assumed to be given by the peak of the C IV emission, regardless of whether there is evidence that the line is shifted or not. Figure 3.2 shows the shape of the C IV-emission in composite spectra constructed from SDSS DR7 quasars with $\text{EW}(\text{C IV})=25\text{--}40 \text{ \AA}$, as a function of C IV blueshift. Quasars classified as BALs, or possessing strong associated absorbers have been excluded, and the composite-spectra shown are derived using an

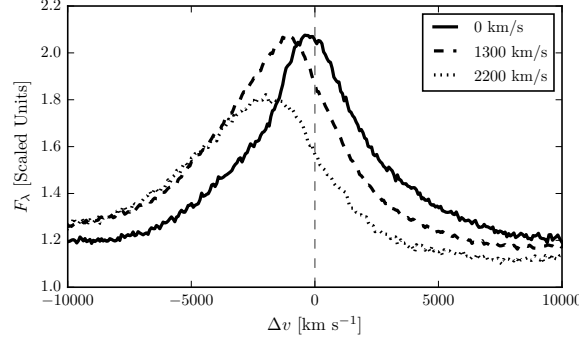


Figure 3.2: Composite spectra of the C IV-emission line as a function of C IV blueshift for SDSS DR7 quasars. The quasars contributing to each composite are indicated in Fig. 3.18b. Virtually the entire C IV-profile appears to shift blueward and the change in line shape is not simply an enhancement of 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.

arithmetic mean of a minimum of 200 spectra at each blueshift. The blueshifts and EWs of the quasars contributing to each of the composites are indicated by the boxes in Fig. 3.18b. The profiles show how, at large values of blueshift ($\gtrsim 2000 \text{ km s}^{-1}$) the C IV-profile is displaced to the blue by amounts comparable to the FWHM of the profile.

A possible origin of the blueshifts is the presence of a disc-wind (see Gallagher et al., 2015; Higginbottom and Proga, 2015, for recent papers) but, irrespective of the physical origin of the high-blueshift C IV-profiles, measures of the emission-line ‘width’ do not relate simply to virialized motions of the emitting gas under the gravitational influence of the BH. On the other hand, Denney et al., (2013) point out that any radiatively driven wind will have a velocity comparable to the escape velocity, i.e. approximately twice the virial velocity. Even if dominated by an outflow component, the C IV line width might therefore still be expected to relate to the BH mass.

The paper is structured as follows. Section 2 presents the extensive set of near-infrared spectra that, combined with optical spectra of the quasar ultraviolet rest-frame, provides our spectroscopic catalogue. The scheme adopted to calculate emission-line parameters, which draws heavily on the methodology of Shen et al., (2011), Shen and Liu, (2012) and Shen, (2016), is described in Section 3. The observational results, where the emission-line properties of the Balmer lines and the C IV emission are compared and a quantitative relationship derived, are included as Section 4. Then, in Section 5, the practical application of the new BH-mass estimation formula and the extent of remaining uncertainties are discussed, and our scheme is compared to others presented in the literature. Finally, we summarise the main

points of the paper and highlight forthcoming improvements to systemic redshift estimates for quasars that should improve the accuracy of BH-masses from rest-frame ultraviolet quasar spectra even further. Throughout this paper we adopt a Λ CDM cosmology with $h_0 = 0.71$, $\Omega_M = 0.27$, and $\Omega_\Lambda = 0.73$. Vacuum wavelengths are used for both rest-frame ultraviolet and optical features. Unless otherwise stated, optical (i.e. SDSS) magnitudes are given in the AB system and infrared magnitudes in the Vega system, following the conventions of the original surveys.

Denney, (2012) presented evidence that the interpretation of the FWHM velocity of the C IV-emission being due primarily to virial motions within the quasar BLR requires care. Specifically, both a low-velocity core component and a blue excess to the C IV-emission, both of which do not reverberate, can be present and Denney, (2012) proposes that a contribution from an accretion disc wind or from a more distant narrow emission line region is important.

Changes in the C IV blueshift and equivalent width are correlated with changes in the velocity widths and strengths of other optical and ultra-violet emission lines. In the spectra of lower-redshift AGN, the FWHM of the broad H β emission line and the relative strengths of optical Fe II and H β have been identified as the features responsible for the largest variance in the population. These parameters form part of ‘Eigenvector 1’ (EV1), the first eigenvector in a principal component analysis which originated from the work of Boroson and Green, (1992). The underlying driver behind EV1 is thought to be the Eddington ratio (e.g. Sulentic et al., 2000; Shen and Ho, 2014). **sulentic00** proposed a two-population model to classify AGN by their EV1 properties. In this scheme AGN with $\text{FWHM}(\text{H}\beta) < 4000 \text{ km s}^{-1}$ and $\text{FWHM}(\text{H}\beta) > 4000 \text{ km s}^{-1}$ are classified as population A and B objects respectively, although there is a continuous distribution of parameter values across this divide. Sulentic et al., (2007) added a measure of the C IV asymmetry to EV1, and found a strong association between blue-asymmetry and their population A quasars.

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. Runnoe et al., (2013a) found the scatter between the C IV and H β line widths to be correlated with 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. Both authors used these correlations to propose empirical corrections to the C IV line width which can improve the consistency between C IV and H β -based virial BH mass estimates. In fact, the shape, peak flux relative to the 1400 Å blend, and blueshift of C IV all correlate with one another and with other parameters in EV1. Therefore, EV1 provides a useful context for understanding systematic trends in C IV velocity widths, and hence virial BH masses.

Currently, the number of reverberation mapped quasars is both small (~ 50 quasars; Park et al., 2013) and, as highlighted by Richards et al., (2011), 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. Much more complete coverage of the C IV-emission properties within the population of luminous quasars will come from the new SDSS-IV reverberation mapping project (Shen et al., 2015) but, for now, additional direct comparison of C IV-emission and low-ionization emission-line properties in the same quasars offers a way forward.

Near-infrared spectra, including the H α emission line, for a sample of 19 quasars, at redshifts $2.0 < z < 2.7$, have been obtained to complement existing SDSS optical spectra covering the C IV emission line. The 19 quasars were chosen to include a broad range of C IV line blueshifts. Our aim is to directly test the reliability of C IV-based BH mass estimates at high redshift for objects with a diverse range of C IV-line shapes. In particular, we will investigate potential systematic effects on the C IV-emission based BH masses for quasars with large, $\gtrsim 1200 \text{ km s}^{-1}$, C IV blueshifts, using the properties of the H α emission line to provide BH-mass estimates for the objects unbiased by non-virial contributions to the emission-line profile. Examining higher redshifts, our work complements other studies which attempt to improve the reliability of BH mass estimates which use the C IV line (e.g. Runnoe et al., 2013a; Denney, 2012). However, the range of C IV blueshifts in our sample is significantly more extended, which will allow us to study systematic biases in C IV-based virial BH masses more directly, i.e. as a function of the C IV blueshift. Established relations to derive BH masses from emission-line properties are employed but an advantage of our approach is that C IV and H α can be directly compared as a function of C IV-emission line shape. In Section ?? we describe how quasars were selected for rest-frame optical spectroscopy, before reviewing the C IV emission-line properties of high luminosity quasars in the SDSS surveys in Section 3.3. The new near-infrared spectroscopic observations are outlined in Section ?. With spectroscopic coverage of the C IV and H α emission lines available, the procedures used to quantify the line parameters are described in Section ?. The results of the spectral analysis in the context of BH masses and C IV-blueshift are presented in Section ?, with a discussion of the significance of the systematic trends identified included in Section 3.8. The paper concludes with a summary of the main conclusions in Section 3.9. Throughout this paper we adopt a Λ CDM

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

3.4 QUASAR SAMPLE

The aim of this work is to measure empirically the systematic bias in C IV-based virial BH mass estimates for high- z quasars as a function of the C IV emission-line blueshift. The basis for the C IV blueshift based correction is a large sample of quasars where it is possible to make a direct comparison of the C IV line-width with the line-width of the low-ionisation Balmer lines H α and H β , which are believed to provide reliable proxies for the virial velocity. Such an approach has not been possible hitherto as spectra that cover both the observed-frame optical (where the redshifted C IV appears) and near-infrared (where H β and H α lie) are required.

We have compiled a sample of 307 quasars at redshifts $1.5 < z < 4$ with both optical and near-infrared spectra to enable such a comparison to be performed. Reliable emission line properties were measured for 230 quasars (Section 3.5.5), with 164 possessing H α line measurements and 144 H β line measurements. The sample is considerably larger than previous studies of the rest-frame optical spectra of high- z quasars (e.g. Shen and Liu, 2012). As we demonstrate in Section 3.7.3, the quasars have C IV blueshifts of up to $\sim 5000 \text{ km s}^{-1}$, and span the full range observed in the population. Part of this data set has been taken from the literature, but a substantial fraction is presented here for the first time. The infrared spectra were acquired using several different telescope and spectrograph combinations and the contributions from each telescope/spectrograph, along with the instrumental configurations, are summarised in Table 4.1. We have sub-divided our sample into two overlapping groups: quasars with reliable H α line measurements (the ‘H α sample’) and quasars with reliable H β measurements (the ‘H β sample’).

In Fig. 4.1 we show the luminosities and redshifts of the quasar sample relative to the redshift-luminosity distribution for the Sloan Digital Sky Survey (SDSS; York et al., 2000) Seventh Data Release (DR7; Schneider et al., 2010). Our sample spans a redshift range $1.5 < z < 4.0$ and a bolometric luminosity range $10^{45.5} - 10^{48} \text{ erg s}^{-1}$. Spectra were obtained within one or more of the JHK pass-bands and the gaps in our sample coverage at $z \sim 1.8$ and $z \sim 3$ are due to the presence of atmospheric absorption. Obtaining near-infrared spectra of adequate resolution and signal-to-noise ratio (S/N) of even moderately bright quasars remains resource intensive. As a consequence, at fixed redshift, the luminosities of the quasars are brighter than the

Table 3.1: The numbers of quasars with reliable H α and H β line measurements, and the spectrographs and telescopes used to obtain the near-infrared spectra

Spectrograph	Telescope	H α Sample	H β 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

average luminosity of the SDSS sample, although the dynamic range in luminosity is a full 1.5 decades.

Below, we present the key elements of the observations of the six quasar sub-samples that make up the full 230-quasar catalogue.

3.4.1 Optical observations

In the previous sections, we described the infrared spectra of the 230 quasars making up our full spectroscopic catalogue. We will now describe the companion optical spectra, which provide coverage of the C iv emission.

Optical SDSS DR7 spectra are employed for 70 quasars in the full catalogue. The SDSS DR7 catalogue contains moderate resolution $\sim 3800 - 9180 \text{ \AA}$ spectra for 105,783 quasars. Shen et al., (2011) have compiled a catalogue of properties for the SDSS DR7 quasars including, at $z > 1.5$, measurements of the broad C iv $\lambda\lambda 1548.2, 1550.8$ emission line. The SDSS DR7 spectra are moderate resolution ($R \simeq 2000$) and S/N ($S/N \simeq 20$) and cover the observed-frame wavelength interval $\sim 3800 - 9180 \text{ \AA}$. Many of the quasars in the SDSS DR7 catalogue have been re-observed as part of the Sloan Digital Sky Survey-III: Baryon Oscillation Spectroscopic Survey (SDSS-III/BOSS; Dawson et al., 2013). As the BOSS-spectra typically have higher S/N than the SDSS DR7 spectra, we have used the BOSS spectra when available (126 quasars). We also use high-resolution optical spectra taken with VLT/UVES (11 quasars) and VLT/XSHOOTER (8 quasars), and Hamburg/ESO spectra for a further 15 quasars. The Hamburg/ESO opti-

cal spectra have a typical $\sim 400 \text{ km s}^{-1}$ spectral resolution and $S/N \gtrsim 10$ per pixel.

In summary, we have assembled a sample of 230 luminous, high- z quasars with optical and near-infrared spectra. This will allow us to directly compare virial BH mass estimates based on the C IV line-width with estimates based on the line-widths of the low-ionisation Balmer lines $H\alpha$ and $H\beta$.

3.5 SPECTRAL MEASUREMENTS

Conventionally, single-epoch virial estimates of the BH mass are a function of the line-of-sight velocity width of a broad emission line and the quasar luminosity. The velocity width is a proxy for the virial velocity in the broad line region (BLR) and, as revealed in reverberation-mapping studies, the luminosity is a proxy for the typical size of the BLR (the $R-L$ relation; e.g. Kaspi et al., 2000; Kaspi et al., 2007). Most reverberation mapping campaigns have employed $H\beta$ time-lags and velocity widths, but the line-widths of $H\alpha$ and $\text{Mg II } \lambda 2800$ have been shown to yield consistent BH masses (e.g. McLure and Jarvis, 2002; Greene and Ho, 2005; Onken and Kollmeier, 2008; Shen et al., 2008; Wang et al., 2009; Rafiee and Hall, 2011; Mejía-Restrepo et al., 2016). In Section 3.6.1 we verify that this is the case for the 99 quasars in our sample with measurements of both $H\alpha$ and $H\beta$ lines.

At redshifts $z > 2.2$, where the hydrogen Balmer lines and Mg II are no longer accessible in many optical spectra, the C IV $\lambda 1550$ emission doublet has routinely been used to provide estimates of the virial velocity (e.g. Shen et al., 2011). As has long been recognised (Gaskell, 1982; Tytler and Fan, 1992) the C IV emission line in many quasars includes contributions from gas that does not straightforwardly relate to virial motions within a stable BLR. A number of studies (e.g. Shen et al., 2008; Richards et al., 2011) have shown that the amplitude of the systematic shift of the C IV emission to shorter wavelengths (relative to the systemic velocity) is strongly correlated with the properties of the emission-lines and the overall spectral energy distributions (SEDs).

In our work, a robust measure of the C IV emission-line ‘blueshift’ provides the basis for the corrected C IV velocity-width measurements, and hence BH masses. The effectiveness of the scheme is validated via a direct comparison of the C IV velocity-widths to the Balmer emission velocity-widths in the same quasars. Our process is as follows. First, an accurate measure of the quasar’s systemic redshift is required, for which we adopt the centre of the Balmer emission, where the centre, 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 III] $\lambda\lambda 4960, 5008$ doublet for the subset of spectra where both are available. Second, the blueshift of the C IV emission line is determined. Again, we adopt the line centroid to provide a robust measure of the C IV emission blueshift. The blueshift (in km s^{-1}) is defined as $c \times (1549.48 - \lambda_{\text{half}}) / 1549.48$ where c is the velocity of light and 1549.48 \AA is the rest-frame wavelength for the C IV doublet, assuming equal contribution from both components. Positive blueshift values indicate an excess of emitting material moving towards the observer and hence out-flowing from the quasar.

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

To minimise the impact of the finite S/N of the quasar spectra and the presence of absorption features superposed on the broad emission lines we first fit a parametric model to the continuum and the emission lines. The purpose of the parametric fits is, however, simply to provide higher S/N representations of the emission features. 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.

Virial BH mass estimators are calibrated using either the FWHM or dispersion (σ ; derived from the second-moment velocity) of a broad emission line (e.g. Vestergaard and Peterson, 2006; Park et al., 2013). Complications which are encountered when measuring line widths include how to model the ‘continuum’ flux, where to define the limits of the line emission, and how to deal with absorption. All of these issues are exacerbated when working with low S/N data (see Denney et al., 2013, for a discussion). In Section 3.3.2 we measured the blueshift of C IV for tens of thousands of SDSS DR7 quasar spectra. This allowed us to quantify the distribution of C IV blueshift values and hence select a subset for near-infrared observations which have C IV blueshifts spanning the full range of this distribution (Fig. 3.18). A non-parametric scheme was employed because, in comparison to recipes involving the fitting of multiple Gaussian (or other parametric) profiles, it was found to be more robust and less sensitive to the range of S/N among the spectra and to the presence of narrow

absorption systems within the C IV-emission profile. In this section we will use a different approach, and measure the line properties by fitting a parametric model to the data. When working with a small number of spectra, it is possible to use careful mask-definition to minimise the effect of absorption features on the profiles used for the parametrization. The purpose of the model fits is purely to best represent the intrinsic line profile, and no physical meaning is attached to the individual model components. We will now describe the parametric model and fitting procedure used for each emission line. The models were fit using a standard variance-weighted least squares minimisation procedure employing the Levenberg-Marquardt algorithm. Prior to the fit, the spectra were visually inspected and regions significantly affected by absorption were masked and excluded.

3.5.1 C IV

The parametrization of the C IV emission line is identical to the one described in Paper I. 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 (1549.98 Å, assuming equal contributions from both components). The parametric model is ordinarily fit within the wavelength interval 1500-1600 Å (corresponding to approximately $\pm 10\,000 \text{ km s}^{-1}$ from the rest-frame transition wavelength), a recipe that is commonly adopted (e.g. Denney et al., 2013). The line-window was extended if more than 5 per cent of the total flux in the profile was present blueward of the short wavelength limit. Narrow absorption features, which are frequently found superimposed on C IV emission (see, for example, the C IV profile of J0942+3523 in Fig. 3.5), were masked out during the fit.

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

In Paper I we found 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 pa-

rameters are also in good agreement with Shen et al., (2011), who employ a multi-Gaussian parametrization. In the Appendix we demonstrate that the mean difference between our FWHM measurements and the measurements of Shen et al., (2011) is just 200 km s^{-1} for the quasars common to both samples, which is much too small to have any significant effect on our results.

We first measure and subtract the local continuum emission, by fitting a power-law to two windows on either side of the line emission, as described in Section 3.3.2. For a small number of objects, absorption features, or artefacts, in the spectrum necessitated modest adjustments to the window extents, which are specified in Table ?? . The continuum-subtracted spectra are then transformed from wavelength units into units of velocity relative to the rest-frame line-transition wavelength for the C IV doublet (1549.48 \AA , assuming equal contributions from both components). The parametric model is ordinarily fit within the same $1500\text{--}1600 \text{ \AA}$ window used in Section 3.3.2, which corresponds to approximately $\pm 10\,000 \text{ km s}^{-1}$ from the rest-frame transition wavelength. The line-window was extended if significant flux in the profile was present blueward of the short wavelength limit. The adopted line-fitting windows, in units of velocity from the rest-frame transition wavelength, are given in Table ?? .

To fit the C IV profile we employed 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 in the GH polynomial model, but in many cases a lower order was sufficient; the polynomial order used for each line is given in Table ?? . It is also a common practice to fit the C IV emission profile with two or three Gaussian components (e.g. Shen et al., 2011). We opted to use a GH-polynomial model primarily because it provided a significantly better fit to the most blueshifted and asymmetric C IV line (in SDSSJ0858+0152). Figure 3.3 shows how a model with three Gaussian components underestimates the flux in the blue wing and overestimates the flux in the red wing of the line profile. Using the Gaussian model rather than the GH polynomial changes the FWHM, line dispersion, and blueshift by -3, -3, and 10 per cent respectively. We have highlighted SDSSJ0858+0152 because, of all the objects in our sample, the choice of model leads to the largest change in C IV line parameters. Even in this case, however, the differences are modest.

For every other C IV line in our sample we found only marginal differences in our best-fit line parameters when, rather than using a GH polynomial model, the C IV emission was fit using a composite model of up to three Gaussians. Our best-fit parameters are also in good agreement with Shen et al., (2011), who employ a multi-Gaussian

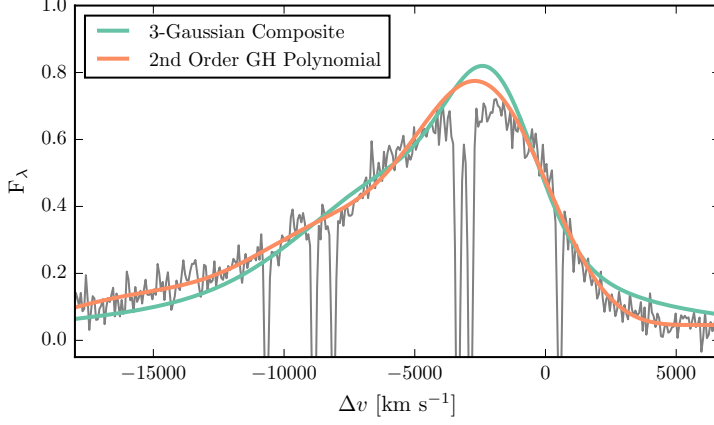


Figure 3.3: 2nd-order Gauss-Hermite (GH) polynomial and 3-component Gaussian fit to the C IV-emission line of SDSSJ0858+0152, which is the most blueshifted in our sample. We derive line parameters from the GH polynomial fit; using the Gaussian model changes the FWHM, line dispersion, and blueshift by -250, -150, and 500 km s⁻¹ respectively. For the C IV lines of all other quasars in our sample the GH polynomial and Gaussian models provide equally good fits.

parametrization⁴. The scatter between the Shen et al., (2011) results and our own is 0.1 dex about the one-to-one relation and, as expected, is larger for lines with smaller EWs.

In this section we verify the accuracy of our parameter estimation scheme by comparing our FWHM measurements to measurements published in Shen et al., 2011, Shen and Liu, 2012 and Shen, 2016.

The parametric model we fit to the H β /[O III] emission region was very similar to the model employed by Shen, (2016). In Fig. 3.4a we plot our H β 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.

In Fig. 3.4b we plot our H α 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, although the scatter is larger than for the H β comparison (0.07 dex), no significant systematic bias.

Finally, in Fig. 3.4c 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). Shen et al., (2011) fit the C IV line profile with a composite model comprising up to three Gaussian components, whereas we used (up to order six) Gauss-Hermite polynomials.

⁴ The Shen et al., (2011) parameters are derived from the SDSS DR7 spectra, whereas 16 out of 19 of our fits are to higher S/N BOSS DR12 spectra.

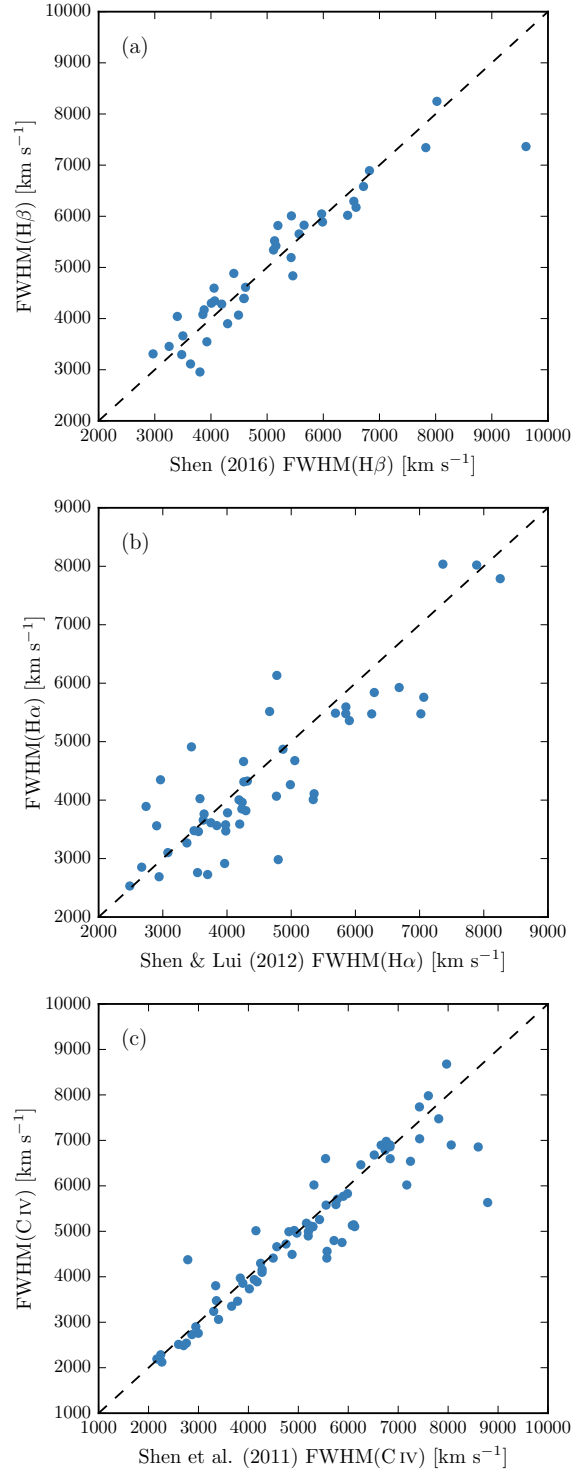


Figure 3.4: Demonstration of the effectiveness of our line parameter estimation scheme via a comparison of (a) the $H\beta$ FWHM with Shen, (2016), (b) the $H\beta$ FWHM with Shen and Liu, (2012), and (c) the C IV FWHM with Shen et al., (2011).

Nevertheless, there is a very strong agreement between our measurements, with a scatter of 0.05 dex.

3.5.2 $H\alpha$

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

The emission-line profiles of both $H\beta$ and $H\alpha$ frequently include a significant narrow component from the physically more extended narrow line region (NLR). Additional Gaussian components were included in our parametric model to fit the narrow component of $H\alpha$ as well as $[\text{N II}]\lambda\lambda 6548, 6584$ and $[\text{S II}]\lambda\lambda 6717, 6731$. This resulted in a better fit to the observed flux in 50 per cent of cases. We impose a 1200km s^{-1} upper limit on the FWHM of all narrow lines and the amplitudes of all components must be non-negative. The relative flux ratio of the two $[\text{N II}]$ components is also fixed at the expected value of 2.96. In 70 per cent of the spectra the $[\text{O III}]\lambda\lambda 4960, 5008$ doublet is detected at moderate S/N in the $H\beta$ region. In these cases the peak of the $[\text{O III}]$ is used to fix the velocity offsets and the FWHMs of the narrow line components in the $H\alpha$ region. For spectra where the $[\text{O III}]$ doublet does not constrain the velocity and FWHM accurately, the narrow emission in the $H\alpha$ and $H\beta$ regions are fitted independently but, for each region, the individual narrow-line velocity offsets and the FWHMs are constrained to be identical. In these objects the narrow line contribution is generally weak, and so does not have a large effect on the line parameters we measure for the broad component.

The model described above is very similar to the one described in Shen, (2016), Shen and Liu, (2012) and Shen et al., (2011), the only major difference being that we do not fit the $H\alpha$ and $H\beta$ emission regions simultaneously. In Appendix ??, we compare our $H\alpha$ line measurements for the subset of our sample taken from Shen, (2016) and Shen and Liu, (2012), and find a scatter of just 0.07 dex.

We employ the same continuum subtraction and fitting method as for C IV, with the continuum and fitting windows as given in Table ?. We adopt a rest-frame transition wavelength of 6564.89 Å to transform wavelengths into equivalent Doppler velocities. We used a simple model with up to three broad Gaussian components to fit the $H\alpha$ emission line. We opted against parametrizing the $H\alpha$ line using a GH polynomial because the extra degrees of freedom in this model

did not improve the quality of the fits⁵. Upon inspection of the residuals from the fit, we also found no evidence that additional model components for narrow $H\alpha$, $N\text{II}\lambda\lambda 6548, 6584$ and $S\text{II}\lambda\lambda 6717, 6731$ were required. Furthermore, narrow $O\text{III}\lambda\lambda 4960, 5008$ emission is relatively weak in these spectra.

The sole exception is the $H\alpha$ line in the spectrum of SDSSJ0738+2710. In addition to having the narrowest $H\alpha$ line, this spectra also has the strongest narrow $O\text{III}$ component ($EW = 63 \text{ \AA}$), which suggests that a contribution from the narrow-line region might be important. Introducing a single Gaussian for the narrow emission, while retaining a double Gaussian for the broad emission, the FWHM of the broad component increases to 3400 km s^{-1} (compared to 1580 km s^{-1} without the narrow component). For consistency, the parameters quoted in Table ?? are from the model with no narrow component. However, because the properties derived from the emission line width (the BH mass and the mass-normalised accretion rate) are strongly biased by the probable contribution from the narrow-line region, SDSSJ0738+2710 is excluded from the analysis in Section ??.

3.5.3 $H\beta$ and $[O\text{III}]$

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

Open science. Provide: Fitting region Continuum region Model Redshift used in fit Number of Gaussians / order

3.5.4 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 re-

⁵ The emission line parameters and subsequent analysis do not depend on whether line parameters from multiple-Gaussian or GH-polynomial model fits are used.

gions significantly affected by absorption or of low S/N were masked out.

In Fig. 3.5 we present our parametric fits to the C IV, H α and H β emission lines in a handful of quasars, which have been chosen to illustrate the range of spectrum S/N and line shapes in the sample. The Doppler velocities have been shifted so that the H α emission line centroid is at 0 km s⁻¹. The y-axes of the data-minus-model residual plots have deliberately been scaled by the spectrum flux errors. The mean reduced chi-squared values in our H α , H β 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 3.2 includes the line parameters of our best-fitting model for each line. The reported line-width measures are corrected for instrumental broadening by subtracting the resolution of the spectrograph in quadrature. The spectrograph resolutions, which we estimate from the line widths in the observed sky spectra, range from 25 km s⁻¹ for XSHOOTER to 477 km s⁻¹ for the low-resolution LIRIS grism and are therefore small relative to the quasar broad line widths.

3.5.5 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 entirely 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 3.3 and the distribution in redshift and luminosity is shown in Fig. 3.6. Unsurprisingly, there is a preferential removal of intrinsically faint quasars, whose spectra can be of poorer S/N, and a loss of quasars at redshifts $z \sim 2.6$ where the H α emission falls at the edge of the K-band. H β is much weaker than H α , and the H β spectra are

Table 3.2: 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 journal.

	Units	Description
NAME		Catalogue name
FWHM_BROAD_HA	km s^{-1}	FWHM of broad H α line
FWHM_BROAD_HA_ERR	km s^{-1}	
SIGMA_BROAD_HA	km s^{-1}	Dispersion of broad H α line
SIGMA_BROAD_HA_ERR	km s^{-1}	
Z_BROAD_HA		Redshift from broad H α line
FWHM_BROAD_HB	km s^{-1}	FWHM of broad H β line
FWHM_BROAD_HB_ERR	km s^{-1}	
SIGMA_BROAD_HB	km s^{-1}	Dispersion of broad H β line
SIGMA_BROAD_HB_ERR	km s^{-1}	
Z_BROAD_HB		Redshift from broad H β line
FWHM_CIV	km s^{-1}	FWHM of C iv doublet
FWHM_CIV_ERR	km s^{-1}	
SIGMA_CIV	km s^{-1}	Dispersion of C iv doublet
SIGMA_CIV_ERR	km s^{-1}	
BLUESHIFT_CIV_HA	km s^{-1}	Blueshift of C iv relative to H α
BLUESHIFT_CIV_HA_ERR	km s^{-1}	
BLUESHIFT_CIV_HB	km s^{-1}	Blueshift of C iv relative to H β
BLUESHIFT_CIV_HB_ERR	km s^{-1}	
LOGL ₅₁₀₀	erg s^{-1}	Luminosity at 5100Å
LOGL ₁₃₅₀	erg s^{-1}	Luminosity at 1350Å

Table 3.3: The number of spectra removed from our sample by the cuts described in Section 3.5.5.

		H α sample	H β sample
	Total	194	279
H α /H β	Wavelength	6	27
	S/N	8	83
C iv	Wavelength	6	5
	S/N	4	12
	Absorption	6	8
	Total remaining	164	144

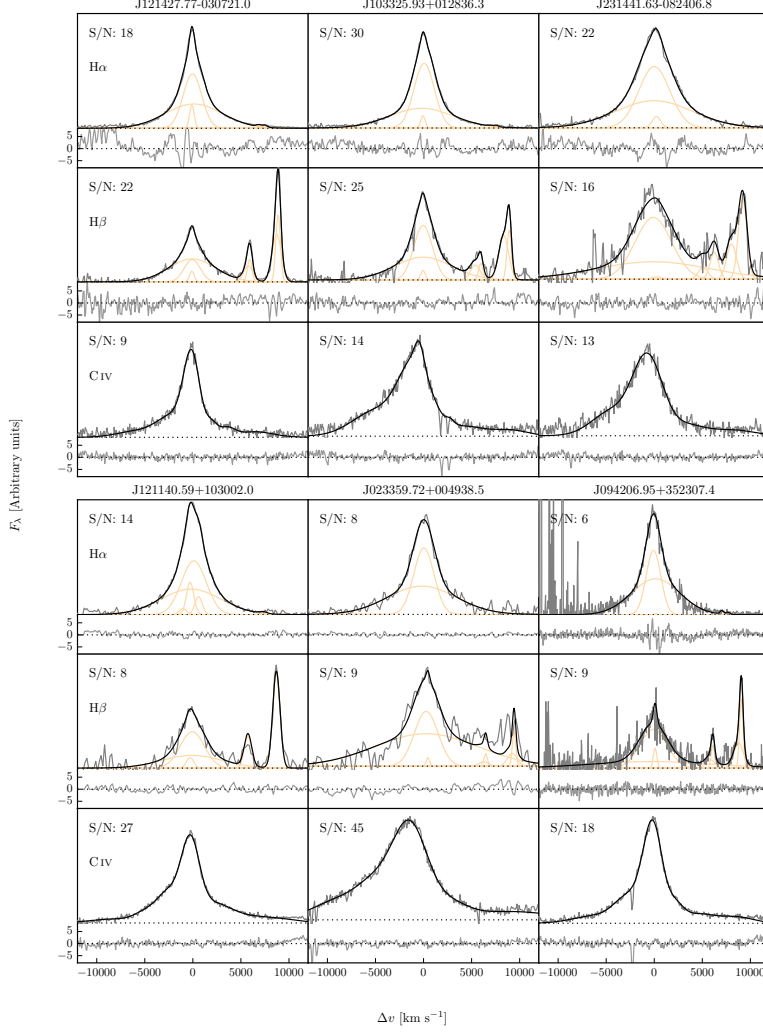


Figure 3.5: Model fits to continuum-subtracted $H\alpha$, $H\beta$, and $C\text{IV}$ emission in six quasars, chosen to represent the range of S/N (indicated in the figure and given per 150km s^{-1} pixel in the continuum) and line shapes present in the catalogue. The data is shown in grey, the best-fitting parametric model in black, and the individual model components in orange. The centroid of the broad $H\alpha$ emission is used to set the redshift, and 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.

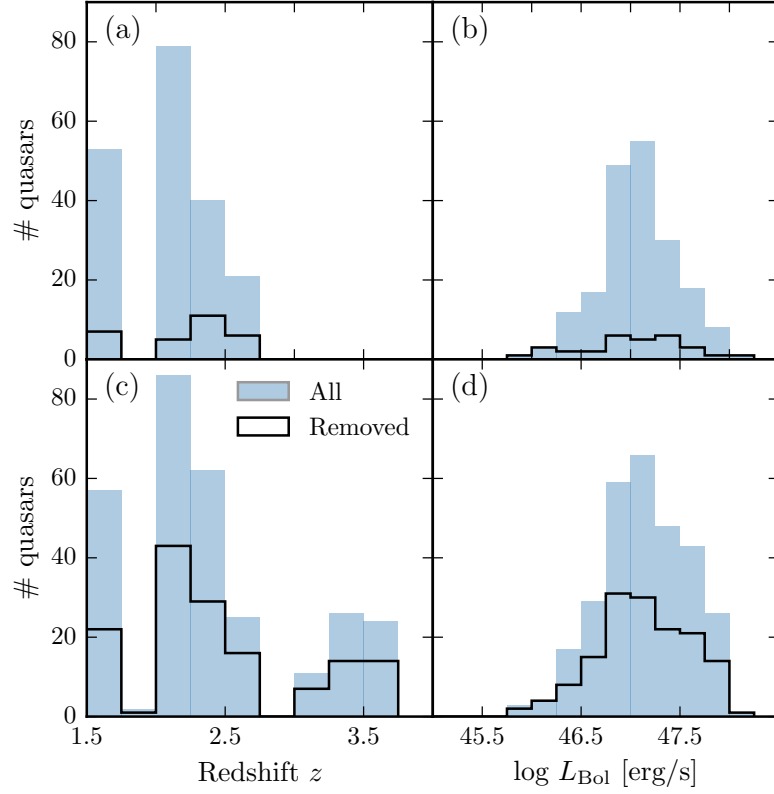


Figure 3.6: The redshift and luminosity distributions of the spectra removed from our H α /C IV (a, b) and H β /C IV (c, d) samples.

generally of lower S/N. As a result, the fraction of H β spectra that are flagged – 39 per cent – is particularly high.

3.5.6 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 fits⁶. We therefore chose to measure the spread in the parameter distribution by 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 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. 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 five thousand artificial spectra. Our parameter uncertainties are defined as the standard deviation of the best-fitting parameter values from these five thousand realisations. The uncertainty on the monochromatic continuum luminosity at 5100 Å was estimated via a very similar method – using the error on the UKIDSS H-band magnitude to run a number of realisations of our SED-fitting routine. The uncertainties on all derived quantities, such as the BH mass, are propagated through by assuming that the uncertainties are uncorrelated and independent.

Because of its sensitivity to the flux in the wings of the line profile, care must be taken to define an appropriate range over which to measure the line dispersion. This is particularly true of Lorentzian-like profiles with extended wings. In spectra of only moderate S/N the line limits are difficult to determine unambiguously, which introduces an extra degree of uncertainty in line dispersion measurements. In common with previous work (e.g. Vestergaard and Peterson, 2006), by default, the dispersion was calculated within $\pm 10\,000\text{ km s}^{-1}$ of the line centre, but this was extended when appropriate to avoid excluding a significant amount of line flux.

3.5.7 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 spectra 20 years or more after the Hamburg/ESO observations⁷. 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.

⁶ In the analysis of the real spectra such fits are identified via visual inspection.

⁷ Time differences in the quasar rest-frame are reduced by a factor of $(1+z)$.

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 $\simeq 500 \text{ km s}^{-1}$. Differences in the S/N of the spectra will make a substantial contribution and the scatter due to true variations in the C IV velocity-width will be significantly smaller than 500 km s^{-1} . We conclude therefore that any intrinsic changes with time do not materially affect the emission line measurements.

3.5.8 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 \AA for C IV and 5100 \AA for H α and H β .

Relative flux-calibration of the infrared spectra as a function of wavelength has been achieved, to $\simeq 10$ per cent, through observations of appropriate flux standards. The absolute flux levels, however, can be in error by large factors due to variable atmospheric conditions combined with the narrow slit widths. For the majority of the quasars we have, therefore, established the absolute flux scale for each near-infrared spectrum using the same quasar SED-model fitting scheme employed in Paper I. The SED model, described in Maddox et al., (2012), gives a very good fit to the SDSS and UKIDSS magnitudes of SDSS DR7 quasars, reproducing the individual magnitudes with a $\sigma < 0.1 \text{ mag}$. For 207 quasars, (Y)JHK passband magnitudes from the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al., 2007) Large Area Survey, the Two Micron All Sky Survey (2MASS; Skrutskie et al., 2006) and the Visible and Infrared Survey Telescope for Astronomy (VISTA) Hemisphere Survey (VHS; McMahon et al., 2013) and Kilo-Degree Infrared Galaxy (VIKING; Edge et al., 2013) survey are available. The SED model was fit to the infrared magnitudes; integrating the SED model through the pass-band transmission functions, to give model magnitudes, and performing a variance weighted least-squares fit to the observed magnitudes. The flux at 5100 \AA was then taken from the normalised model.

For 19 of the remaining 23 quasars, where near-infrared photometry was not available, the quasar SED model was fit to the SDSS spectra, the flux calibration of which are known to be excellent. The fit was done using a simple variance-weighted chi-squared minimisation procedure in emission line-free intervals of the optical spectra.

The model includes a reddening, $E(B - V)$, based on a Small Magellanic Cloud-like extinction curve, and an overall normalisation of the model as free parameters. In practice, the quasars possess only very modest reddenings, with $E(B - V) \simeq 0.0-0.1$. The flux at 5100Å was then, again, taken from the normalised SED model. For the four remaining quasars, which possess neither near-infrared photometry nor SDSS DR7 spectra, we fit the SED model to the BOSS DR12 spectra. To avoid the known issues in the flux calibration of the BOSS DR12 quasar spectra at observed-frame blue wavelengths (Lee et al., 2013), our fitting was confined to rest-frame wavelengths long-ward of 1275Å .

Comparison of the 5100Å luminosity, computed using the photometry- and spectrum-based methods for 177 quasars, showed a scatter of just ~ 0.1 dex. We therefore assume 0.1 dex to be the measurement uncertainty on the 5100Å luminosities.

For 34 quasars in the catalogue the optical spectra come from surveys other than SDSS/BOSS and optical magnitudes from recent epochs are not available. In order to obtain an estimate of the luminosity at 1350Å for the 30 quasars, we take the standard Maddox et al., (2012) quasar SED model, normalised to the near-infrared magnitudes, and read off the flux at 1350Å .

For all the catalogue quasars, the optical and near-infrared spectra as well as the near-infrared photometry were obtained at different epochs, with rest-frame time differences of up to ~ 5 years. Intrinsic quasar photometric variability in the rest-frame ultraviolet and optical will therefore add additional scatter of ~ 0.2 mag (e.g. MacLeod et al., 2010) to the derived 1350 - and 5100Å -luminosities.

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.

The luminosity at 1350Å was taken from the spectral fits of Shen et al., (2011). The quasar rest-frame continuum at 5100Å often lies at the edge, or beyond, the wavelength coverage of the LIRIS spectra. Monochromatic 5100Å luminosities were therefore calculated from the fit of our parametric quasar model (described in Maddox et al., (2012)) to the UKIDSS H- broadband magnitude for each quasar. The model fits to the quasars are excellent, with residuals in SDSS and UKIDSS passbands under 10 per cent.

3.5.9 Characterising the emission-line profiles

Our definition of the C IV blueshift differs slightly in two ways from the values plotted for the quasar population in Fig. 3.18. Firstly, we use the peak of our parametric model fit to the $H\alpha$ line to define the

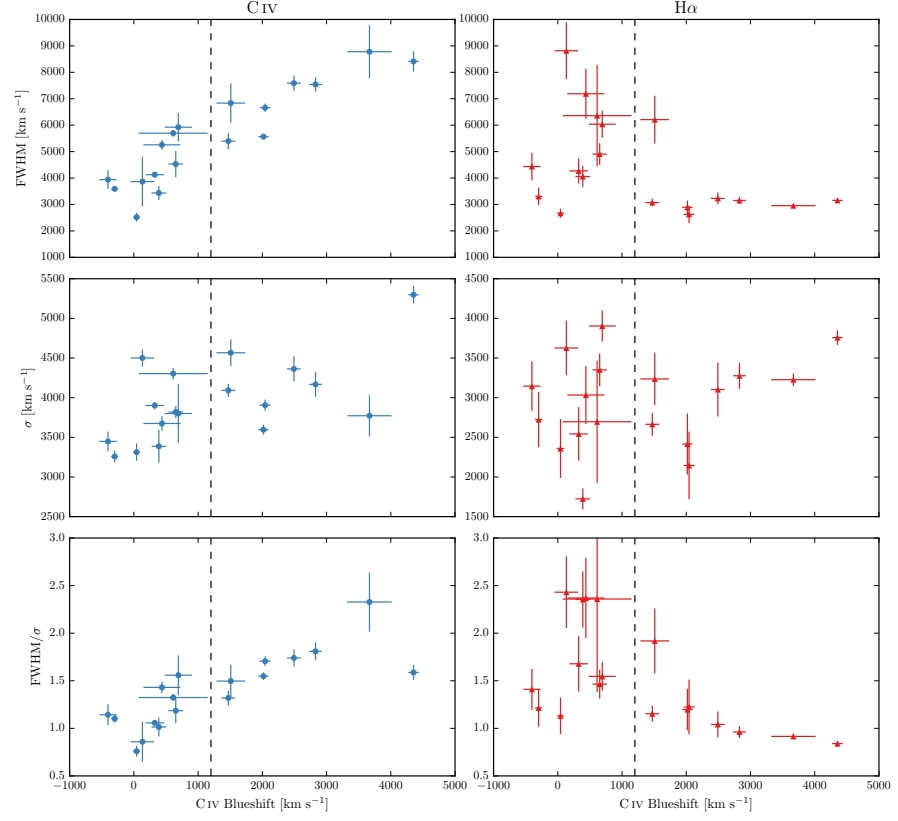


Figure 3.7: The FWHM, dispersion (σ) and shape (FWHM/ σ) of C IV and H α as a function of the C IV blueshift. The vertical line demarcates the ‘high’ and ‘low’ C IV blueshift regimes discussed in the text. At high blueshift it is clear that BH masses estimated from the C IV FWHM (as is typically done at the redshifts considered) will be significantly larger than those estimated from the H α FWHM.

quasar systemic-redshift⁸. Secondly, the centre of the C iv line is now defined as the wavelength that bisects the cumulative total flux of our best-fit GH-polynomial model rather than of the data.

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 relatively more sensitive to any low-velocity core present and the line dispersion relatively more sensitive to the high velocity wings. The shape of the line can be characterised by the ratio FWHM/σ . $\text{FWHM}/\sigma \simeq 2.35$ for a Gaussian profile, while $\text{FWHM}/\sigma \simeq 1$ for a peakier Lorentzian profile⁹. 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 the left panels of Fig. 3.7 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

⁸ The H α -derived redshifts are very closely in agreement with those from the forthcoming Allen & Hewett redshifts, which are plotted in Fig. 3.18c.

⁹ Strictly $\text{FWHM}/\sigma \rightarrow 0$ for a Lorentzian profile, but values close to unity are typical when the dispersion is calculated over a velocity range, $\simeq \pm 10\,000\text{km s}^{-1}$, used to parametrize broad emission lines in quasar spectra.

the emission-line profile shape than is the case for the dispersion, which varies by a factor of only $\simeq 2$. Adopting a value of 1200 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. 3.8 we overlay the C IV line profiles of SDSSJ1236+1129 and SDSSJ1525+2928, whose dispersions (Table ??) are indistinguishable (4168 ± 271 and $4303 \pm 128 \text{ km s}^{-1}$ respectively). Notwithstanding the very similar dispersion values, the emission-line velocity fields differ dramatically and, therefore, the dispersion values cannot be measuring accurately the virial-induced velocity spread of the C IV emission in both quasars.

The analysis here, building on earlier work (including 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). As a first step towards the goal, below (Sec. 3.7.5) we investigate the apparent systematic trends in the H α FWHM and line shape as a function of C IV blueshift (shown in the right of Fig. 3.7).

3.6 AN EMPIRICAL CORRECTION TO C IV-BASED VIRIAL BH-MASS ESTIMATES

3.6.1 H α /H β FWHM comparison

BH-mass calibrations which use the width of the broad H β emission line as a proxy for the virial velocity are widely regarded as the most reliable, since most reverberation mapping employs the H β line and the R – L relation has been established using H β . When H β is not

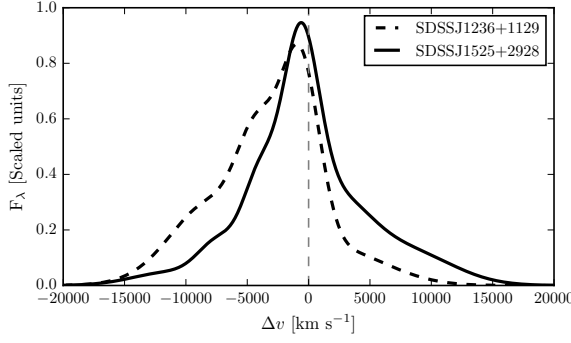


Figure 3.8: Comparison of the CIV line profiles of SDSSJ1236+1129 and SDSSJ1525+0426. Notwithstanding the essentially identical dispersion values, the emission-line velocity fields differ dramatically and, therefore, the dispersion values cannot be measuring accurately the virial-induced velocity spread of the CIV emission in both quasars.

available, H α has been shown to be a reliable substitute (e.g. Greene and Ho, 2005; Shen et al., 2011; Shen and Liu, 2012).

In our sample, we have 99 quasars with reliable measurements of both H α and H β lines. The 99 objects include 21 quasars which were excluded from the main 308-object catalogue because adequate measurements of the CIV FWHM and blueshift could not be acquired. The line widths are compared in Fig. 3.9 and, as expected, a tight correlation is observed. Greene and Ho, (2005), using a sample of 162 quasars with high S/N SDSS spectra at $z < 0.35$, established the following relation between the H α and H β FWHMs

$$\text{FWHM}(\text{H}_\alpha) = (1.07 \pm 0.07) \times 10^3 \left(\frac{\text{FWHM}(\text{H}_\beta)}{10^3 \text{ km s}^{-1}} \right)^{(1.03 \pm 0.03)} \quad (3.1)$$

The relation is shown as the dashed line in Fig. 3.9. 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 H β line-widths we measure are on average larger by 270 km s^{-1} than predicted by the Greene and Ho, (2005) relation. As our sample covers higher redshifts and luminosities than the sample in Greene and Ho, (2005), we derive a new relation between the H α and H β 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 parameters 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

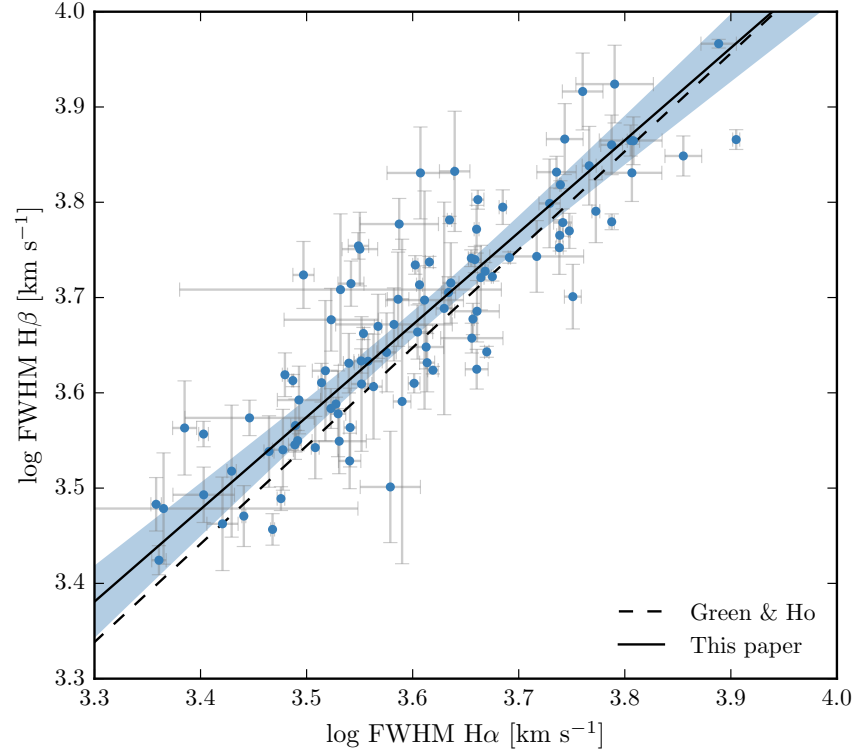


Figure 3.9: Comparison of H α and H β FWHM measurements for 99 quasars. The solid line is our best-fitting power-law model, and the blue-shaded region shows the $2\text{-}\sigma$ uncertainties on the model parameters. The dashed line is the relation found by Greene and Ho, (2005) using a sample of $z < 0.35$ SDSS AGN.

variance V representing the intrinsic variance in the data. The log-likelihood is then given by

$$\ln \mathcal{L} = - \sum_{i=1}^N \frac{1}{2} \ln [2\pi (\Sigma_i^2 + V)] - \sum_{i=1}^N \frac{\Delta_i^2}{2[\Sigma_i^2 + V]} \quad (3.2)$$

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 suggestion by Hogg, Bovy, and Lang, (2010), the linear model was parametrized in terms of (θ, 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. 3.10. The solid line in Fig. 3.9 is the maximum likelihood solution

$$\text{FWHM}(H_\gamma) = (1.23 \pm 0.10) \times 10^3 \left(\frac{\text{FWHM}(H_\alpha)}{10^3 \text{km s}^{-1}} \right)^{0.97 \pm 0.05} \quad (3.3)$$

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

As discussed above, our relation is displaced to slightly higher $H\beta$ FWHM than the Greene and Ho, (2005) relation – the offset is 210km s^{-1} for a quasar with $H\alpha$ FWHM 4500km 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 dex. This is smaller than the total scatter seen in Fig. 3.9 (0.06 dex), which suggests that measurement errors make a significant contribution to the total scatter in the relation.

For 19 of the 99 quasars with $H\beta$ and $H\alpha$ emission profiles, one of the two Gaussians used to reproduce the $H\beta$ profiles has a FWHM greater than 20000km s^{-1} and a fractional contribution to the total $H\beta$ broad line flux of >0.3 (Marziani et al., 2009; Marziani et al., 2013). Such a broad component is not seen in the $H\alpha$ profiles and the very broad $H\beta$ -component may be an artifact of the fitting scheme. A particular issue for $H\beta$ 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 $\lambda 4922, 5017$ doublet; Véron, Gonçalves, and Véron-Cetty, 2002; Zamfir et al., 2010).

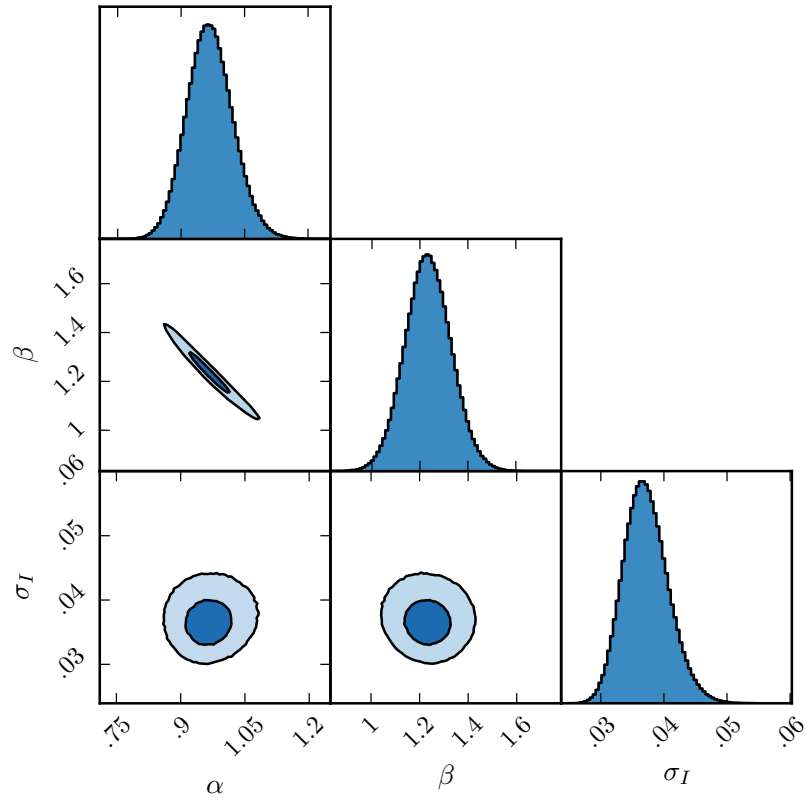


Figure 3.10: One- and two-dimensional projections of the MCMC sampling of the posterior distribution from the fit in Fig. 3.9. α is the power-law index, 10^β is the normalisation, and σ_I is the intrinsic scatter. In the two-dimensional projections, 1- and 2- σ contours are shown.

In Sec. 3.6.3 we use the whole of the H β profile to derive an unbiased BH mass. If, instead, the FWHM is calculated from the narrower of the two Gaussian components rather than the composite profile, then the H β FWHM decreases by 630km s^{-1} on average. The H α FWHM, which are calibrated against the H β FWHM, will also decrease by the same amount on average. This will enhance the C IV FWHM relative to the H α /H β FWHM by ~ 15 per cent and increase the size of the correction which must be applied to the C IV-based BH masses by ~ 30 per cent.

Virial BH mass estimators are typically based on the width of H β . However, the H α and H β emission is believed to originate from the same gas and the transformation between the emission-line velocity widths is expected to be well defined. Greene and Ho, (2005), using a sample of 162 quasars with high S/N SDSS spectra, established the following relation between the H α and H β FWHM:

$$\text{FWHM}(\text{H}\beta) = (1.07 \pm 0.07) \times 10^3 \left(\frac{\text{FWHM}(\text{H}\alpha)}{10^3 \text{ km s}^{-1}} \right)^{(1.03 \pm 0.03)} \quad (3.4)$$

Greene and Ho, (2005) found the root-mean-square scatter about this relation to be ~ 0.1 dex. We do not have a sufficient number of robust H β line measurements to test this relation directly. However, we are in the process of acquiring a much larger sample of quasars with near-infrared spectra covering H α and H β at similar redshifts and luminosities to the sample presented here. The H α and H β line widths of this sample are in excellent agreement with the Greene and Ho, (2005) relation. To indirectly test the H α /H β line width relation for the sample presented here, we first constructed mean composite spectra in the H α and H β emission line regions to increase the S/N. The individual rest-frame spectra (defined using the wavelength of the H α centroid) were interpolated on to a common wavelength grid. The spectra were then normalised using the continuum flux under the line centre, which was found by linearly interpolating between two emission-line free windows on either side of the line. Figure. 3.11 shows the composite H α and H β line regions overlaid and, as expected, the line profiles are closely matched.

3.6.2 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, where the centroid, λ_{half} , is the wavelength that bisects the cumulative total flux, are available for all quasars in the catalogue and so we use this to define the systemic redshift.

For 83 and 120 quasars in the H α and H β samples respectively narrow [O III] emission is also detected. In the model fit to the H β region

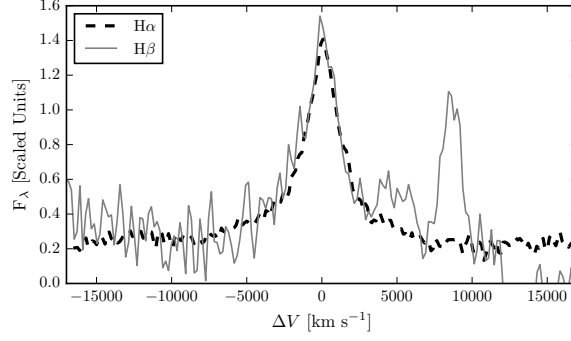


Figure 3.11: The H α and H β emission line regions in the median composite spectrum, shown as function of the velocity shift from the respective predicted line peak wavelengths. The line fluxes have been scaled in order for the profile shapes to be readily compared. The H α and H β line profiles are very similar, which suggests a tight correlation between the H α and H β line widths. Quasar narrow-line emission from O III λ 5008.2 is visible but, overall, the O III λ 4960,5008 emission is relatively weak in these spectra.

the velocity centroids of the broad H β -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 Balmer broad emission and the narrow [O III] emission to have a dispersion of 360 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 270 km s^{-1} relative to the centroid of the broad Balmer line. Applying our parametric model fitting routine to the composite spectrum from Hewett and Wild, (2010), which is constructed using relatively low redshift SDSS quasars with $L_{\text{Bol}} \sim 10^{44} \text{ erg s}^{-1}$, the centroids of the broad component of H β 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. 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.

3.6.3 Balmer/C IV line widths as a function of C IV-blueshift

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. However, the C IV line profiles of the quasars in our sample with the largest C IV blueshifts

indicate that non-virial motions, very likely due to outflows, are having a significant effect on the observed C iv emission velocity profile (e.g. Gaskell, 1982; Baskin and Laor, 2005a; Sulentic et al., 2007; Richards et al., 2011; Wang et al., 2013). As shown in Fig. 3.2, 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.

The main aim of this paper is to investigate potential systematic trends in C iv-based single-epoch virial BH masses as a function of the C iv blueshift. Calibrations using H β (and therefore also H α) are generally accepted to be the most reliable, since most reverberation mapping employs the H β line and the R – L relation has been established using H β . Therefore, we will test the reliability of the C iv-based estimates by comparing C iv line profiles to H α profiles in the same quasars.

In this section we directly compare the C iv and H α /H β line widths as a function of the C iv blueshift. Because virial BH mass estimates are generally based on the H β FWHM, we first convert our H α FWHM measurements to equivalent H β FWHM using Eq. 3.3. In Fig. ??a and b we show the C iv FWHM relative to both the (H β -scaled) H α FWHM and the H β FWHM, as a function of the C iv blueshift.

Employing the same Bayesian fitting framework described in Section 3.6.1, we fit independent linear models to the C iv FWHM relative to the H α and H β 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. 3.14 we show the one- and two-dimensional projections of the posterior distribution from the linear fit to the FWHM C iv/H α ratio. The projections from the FWHM C iv/H β fit (not shown) have very similar appearances. In Fig. ??a we plot the maximum likelihood model and the 2σ uncertainties on the model parameters. The maximum likelihood line is given by

$$\text{FWHM}(\text{C iv, Corr.}) = \frac{\text{FWHM}(\text{C iv, Meas.})}{(0.41 \pm 0.02) \left(\frac{\text{C iv Blueshift}}{10^3 \text{ km s}^{-1}} \right) + (0.62 \pm 0.04)} \quad (3.5)$$

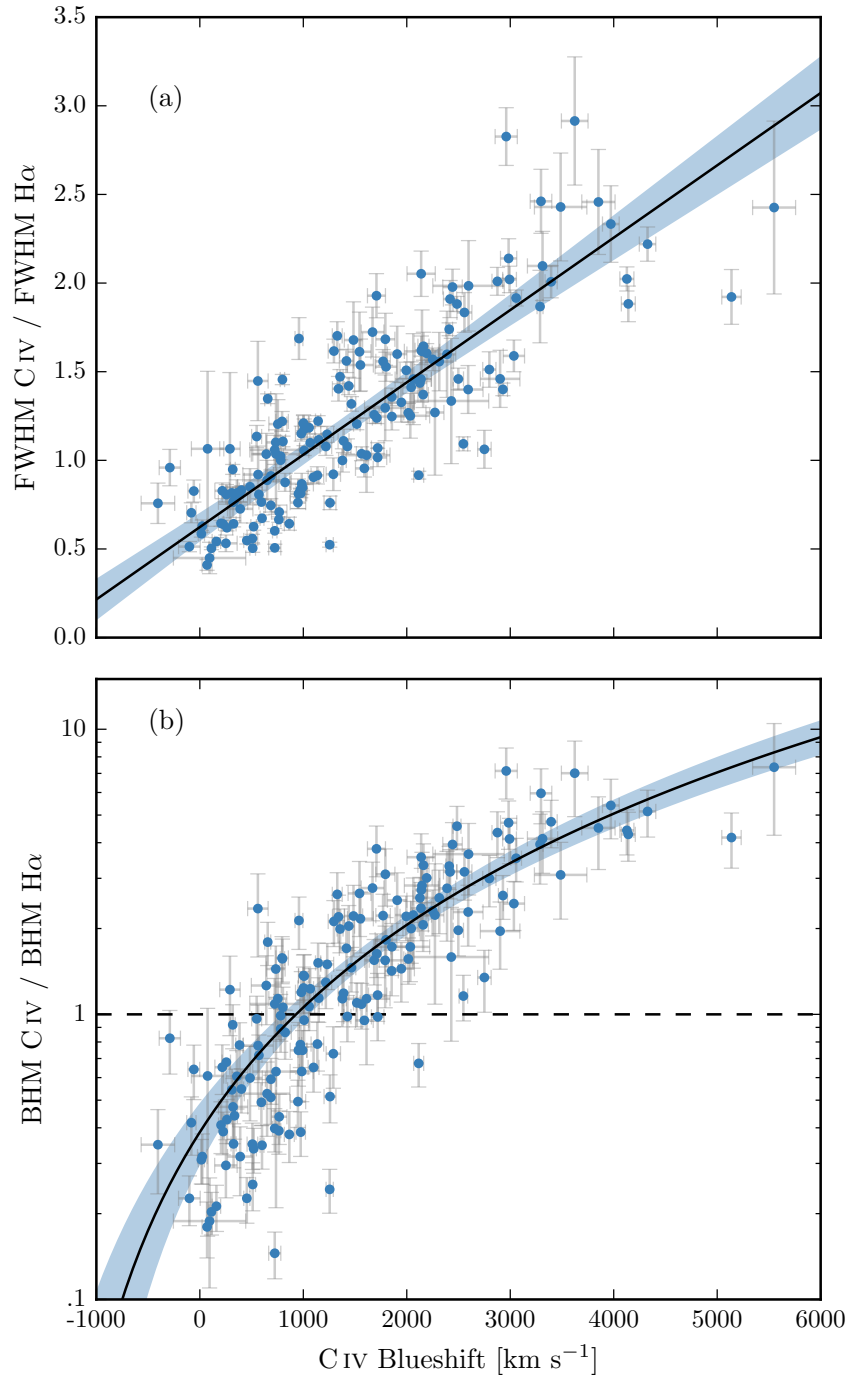


Figure 3.12: C IV FWHM relative to H α FWHM (a), and C IV based BH mass (BHM) compared to H α based mass (b), both as a function of the C IV blueshift. The black line is our best-fit linear model, and the shaded region shows the 2- σ uncertainties on the slope and intercept. The H α FWHM have been scaled to match the H β FWHM using Eq. 3.3.

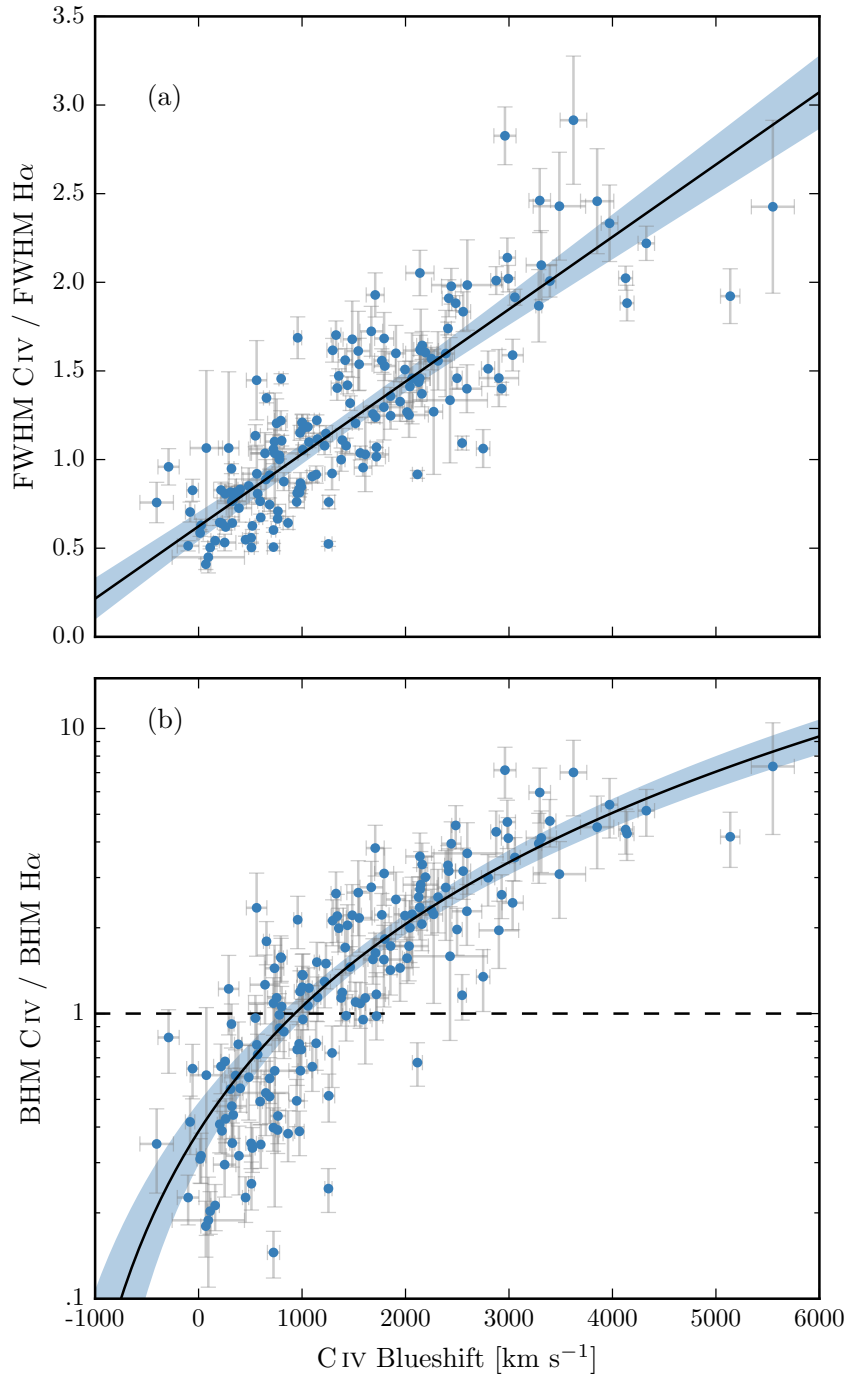


Figure 3.13: C IV FWHM relative to H β FWHM (a), and C IV based BH mass (BHM) compared to H β based mass (b), both as a function of the C IV blueshift.

for the C IV/H α fit and

$$\text{FWHM}(\text{C IV, Corr.}) = \frac{\text{FWHM}(\text{C IV, Meas.})}{(0.36 \pm 0.03) \left(\frac{\text{C IV Blueshift}}{10^3 \text{ km s}^{-1}} \right) + (0.61 \pm 0.04)} \quad (3.6)$$

for the C IV/H β 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 H α and H β fits respectively. The intrinsic scatter for the H α fit is represented by the Normal probability density distribution shown in Fig. 3.15. In the same figure we show the distribution of the orthogonal displacement of each data point from the best-fitting linear relationship. The two distributions are well-matched, which demonstrates that our model is a good representation of the data and the measurement errors on the data points are small relative to the intrinsic scatter.

The overall (intrinsic and measurement) scatter about the best-fitting model is slightly higher when the C IV line-widths are compared to H β (0.12 dex) than when compared to H α (0.10 dex). This is likely due, at least in part, to the generally higher S/N of the H α emission. In addition, contributions from the strong [O III] doublet in the vicinity of H β make de-blending the H β emission more uncertain. As a consequence, for quasars where H α and H β are both measured, the mean uncertainty on the H α FWHM is 130 km s^{-1} , compared to 340 km s^{-1} for H β .

In the next section we use both the H α and H β lines to calculate unbiased BH masses. We use the H α measurements to derive an empirical C IV blueshift based correction to the C IV masses (Eq. 3.8) because of the issues related to the accurate modelling of the H β -profile just described. An extra advantage, which is evident in Fig. ??, is that the H α sample has a better C IV blueshift coverage. However, as can be seen from the similarity of Equations 3.5 and 3.6, our results would not change significantly were we instead to use the H β sample.

3.6.4 C IV-derived BH masses at low C IV blueshift

The H α and C IV FWHM (dispersion) of the 10 quasars with C IV blueshifts $< 1200 \text{ km s}^{-1}$ are linearly correlated, as expected if the dynamics of the BLR clouds are dominated by virial motions. The median C IV/H α FWHM (dispersion) ratio is 0.91 (1.22) with standard deviation 0.17 (0.28). Thus, the dispersion-based comparison results in a median C IV/H α consistent with the value of $\simeq 1.4$ from assuming a virial origin for the emission but with a relatively large standard deviation. As predicted in Section ??, the FWHM-based comparison results in a systematically lower median C IV/H α . However, the corre-

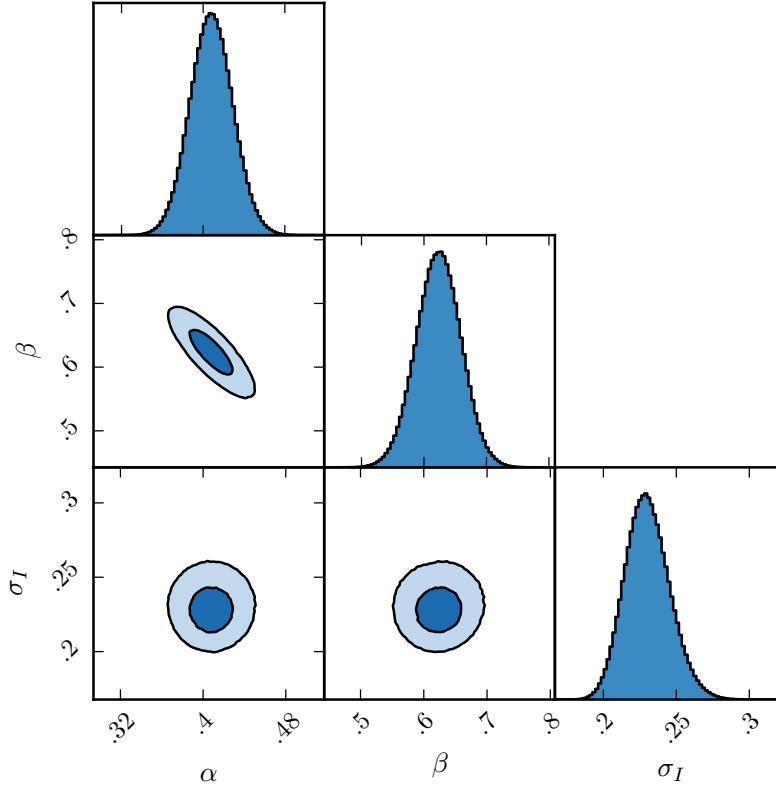


Figure 3.14: One- and two-dimensional projections of the MCMC sample of the posterior distribution for a linear fit to the FWHM C IV/H α ratio as a function of the C IV blueshift. In the two-dimensional projections we show 1- and 2- σ contours. The posterior distribution for the linear fit to the FWHM C IV/H β ratio, which we do not show, has a very similar appearance.

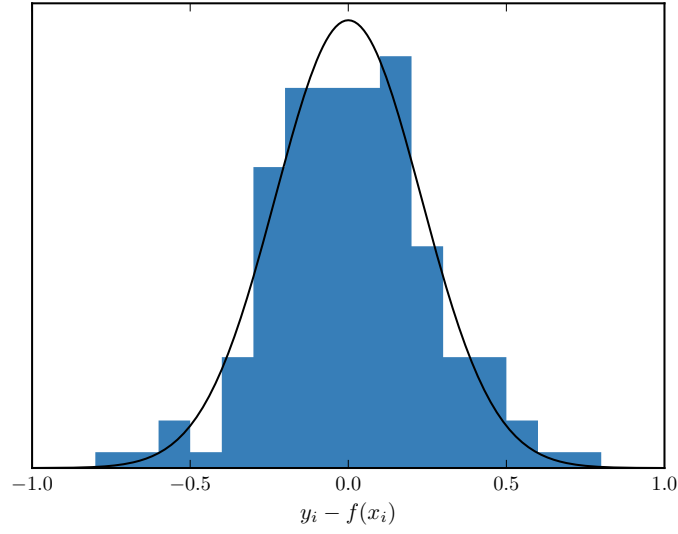


Figure 3.15: The distribution of the orthogonal displacement of each data point from the best-fitting linear relationship in the fit to $\text{FWHM}(\text{C IV})/\text{FWHM}(\text{H}\alpha)$ as a function of the C IV blueshift (blue histogram). The black curve is a Normal distribution with a width equal to the intrinsic scatter in the population inferred from the fit. The two distributions are well-matched, which demonstrates that our model is a good representation of the data and the measurement errors on the data points are small relative to the intrinsic scatter.

lation between the C IV and H α FWHMs is significantly tighter, lending support to the proposal that corrections to BH-mass estimates based on the C IV emission line properties may be possible.

Virial BH masses were calculated using the widely adopted Vestergaard and Peterson, (2006) calibrations. The Vestergaard and Peterson, (2006) C IV FWHM calibration uses the monochromatic continuum luminosity at 1350 Å to predict the BLR radius and corresponds to ($a = 6.66$, $b = 2$, $c = 0.53$) in Eq. 3.7. The calibration coefficient $a = 6.73$ in their equivalent dispersion-based relation. For the H β 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 Tables ?? and ??, and we convert our H α emission-line velocity-width measures to predicted H β widths using Eq. 3.4.

As a direct consequence of the empirically small C IV/H α FWHM ratio, the C IV-derived BH mass estimates are systematically lower than the corresponding H α -derived masses when the blueshift is small. This can be seen in Fig 3.16, where for every quasar with a C IV blueshift $< 1200 \text{ km s}^{-1}$ (i.e. to the left of the dashed line), the C IV-derived BH mass is smaller than the corresponding H α -derived mass. The median fractional difference between the two estimates is 0.60.

For the 10 quasars with low C IV blueshifts, we looked for correlations of the C IV/H α FWHM ratio with other spectral properties. We found marginal evidence for an anti-correlation with the H α FWHM (Spearman coefficient 0.58 with p-value 0.08). Among the quasars with H α FWHM $> 4000 \text{ km s}^{-1}$ we found the mean C IV/H α FWHM ratio to be 0.83, compared to 1.01 for the quasars with H α FWHM $< 4000 \text{ km s}^{-1}$. Similar trends have been observed at low- z ; in a sample of Boroson and Green, (1992) quasars, Baskin and Laor, (2005a) found the C IV line to be broader than H β when the H β FWHM $< 4000 \text{ km s}^{-1}$ and narrower when the H β FWHM $> 4000 \text{ km s}^{-1}$.

3.6.5 C IV based virial BH mass estimates

Single-epoch virial BH mass estimates normally take the form

$$M_{\text{BH}} = 10^a \left(\frac{\Delta V}{1000 \text{ km s}^{-1}} \right)^b \left[\frac{L_\lambda}{10^{44} \text{ erg s}^{-1}} \right]^c \quad (3.7)$$

where ΔV is a measure of the line width (from either the FWHM or dispersion), L_λ is the monochromatic continuum luminosity at wavelength λ , and a , b , and c are coefficients, determined via calibration against a sample of AGN with reverberation-mapping BH mass estimates. Several calibrations have been derived using different lines

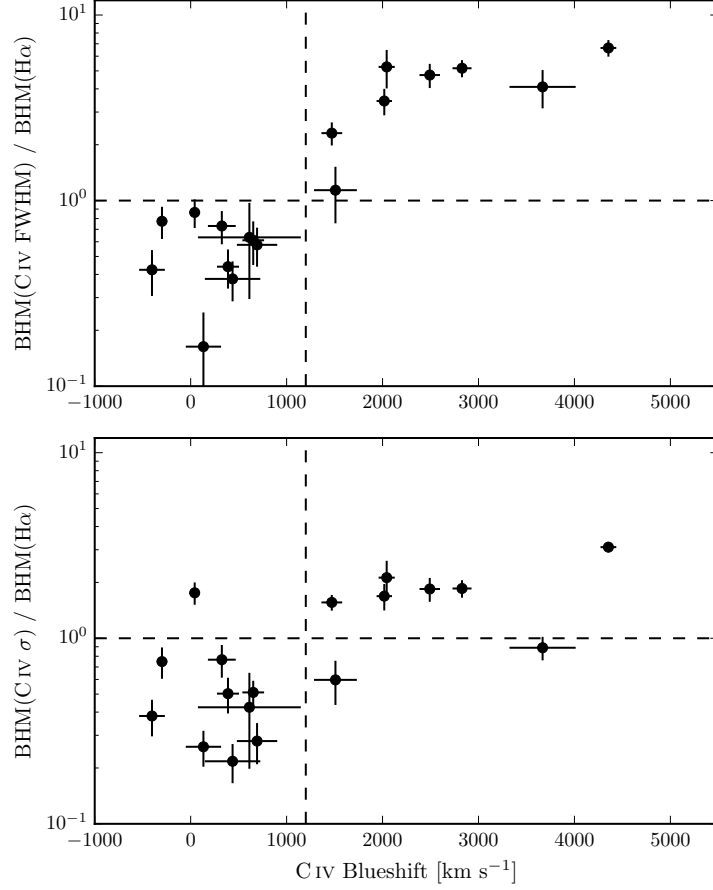


Figure 3.16: Comparison of virial BH mass estimates based on the C IV FWHM (*top*) and dispersion σ (*bottom*) and H α FWHM as a function of the C IV blueshift. The horizontal line indicates agreement between C IV and H α BH masses, and the vertical line demarcates the ‘high’ and ‘low’ C IV blueshift regimes discussed in the text. The BH masses of quasars with moderate C IV blueshifts are underestimated when using the C IV FWHM, while the masses of quasars with large blueshifts are severely overestimated. This situation cannot be corrected by changing the exponent on the FWHM (e.g. Rafiee and Hall, 2011; Park et al., 2013) or the overall scaling in standard virial BH mass relations.

(e.g. H β , Mg II, C IV) and different measures of the line width (FWHM or dispersion) (e.g. Vestergaard, 2002; McLure and Jarvis, 2002; Vestergaard and Peterson, 2006; McGill et al., 2008; Wang et al., 2009; Rafiee and Hall, 2011; Park et al., 2013).

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

We calculate virial BH mass estimates from C IV, H α and H β using the widely-adopted Vestergaard and Peterson, (2006) scaling relations (their equations 5 and 7 for H α /H β and C IV respectively). In Figs. ??c and d the C IV-based estimates are compared to the H α /H β estimates as a function of the C IV blueshift. There is a strong systematic error in the C IV-based masses as a function of blueshift, which is a direct consequence of the FWHM trend described in the previous section. The C IV emission-based BH-masses are in error by a factor of more than five at 3000km s^{-1} in C IV emission blueshift and the overestimate of the BH-masses reaches a factor of 10 for quasars exhibiting the most extreme blueshifts, $\gtrsim 5000\text{km s}^{-1}$.

The virial product is the product of the virial velocity squared and the BLR radius (e.g. Shen, 2013), and is proportional to the BH mass. We use the corrected C IV FWHM given by Eq. 3.5 as an indicator of the virial velocity, and adopt the same $R - L$ relation for the 1350\AA 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 in to a BH mass we compute the inverse-variance weighted mean difference between the virial products and the H α -based masses. The virial BH mass can then be expressed in terms of the corrected C IV FWHM and monochromatic continuum luminosity at 1350\AA

$$\text{MBH}(\text{C IV, Corr.}) = 10^{6.71} \left(\frac{\text{FWHM}(\text{C IV, Corr.})}{10^3 \text{ km s}^{-1}} \right)^2 \left(\frac{-L_{\text{c}}(1350\text{\AA})}{10^{44} \text{ erg s}^{-1}} \right)^{0.53} \quad (3.8)$$

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.

In Section 3.5.9 we have shown that the C iv emission at large C iv-blueshift is not dominated by virial-induced motions due to the BH. The empirically derived increase in the C iv emission FWHM with blueshift leads directly to an overestimate of BH-mass if the trend with blueshift is not taken into account. The availability of the H α -spectra for the sample allows the quantification of the bias in inferred BH-mass under the assumption that the H α emission line provides a reliable BH mass.

Figure 3.16 shows the ratio of the C iv- and H α -FWHM derived BH masses as a function of the C iv blueshift. We see that for quasars with C iv blueshifts $> 2000 \text{ km s}^{-1}$, the C iv-based masses overestimate the H α -based masses by as much as a factor of ~ 5 .

The existence of a trend in the C iv-dispersion values with C iv blueshift is evident from inspection of the bottom left panel of Fig. 3.7 but the systematic trend relative to the spread at fixed blueshift is significantly smaller than when using C iv FWHM. Similarly, Fig. 3.16 shows, at most, only a weak increase in the ratio of C iv- and H α -derived masses. Without knowledge of the C iv blueshifts the distribution of C iv- and H α -dispersion based BH masses could be taken to be reassuring. Including the C iv-blueshift information, however, demonstrates that any such interpretation is inherently flawed because the origin of the C iv emission velocity width is not due to virial-motions for a significant range of C iv blueshift. To reiterate the point made above (Sec. 3.5.9), we believe that using a greater knowledge of the line profile (i.e. both the FWHM and blueshift) is a better motivated (and more practical) approach to obtaining more reliable virial BH mass estimates from the C iv line.

The number of objects in our sample is small but an important factor contributing to the significant correlation evident in the FWHM version of Fig. 3.16 is a change in the emission-line shape of H α as the C iv-blueshift increases. By comparing the distributions of the H α FWHM and dispersion as a function of C iv-blueshift (shown in the right-panels of Fig. 3.7), there is trend for the H α lines to become peakier (with FWHM/ σ approaching unity) as the C iv blueshift increases. Whether the size of the true systematic bias in BH masses inferred from C iv-emission FWHM is as large as shown in Fig. 3.16 will depend on the future parametrization of the reverberation-component present in H β (and H α) profiles for quasars with high luminosities and large C iv blueshifts.

In summary, Fig. 3.16 illustrates the extent to which key derived physical parameters, including the BH mass and L/L_{Edd} , could be systematically in error when C iv-FWHM measures are used without incorporating the information from the C iv blueshifts. Other authors have proposed empirical corrections to C iv-based BH masses based on similar systematic trends seen in the C iv line shape (Denney, 2012) and 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 (Runnoe et al., 2013a). In Section 3.8.1 we apply these corrections the quasars in our sample, and discuss the effect they have on the systematic bias seen in Fig. 3.16.

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

3.7.1 *Recipe for unbiased C IV based BH masses*

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 3.5.1. 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.

Shen et al., (2016) and our own work shows that there is an intrinsic variation of $\sigma \simeq 220 \text{ km s}^{-1}$ in the velocity centroids of the broad-line region relative to a systemic-frame defined by the quasar narrow-line regions. As we showed in Paper I, 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 (see figure 1 in Paper I). The redshift-determination scheme of Hewett and Wild, (2010) provided much improved redshifts for the SDSS DR7 quasar catalogue, not least because the redshift estimates for the majority of quasars were derived using emission-lines other than the C IV-line itself. The redshifts for quasars in the SDSS DR10 and DR12 catalogues (Pâris et al., 2014; Pâris et al., 2017) possess errors of $\simeq 500\text{--}750 \text{ km s}^{-1}$ (Pâris et al., 2012; Font-Ribera et al., 2013). The impact of low spectrum S/N for fainter quasars in all the SDSS data releases increases the uncertainty further. Table 3.4 includes the values for the fractional error in the corrected

Table 3.4: 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

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 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 (≈ 300 km s⁻¹) SED-dependent redshift errors. Redshift uncertainties as large as ≈ 1000 km s⁻¹ for such quasars are unusual and the large relative error in the corrected C iv BH-mass given in Table 3.4 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 Fig. ??). 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. Given the importance of generating more accurate redshifts for the SDSS DR7 and DR12 quasar samples we will publish a catalogue of more accurate redshifts in due course (see Section 3.9).

3.7.2 Systematic trends in residuals

The scatter about the best-fitting line in the C iv/H α 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 H α FWHM (Fig. 3.17). Quasars with broad H α lines tend to lie below the relation while quasars with narrow H α tend to lie above it. One possibility is that this correlation is simply due to random scatter (either intrinsic or measurement error) in the H α FWHM which, with the other quasar properties fixed, would naturally produce a correlation between FWHM(C IV)/FWHM(H α) and FWHM(H α). However, the fact that we see no such correlation between the model residuals and the C IV FWHM suggests that the H α FWHM correlation could be revealing something more fundamental. The H α /H β 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. 3.17 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 residuals and the H α FWHM also correlate with the shape of the line (FWHM/ σ , where σ is the dispersion, derived from the second moment velocity; e.g. Kollatschny and Zetzl, 2011; Kollatschny and Zetzl, 2013). 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 H α 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). We matched our sample to the FIRST radio catalogue (White et al., 1997) in an attempt to identify orientation-dependent signatures. Following Shen et al., (2011), we classified quasars with matches within 5 arcseconds as core-dominated, while, if multiple matches were found within 30 arcseconds, quasars were classified as lobe-dominated. Twenty core- quasars and six lobe-dominated quasars resulted but no statistically significant differences in the H α 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).

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

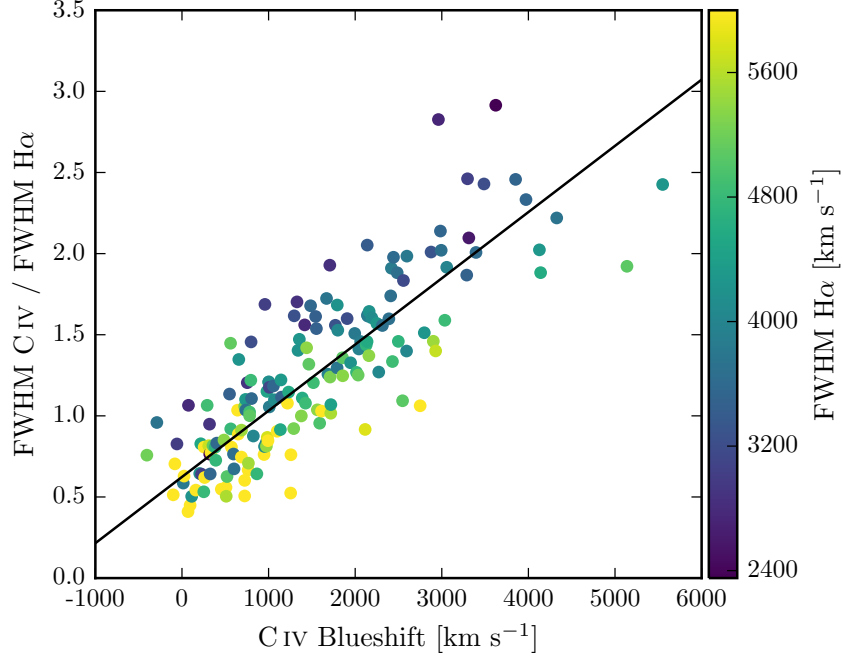


Figure 3.17: Same as Fig. ??a, with the marker colour representing the $H\alpha$ FWHM. At fixed $C\text{ IV}$ blueshift, there is a clear $H\alpha$ FWHM dependent systematic in the model residuals.

(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 $C\text{ IV}$ blueshifts and lead to a greater understanding of the relation between the Balmer line profile and the BH mass.

3.7.3 Effectiveness of the $C\text{ IV}$ blueshift based correction to BH masses

Figure 3.18 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 $C\text{ IV}$ blueshifts for our catalogue relative to the SDSS quasars as a whole is a result of the higher mean luminosity relative to the SDSS sample (Fig. 4.1). Our sample includes 21 quasars with $C\text{ IV}$ blueshifts $> 3000 \text{ km s}^{-1}$, and extends to $\sim 5000 \text{ km s}^{-1}$, i.e. at the very extreme of what is observed in this redshift and luminosity range. Our investigation thus demonstrates that the $C\text{ IV}$ -blueshift based correction derived in this paper is applicable to very high blueshifts. Conversely, there are no quasars in our catalogue with $C\text{ IV}$ blueshifts $\lesssim 0 \text{ km s}^{-1}$ and we caution against extrapolating the correction formula to negative blueshifts.

Figure 3.19 compares the $C\text{ IV}$ - and $H\alpha$ -based BH masses before and after applying the blueshift-based correction to the $C\text{ IV}$ FWHM. Before the correction, the correlation between the $C\text{ IV}$ - and $H\alpha$ -based

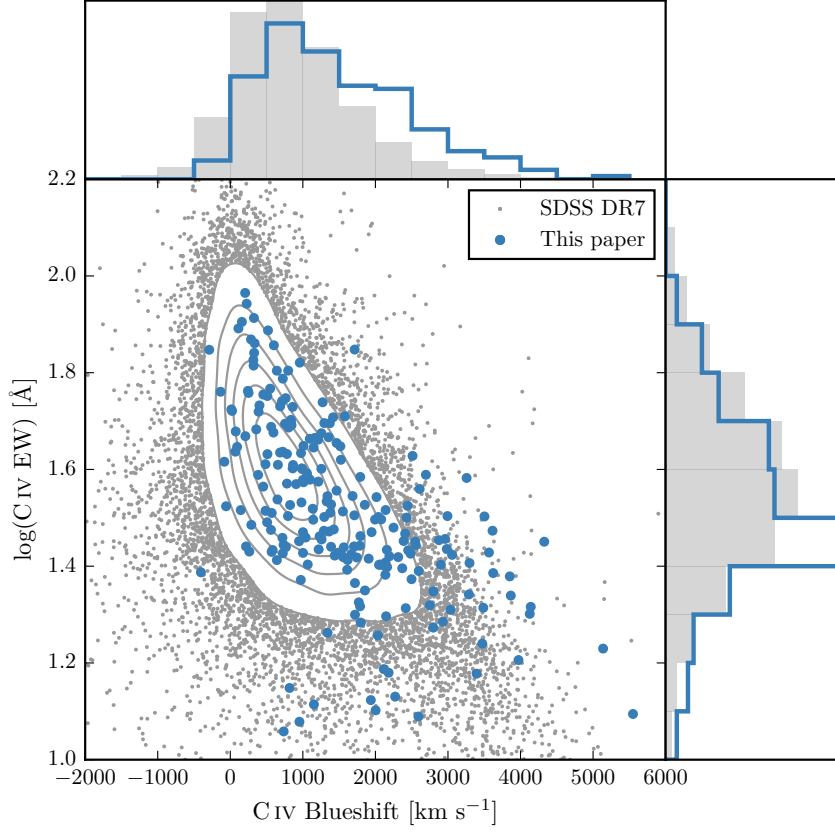


Figure 3.18: Rest-frame EW versus blueshift of the broad C IV-emission line for 32,157 SDSS DR7 quasars at $1.6 < z < 3.0$ (grey) and our sample (blue). For the SDSS quasars, the systemic redshifts used to calculate the blueshifts are from Hewett and Wild, (2010) and C IV emission properties are described in Paper I. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator. 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.

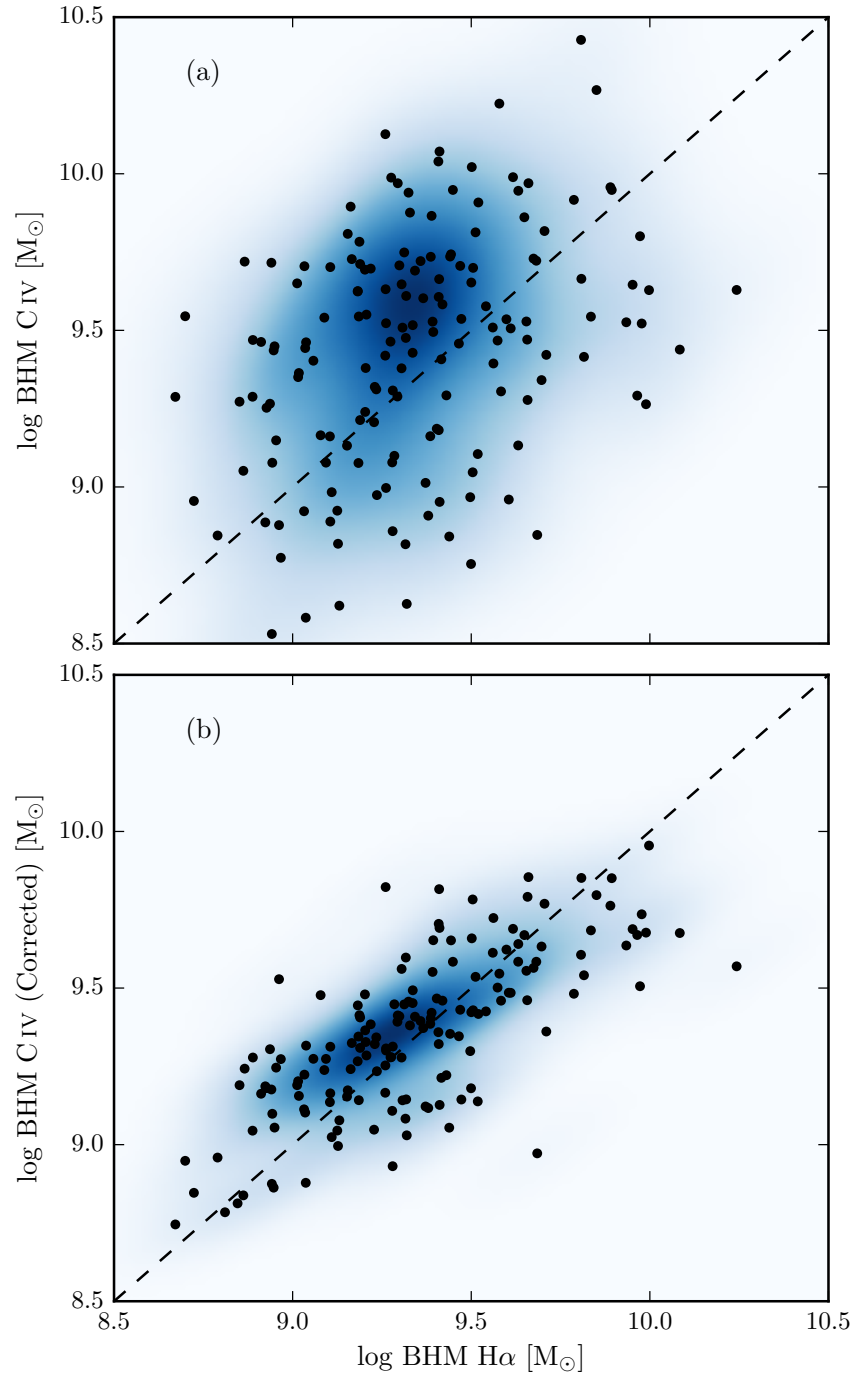


Figure 3.19: Comparison of the C IV- and H α -based BH masses before (a) and after (b) applying the C IV blueshift-based correction to the C IV FWHM. The density of the plotted points (estimated using a Gaussian kernel density estimator) is represented by the colour. The correction to the C IV BH masses decreases the scatter by from 0.4 to 0.2 dex.

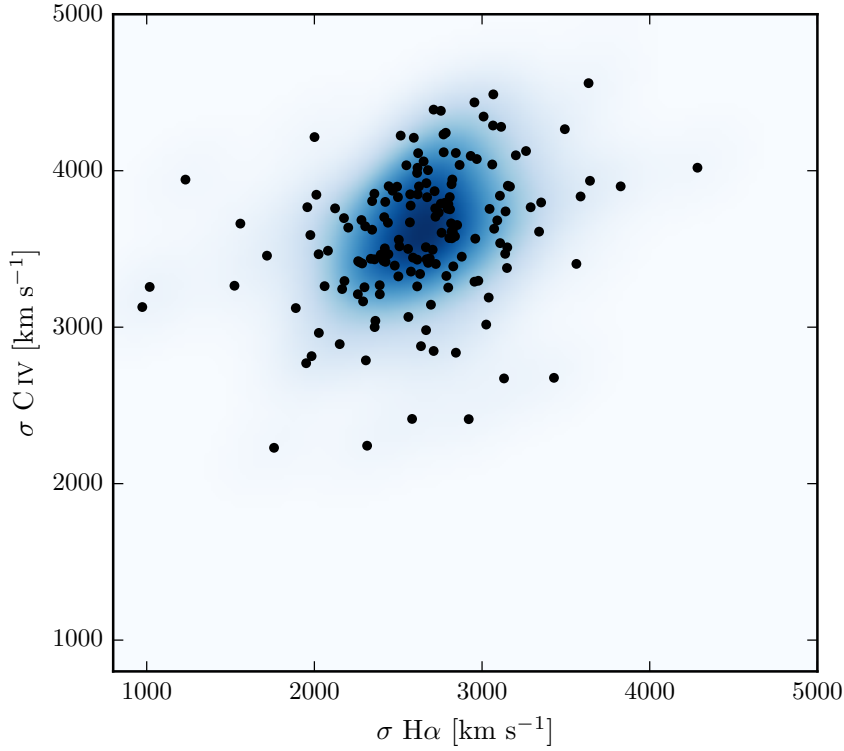


Figure 3.20: Comparison of the C IV and H α line dispersion, σ . The density of the plotted points (estimated using a Gaussian kernel density estimator) is represented by the colour. Estimating a reliable BH mass from the C IV FWHM and blueshift line is substantially more effective than using the C IV line dispersion with, or without, the line blueshift. The C IV dispersion values are larger than the corresponding H α measurements by a factor of 1.4 on average, which is consistent with reverberation mapping measurements (Vestergaard and Peterson, 2006).

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

There has been a considerable amount of attention regarding the relative merits of using the FWHM or dispersion to characterise the velocity width (e.g. Denney et al., 2013). As we showed in Paper I, the line dispersion is relatively insensitive to the blueshift and shape of the C IV line. Therefore, without the blueshift information, using the line dispersion would yield a more accurate BH mass than the FWHM (Fig. 3.20). The correlation between the H α and C IV line dispersion is, however, weak. The Pearson coefficient for the correlation is 0.36 (and just 0.15 when the H β measurements are used in place

of $H\alpha$). Furthermore, there is little dynamic range in the line dispersion: the scatter is just 480 and 460 km s^{-1} for $H\alpha$ and C IV respectively. The observation suggests that the line dispersion does not fully trace the dynamic range in BH mass present in the quasar population. At least part of the reason is that the line dispersion is difficult to measure reliably in current survey-quality data, particularly because of the sensitivity to flux ascribed to the wings of the emission line (e.g. Mejía-Restrepo et al., 2016). Figures 3.19 and 3.20 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.

3.7.4 Comparison to previous prescriptions

In Fig. 3.21 we compare the C IV blueshift-based correction presented in this paper 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 $H\alpha$ -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 $\sim 1400 \text{ \AA}$ (see below). Therefore, our analysis is based on the 123 quasars which satisfy this requirement.

In Fig 3.21a the C IV BH masses have been corrected using the C IV shape (FWHM/ σ) based correction proposed by Denney, (2012). The correction is not applicable at large C IV blueshifts, since it was calibrated on a sample of low-luminosity AGN which does not include any such objects. Therefore, while the consistency between the $H\alpha$ - and C IV-based masses at low C IV blueshifts is improved, at high C IV blueshifts the C IV-based masses remain seriously overestimated.

Runnoe et al., (2013a) used the continuum-subtracted peak flux ratio of the ultraviolet emission-line blend of Si IV+O IV (at 1400 \AA) relative to C IV to correct for non-virial contributions to the C IV velocity-width. 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. 3.18, 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. 3.21b). 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.,

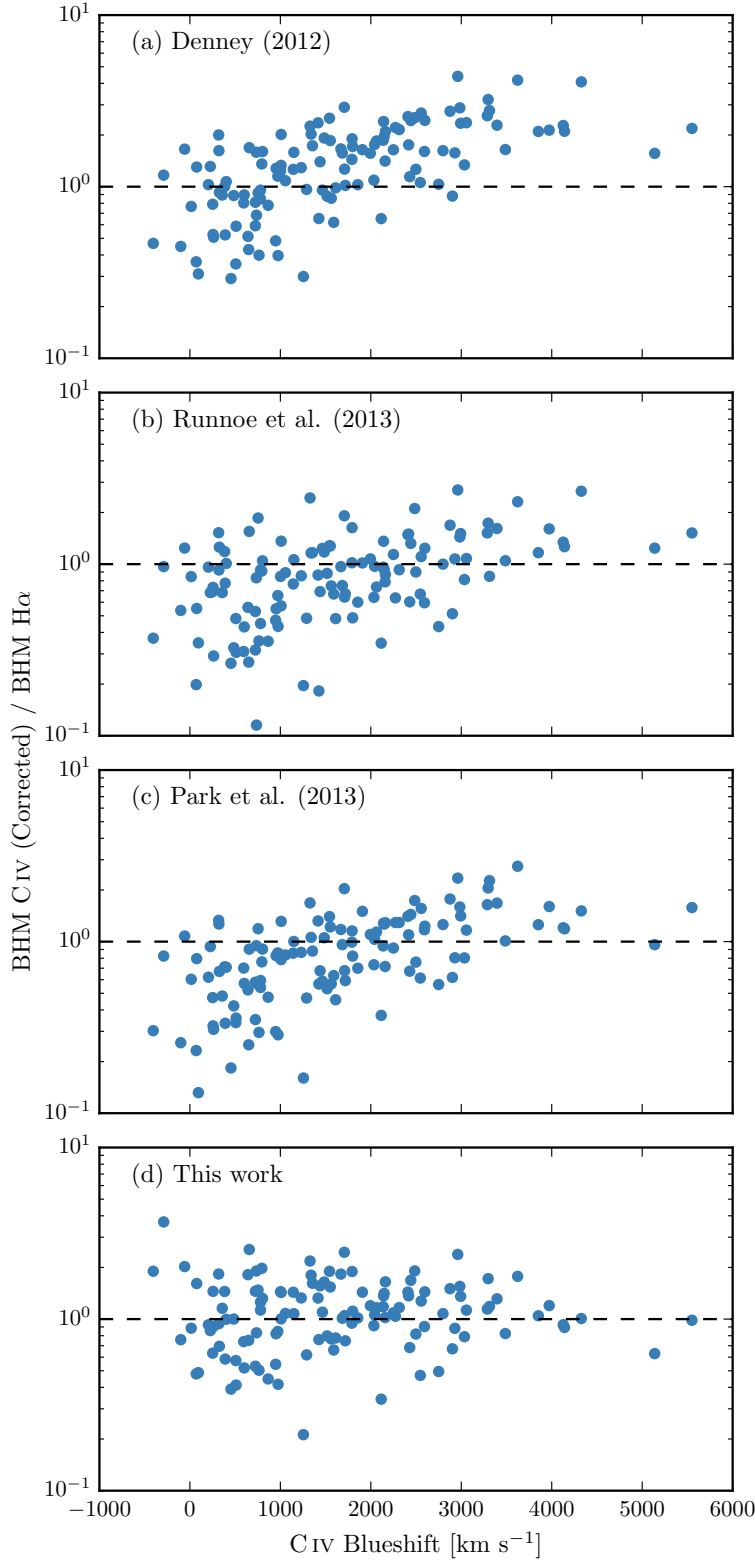


Figure 3.21: Comparison of BH mass estimates derived from C IV and H α as a function of the C IV blueshift. Corrections to the C IV-based masses have been applied based on the shape (FWHM/ σ) of the C IV emission line (a; Denney, 2012), the peak flux ratio of the Si IV+O IV blend relative to C IV (b; 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 paper).

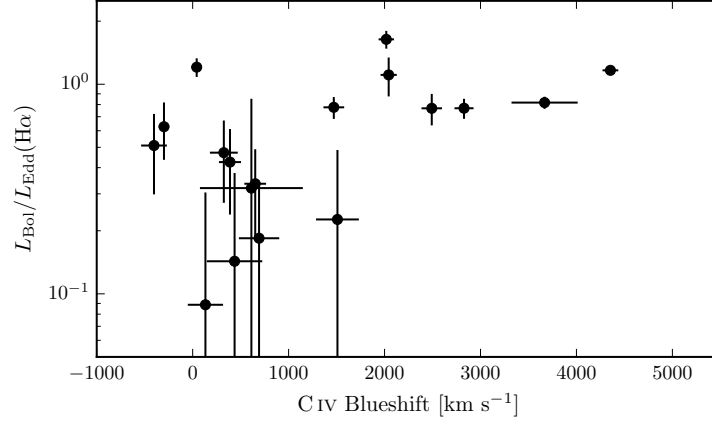


Figure 3.22: $H\alpha$ -derived Eddington ratio versus C IV blueshift. At blueshift $\gtrsim 2000 \text{ km s}^{-1}$ all quasars have high accretion rates ($L/L_{\text{Edd}} \simeq 1$). This is in agreement with Kratzer and Richards, (2015), but in contrast to what one would derive from naive use of C IV-based BH mass scaling relations.

(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. 3.21c). However, the systematic error in the C IV-based BH masses as a function of C IV blueshift remains.

An alternate virial BH-mass calibration is proposed by Park et al., (2013), using an improved sample of AGN with reverberation mapped masses. A major difference from the calibration of Vestergaard and Peterson, (2006) is that Park et al., (2013), recognising the poor correlation sometimes observed between the C IV and $H\beta$ FWHM, allow the exponent on the velocity width (b in Eq. 3.7) to vary. Calibrating Eq. 3.7 against reverberation BH masses, they find a best-fit value of $b = 0.56$, which is much less than the $b = 2.0$ in the strict virial regime. As a result, the derived BH masses are much less sensitive to the C IV-emission line properties.

As a comparison, the C IV-based masses shown in Fig 3.21d have been corrected using to the C IV blueshift-based procedure presented in this paper. No systematic in the BH masses as a function of the C IV blueshift is evident.

3.7.5 Population trends with C IV blueshift

Even with the caveat regarding the small sample size, the differences in the $H\alpha$ emission-profile as a function of C IV-blueshift (Fig. ??) appear to be systematic. At C IV-blueshift $< 1200 \text{ km s}^{-1}$, the $H\alpha$ FWHM range is $\simeq 2700 - 8800 \text{ km s}^{-1}$, with mean $\simeq 5200 \text{ km s}^{-1}$. However,

amongst the six quasars with C iv-blueshift $>2000 \text{ km s}^{-1}$, the mean $\text{H}\alpha$ FWHM $= 3000 \text{ km s}^{-1}$, with a scatter of just 200 km s^{-1} . The apparent trend of peakier $\text{H}\alpha$ -emission, with FWHM/ σ close to unity, at large C iv-blueshift is enhanced by the modest increase in $\text{H}\alpha$ EW with blueshift (Table ??). Amongst the low-C iv-blueshift population there are in addition quasars with broader and more Gaussian-like $\text{H}\alpha$ line profiles, with FWHM/ $\sigma \simeq 2$.

The change in the $\text{H}\alpha$ emission-line profiles as a function of C iv-blueshift means that the $\text{H}\alpha$ -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, in to a bolometric luminosity we use the bolometric correction factor given by Richards et al., (2006a) ($L_{\text{bol}} = 9.26 L_{5100}$). 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. 3.22, show that at large blueshifts quasars are accreting at around their Eddington limits (Fig. 3.22). The converse is, however, not true, i.e. not all quasars with high Eddington ratios possess large C iv blueshifts (see Baskin and Laor, 2005a).

3.8 DISCUSSION

3.8.1 Biases in single-epoch C iv-based BH-mass estimates

The C iv line profiles of the quasars with the largest C iv blueshifts (in the bottom right of Fig. 3.18) demonstrate that non-virial motions are having a significant effect on the C iv BLR dynamics. At fixed emission-line EW, almost 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 symmetric component. 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.

Quantitatively, the C iv-shape change is most apparent from the FWHM values, which are strongly correlated with the C iv-blueshift. This trend has previously been identified, by comparison with Mg II at lower-redshifts (Shen et al., 2008; Shen et al., 2011) and $\text{H}\beta$ at higher redshifts (Shen and Liu, 2012). We find that virial BH mass estimates based on the C iv FWHM will overestimate the true mass by a factor of ~ 5 for objects exhibiting the largest C iv blueshifts.

In contrast, the C iv line dispersion does not show a similarly strong dependence on the blueshift. This is a result of the shape of the

C iv line profile being dependent on the blueshift; the low-blueshift profiles are peakier ($\text{FWHM}/\sigma \simeq 1$) than the high-blueshift profiles ($\text{FWHM}/\sigma \simeq 2$). 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. Their investigation 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 FWHM/σ -based correction to the C iv FWHM proposed by Denney, (2012) increases the inferred BH masses by ~ 0.2 dex at the low end of our C iv blueshift distribution, thereby reducing part of the systematic trend in the BH mass (Fig. 3.23). However, it is not applicable at the high C iv blueshift end, where velocity widths are likely dominated by non-virial motions. Based on the typical line shape of C iv in these high-blueshift quasars ($\text{FWHM}/\sigma \simeq 2$), the Denney, (2012) correction decreases the predicted BH masses by ~ 0.3 dex, which has only a moderate decreasing effect on the strong systematic (Fig. 3.23).

While the C iv-line dispersion is largely independent of the blueshift, it does not follow that dispersion-based BH-mass estimates are correct, because the underlying assumption regarding the virial-origin of the C iv emission profile breaks down at large blueshifts. Furthermore, given the difficulty in obtaining reliable dispersion measurements from survey-quality spectra with limited S/N (e.g. Mejía-Restrepo et al., 2016), virial BH-mass estimates for existing large samples of high-redshift quasars are usually based on the FWHM (e.g. Shen et al., 2011). Our work therefore suggests that a viable recipe for obtaining more reliable BH mass estimates for large numbers of quasars at high redshift is to measure both the FWHM and the blueshift, which together can be used to derive a FWHM corrected for the non-virial contribution.

Although we do not have enough quasars in our sample to derive a reliable quantitative correction to BH-masses as a function of C iv blueshift, we are assembling a large sample of quasar spectra with coverage of both C iv and $\text{H}\beta/\text{H}\alpha$ in order to empirically validate such a correction. Both Runnoe et al., (2013a) and Shen and Liu, (2012) considered a similar approach, but concluded that a low-ionisation broad line (e.g. Mg II), or features from the quasar NLR, are required to determine the systemic redshift and hence the C iv blueshift. Empirical tests of the reliability of the improved Hewett and Wild, (2010) redshifts for the SDSS DR7 quasars (Shen et al., 2016) and the availability of the largely SED-independent principal component analysis redshifts for DR10 and DR12 (Pâris et al., 2014; Pâris et al., 2017) already allow meaningful corrections to BH-mass estimates for quasars exhibiting large C iv-blueshifts. Our intention is, however, to determine a definitive correction formula using the redshifts from Allen & Hewett (2016, in preparation) for both DR7 and DR12.

Given the difficulty of measuring reliable C IV blueshifts, 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/H β residuals, as well as with the strengths of optical O III and Fe II. The strengths of optical O III and Fe II, being parameters in EV1, are also correlated with the C IV blueshift (Sulentic et al., 2007). As the C IV blueshift increases the EW decreases systematically (Fig. 3.18). The Si IV+O IV emission-line blend, however, shows significantly less systematic variation. Therefore, while the Runnoe et al., (2013a) Si IV+O IV-based correction is effective in practice, the C IV blueshift measurement provides a more direct measure of the non-virial contributions to the C IV velocity width.

In our sample, we find the 1400 Å/C IV peak flux ratio to be strongly correlated to the C IV blueshift (the Spearman coefficient for the correlation is 0.82, p-value 2E-5). As such, the correction to C IV-based virial masses proposed by Runnoe et al., (2013a, their equation 3) removes a large part of the systematic in the H α /C IV FWHM residuals with the C IV blueshift (Fig. 3.23); the median C IV/H α FWHM ratio at large C IV blueshifts ($>1200\text{km s}^{-1}$) is reduced from 4.5 to 1.5. However, at low ($<1200\text{km s}^{-1}$) C IV blueshifts, the trend for C IV to predict lower BH masses persists, and the scatter between the C IV- and H α -based masses increases by a factor of two. In accordance with our expectations, we find the FWHM C IV/H α residuals to be more tightly correlated to the C IV blueshift (Spearman coefficient 0.82, p-value 3E-5) than to the 1400 Å/C IV peak flux ratio (Spearman coefficient 0.72, p-value 7E-4).

3.8.2 Possible systematic trends in H α BH-mass estimates

In Section 3.7.5, we found that the quasars with large C IV blueshifts have systematically narrower H α FWHM (Fig. 3.7). Using the standard virial BH mass calibrations this implies that the high C IV blueshift population have high accretion rates ($L/L_{\text{Edd}} \simeq 1$; Fig. 3.22). This interpretation requires some caution since the emission-line shape (characterized by the value of FWHM/σ) of H α is also changing as a function of the C IV blueshift (Fig. 3.7). At low C IV blueshifts there are a range of shapes, but all of the quasars exhibiting large C IV blueshifts have peaky H α profiles with $\text{FWHM}/\sigma \simeq 1$. This raises the question of whether the H α FWHM is a reliable proxy for the virial-induced velocity dispersion for the full range of H α 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 factor, f , is a factor ~ 2

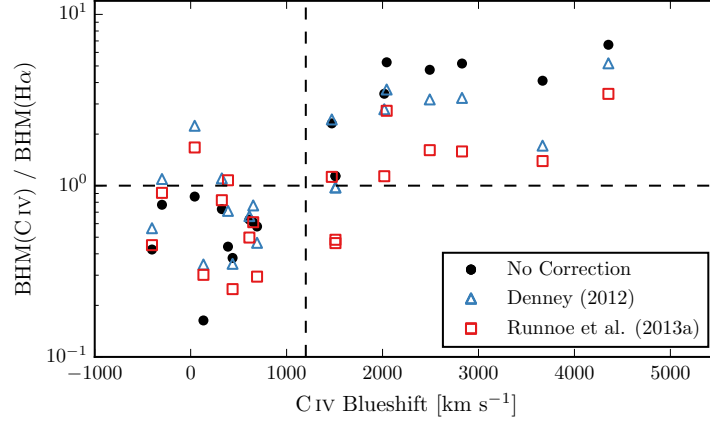


Figure 3.23: Comparison of BH mass estimates derived from the FWHM of C IV and H α as a function of the C IV blueshift (*black circles*), and after applying corrections to the C IV-derived mass based on the line shape (*blue triangles*; Denney, 2012) and the peak flux ratio of the Si IV+O IV blend relative to C IV (*red squares*; Runnoe et al., 2013a). While the shape-based correction improves the consistency between BH mass estimates in the low-blueshift population, only the Si IV+O IV/C IV peak flux-based correction is effective at high blueshifts (although a weak systematic remains).

larger for their Population ‘1’ sources (with $\text{FWHM}/\sigma < 2.35$ and essentially equivalent to population A of Sulentic and co-workers and to the high-blueshift quasars here) than for their Population 2 (with $\text{FWHM}/\sigma > 2.35$). For single-epoch BH-mass estimates, assuming a constant value of f , as is normally done (e.g. Vestergaard and Peterson, 2006), means that Population 1 masses will be underestimated and Population 2 will be overestimated. In the context of this result from Collin et al., (2006), our high-blueshift objects all possess peaky H α -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 Fig. 3.16 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 $H\beta$ FWHM relates not to a spread in BH masses, but rather to the orientation of the BLR relative to the line-of-sight. 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 $H\beta$ FWHM and hence with the BH-mass estimates. This raises the question of whether the narrow $H\alpha$ emission lines observed in the quasars with the largest C IV blueshifts could be an orientation effect. However, there is no evidence that the C IV blueshift is dependent on the orientation (inferred from the radio core-dominance; Richards et al., 2011; Runnoe et al., 2014). Furthermore, Leighly, (2004) showed that the He II $\lambda 1640$ emission-line properties of quasars with large C IV blueshifts are more consistent with differences in the SED rather than differences in the orientation. Collin et al., (2006) showed that orientation effects were also sub-dominant to the Eddington ratio in determining the shape of the $H\beta$ line and the $H\alpha$ line shape trend we observe is consistent with the finding of Marziani et al., (2003) that the $H\beta$ 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 $H\alpha$ line profile.

3.8.3 Accretion-rate trends in the quasar population

The blueshifting of C IV is usually interpreted as 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. This picture is therefore consistent with our finding that the quasars with large C IV blueshifts have high accretion rates.

All of the objects in our sample which exhibit large C IV blueshifts would be classified as population A in the Sulentic et al., (2000) scheme based on the $H\alpha$ FWHM (see Section ??). 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, 2005a).

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).

3.8.4 *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 (Section ??) 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-H α 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 paper (see Runnoe et al., 2013b, for such a study).

3.8.5 *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 DR7 quasar catalogue, from which our quasar sample is effectively drawn (Section ??). In order to estimate the size of the selection effect, we considered the detection probability for a much-simplified quasar population. We assume that all quasars with C iv blueshifts $> 1200 \text{ km s}^{-1}$ have enhanced accretion rates relative to the ‘normal’

population (with C iv blueshifts $< 1200 \text{ km s}^{-1}$). If the accretion rate of the high-blueshift population is double the rate of the low-blueshift population (which is true in an average sense – see Fig. 3.22), 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 continuous¹⁰ we thus use the number counts at $m_i < 19.1$ and $m_i < 18.35$, which differ by a factor of $\simeq 4$.

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

3.9 CONCLUSIONS

The main results of this paper are as follows:

- We have analysed the spectra of 230 high-luminosity ($10^{45.5} - 10^{48} \text{ erg s}^{-1}$), redshift $1.5 < z < 4.0$ quasars for which spectra of the Balmer emission lines and the C iv emission line exist. The large number of quasars in our spectroscopic catalogue and the wide range in C iv blueshifts the quasars possess has allowed us to directly investigate biases in C iv-based BH mass estimates which stem from non-virial contributions to the C iv emission as a function of the C iv blueshift, which, in turn, depends directly on the form of the quasar ultraviolet SEDs (Richards et al., 2011).
- The C iv emission-based BH-masses are systematically in error by a factor of more than five at 3000 km s^{-1} in C iv emission blueshift and the overestimate of the BH-masses reaches a factor of 10 for quasars exhibiting the most extreme blueshifts, $\gtrsim 5000 \text{ km s}^{-1}$.
- We have derived an empirical correction formula for BH-mass estimates based on the C iv emission line FWHM and blueshift.

¹⁰ The luminosity function and number-counts vary only smoothly (e.g. Ross et al., 2013) for the magnitude and redshift range used here.

The correction may be applied using equations 4 and 6 in Section 3.6.3. The large SED-dependent systematic error in C iv-based BH-masses is removed using the correction formulae. The remaining scatter between the corrected C iv-based masses and the H α -based masses is 0.24 dex at low C iv blueshifts ($\sim 0 \text{ km s}^{-1}$) and 0.10 dex at high blueshifts ($\sim 3000 \text{ km s}^{-1}$). This is a significant improvement on the 0.40 dex scatter observed between the un-corrected C iv and H α 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 $z \sim 5$.

As discussed in Section 3.7.1, uncertainties in redshift estimation and hence the definition of the systemic rest-frame for quasars impact on the accuracy of the corrected BH-masses. 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 3.6.3 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. Allen & Hewett will publish improved redshifts for all quasars in the SDSS DR7 and DR12 which will reduce SED-dependent systematic errors below the apparent inherent dispersion of $\simeq 220 \text{ km s}^{-1}$ associated with broad emission line redshifts (Shen et al., 2016). At the same time we will publish catalogues of unbiased BH masses for both 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.

We have presented an analysis of biases in C iv-derived virial BH masses of high-luminosity ($L_{\text{bol}} \sim 47 \text{ erg s}^{-1}$) quasars at redshifts ~ 2.5 from the Sloan Digital Sky Survey. Many authors have reported a large scatter between C iv- and H α /H β -based masses, and part of this scatter has been shown to correlate with the C iv blueshift (Shen and Liu, 2012). The blueshifting of C iv is usually interpreted as evidence for strong outflows which, most likely, result from the presence of a radiation line-driven accretion-disc wind. Our study is the first to examine this bias systematically across the full range of C iv-emission line blueshifts observed in the SDSS sample. In particular, we have used rest-frame optical spectra of 19 quasars in the redshift range

$2 < z < 2.7$ to directly compare C iv and H α emission properties as a function of the C iv blueshift. We reach the following conclusions:

- A strong correlation between C iv-velocity width and blueshift is found and at large blueshifts, $>2000 \text{ km s}^{-1}$, the velocity-widths are dominated by non-virial motions. This suggests that the assumption that velocity-widths arise from virial-induced motions, on which single-epoch BH-mass measurements are predicated, is not straightforwardly applicable to these high-blueshift quasars.
- We use the H α emission line to provide BH-mass estimates that are unbiased by non-virial contributions to the velocity width. We find that the C iv-based BH masses of quasars with low C iv blueshifts are systematically underestimated (by a factor of ~ 1.7) whereas the masses of quasars with large blueshifts are severely overestimated (by a factor of ~ 5).
- We find a systematic change in the shape of the H α line profile as a function of the C iv blueshift. Specifically, the H α line profiles of the quasars with high C iv blueshifts are all ‘peaky’ with FWHM/σ close to unity.
- We suggest that the high C iv blueshift quasars are high Eddington-ratio objects that are inherently rare (comprising $\sim 15\%$ of the UV-selected sample), but are being boosted in number by a factor of ~ 4 in the flux-limited SDSS sample.

With a relatively small sample of 19 quasars we have been able to uncover systematic trends in the C iv and H α emission line shapes as a function of the C iv blueshift. This confirms the prospect of developing a blueshift-dependent correction to C iv-based single-epoch BH-mass estimates using a larger samples of luminous quasars with both rest-frame UV and rest-frame optical spectroscopy. We are currently in the process of assembling such a sample, which will contain ~ 300 luminous quasars, 80 per cent at redshifts $z \geq 2$. A new SED-independent redshift-estimation algorithm (Allen & Hewett 2016, in preparation) makes it possible to quantify the distribution of C iv-blueshifts in the observed quasar population as a whole, thereby allowing us to make empirical corrections to C iv-based BH-masses for all luminous, high-redshift SDSS/BOSS quasars.

NARROW LINE REGION PROPERTIES

4.1 INTRODUCTION

Correlations between the masses of super-massive black holes and properties of the host galaxy bulges in nearby galaxies (Gebhardt et al., 2000; Ferrarese and Merritt, 2000) and the similarity of the cosmic evolution of star formation and black hole activity (Boyle and Terlevich, 1998; Madau and Dickinson, 2014) suggest that the formation and evolution of supermassive black holes and their host galaxies are linked. Active galactic nuclei (AGN) and quasar feedback, in which star formation in the host galaxy is suppressed by the energy output from the quasar, is a favoured model. This has motivated a considerable amount of observational work searching for feedback signatures (for recent reviews, see Alexander and Hickox, 2012; Fabian, 2012; Heckman and Best, 2014).

The [O III] $\lambda\lambda 4960, 5008$ doublet is the strongest narrow line found in optical quasar spectra. It is forbidden line which traces gas on kiloparsec scales. High velocity dispersions or strong asymmetries in forbidden lines are evidence of high velocity extended ionized outflows. Recent studies have provided constraints on the prevalence of ionized outflows traced by [O III] emission in low-redshift type 2 AGN (e.g. Mullaney et al., 2013; Zakamska and Greene, 2014) using spectra from the Sloan Digital Sky Survey (SDSS; York et al., 2000). However, the signatures of feedback are expected to be stronger at higher redshifts ($z \sim 2$) when the black hole accretion and star formation rates both peak.

In this paper we analyse the [O III] properties of a sample of 358 high-luminosity, redshift $1.5 < z < 4$ quasars. In particular, we search for signs of outflowing gas, which include asymmetries and broad velocity widths. Recent near-IR spectroscopy of $z > 1.5$ quasars often report exceptionally large [O III] widths, with $\text{FWHM} > 1000 \text{ km s}^{-1}$ (e.g. Netzer et al., 2004; Nesvadba et al., 2008; Kim et al., 2013; Brusa et al., 2015; Carniani et al., 2015; Perna et al., 2015). **Add reference to WIISH paper.**

To date, the focus has been extreme, rare, AGN sub-populations, e.g. highly reddened and obscured AGN (Banerji et al. 2012, 2013, 2014, Eisenhardt et al. 2012, Ross et al. 2014, Zakamska et al. 2016, WIISH paper). These populations provide insight into relatively short-lived phases in the AGN activity cycle. A more complete understanding of AGN-evolution and the physical link with the host galaxies requires

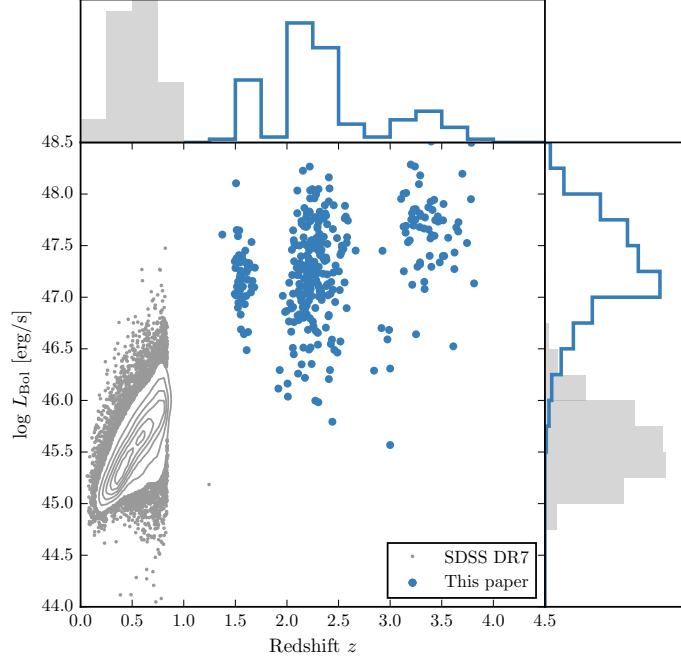


Figure 4.1: The ranges in redshift and luminosity covered by our sample, relative to the redshift-luminosity distribution of the SDSS DR7 quasar catalogue with measured [O III] line properties (Shen et al., 2011). The gaps in our sample coverage at $z \sim 1.8$ and $z \sim 3$ are due to the near-infrared transmission windows. In regions of high point-density, contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator.

the rare, extreme, sub-populations to be placed in the context of the AGN population as a whole.

For a subsample where we have spectra covering the rest-frame ultra-violet line C IV, we can test whether the strong outflows, inferred the blueshifting of C IV relative to the systemic redshift, have any effect on the host galaxy on larger scales.

Throughout this paper we adopt a flat Λ CDM cosmology with $h_0 = 0.69$ and $\Omega_M = 0.29$.

The paper is structured as follows.

4.2 QUASAR SAMPLE

We have assembled a catalogue of 358 high-luminosity, redshift $1.5 < z < 4$ quasars, with spectra taken in the near-infrared. At these redshifts the spectra cover the rest-frame optical region, which includes the broad Balmer H β line and the strong, narrow [O III] doublet. There is a significant overlap between the sample presented here and the one used in Coatman et al. (2016b), hereafter paper I. We briefly

describe the catalogue below, but refer the reader to Coatman et al. (2016b) for further details, including the reduction procedure.

Two hundred are selected from the SDSS Seventh Data Release (DR7; Schneider et al., 2010). These are observed with LIRIS (Machado et al., 1998) on the William Herschel Telescope (WHT), Triple-Spec (Wilson et al., 2004) on the Astrophysics Research Consortium (ARC) 3.5 m telescope, FIRE (Simcoe et al., 2010) on the Magellan-Baade telescope, the SINFONI integral field spectrograph (Eisenhauer et al., 2003; Bonnet et al., 2004) on the European Southern Observatory (ESO) Very Large Telescope (VLT), SOFI (Moorwood, Cuby, and Lidman, 1998) on the New Technology Telescope (NTT), and TRIPLE-SPEC on the Palomar 200-inch Hale telescope (P200). One hundred and seventeen were observed to measure accurate redshifts of quasar pairs at very close projected separations (Prochaska and Hennawi, 2009; Lau, Prochaska, and Hennawi, 2015; Hennawi et al., 2015). Spectra were taken with GNIRS (Elias et al., 2006) on the Gemini North telescope, ISAAC (Moorwood et al., 1998) on the European Southern Observatory (ESO) Very Large Telescope (VLT), NIRSPEC (McLean et al., 1998) on the Keck-II telescope, NIRI (Hodapp et al., 2003) also on Gemini North, and XSHOOTER (Vernet et al., 2011), again, on the VLT. Seventy-nine are bright Hamburg-ESO quasars observed with SINFONI, ninety are quasars with archival high-resolution optical spectra observed with SOFI, and four are observed with XSHOOTER and presented in the XQ-100 catalogue (Lopez et al., 2016).

This is the largest study of the narrow line region properties of high- z quasars ever undertaken. The redshift and luminosity coverage of the quasars in our sample is shown in Fig. 4.1, and the quasar sample is summarised in Table 4.1. Our sample covers much higher redshifts and luminosities than the SDSS sample.

4.3 PARAMETERIC MODEL FITS

In this section we describe how we measure the [O III] velocity-width and parameterize asymmetries in the line. We first fit a parameteric model to the [O III] emission, and the nearby $H\beta$ peak. This step is taken solely to enhance the signal-to-noise (SNR) of the spectra and to decompose the emission from $H\beta$ and the [O III] doublet. We then calculate non-parameteric measures of the line profile from the best-fitting model.

Before the model can be fit, the spectra must first be transformed into the (approximate) quasar rest-frame. This transformation is only for practical purposes and the emission line parameters we will go on to derive are independent of the exact redshift used. The redshift will later be refined using the results from the fit. The redshift used in this transformation is either derived from a multi-component Gaussian fit to the broad $H\alpha$ emission (~ 40 per cent of our sample) or, when this

Table 4.1: The numbers of quasars with [O III] line measurements and the spectrographs and telescopes used to obtain the near-infrared spectra. Further details on the instrumental configurations are given in paper I. **Check numbers.**

Spectrograph	Telescope	Number
FIRE	MAGELLAN	32
GNIRS	GEMINI-N	29
ISAAC	VLT	8
LIRIS	WHT	5
NIRI	GEMINI-N	29
NIRSPEC	Keck II	3
SINFONI	VLT	76
SOFI	NTT	78
TRIPLESPEC	ARC-3.5m	27
TRIPLESPEC	P200	46
XSHOOTER	VLT	25
		144

is not possible, from a preliminary fit to the broad H β emission (~ 40 per cent) or narrow [O III] emission (20 per cent). **Should say whether we use peak/median etc.**

The continuum and optical Fe II emission is modelled and subtracted using a recipe that is identical to the one described in paper I. The following model is then fit to the spectra in the wavelength interval 4700–5100 Å. The fit is done as a function of the Doppler velocity shift, and we adopt the wavelength 4862.721 Å (the laboratory H β wavelength) to transform wavelengths into equivalent Doppler velocities.

H β is modelled by two Gaussians with non-negative amplitudes and FWHM greater than 1200 km s^{-1} . A handful we fix the centroids of the two Gaussians to be the same, normally because of very low S/N or because the blue wing is below the lower wavelength limit of the spectrograph. Any contribution to the H β emission from the narrow-line region is weak in the vast majority of our sample, and so we do not include an additional Gaussian component to model this emission. Note: we do in fact include narrow components for eight objects in our sample. Although, for some objects, this could bias our estimate of the velocity-width of the broad component, this information is not used in the analysis presented in this paper.

Each component of the [O III] doublet is fit with one or two Gaussians, depending on the fractional reduced χ^2 difference between the one- and two-component models. If the addition of the second

Table 4.2: Approximate physical origin of the ICA components.

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

Gaussian decreases the reduced χ^2 by more than 5 per cent then the double-Gaussian model is accepted. One hundred and twenty-five are fit with a single Gaussian, 147 with two Gaussians, and [O III] is undetected in a further 78 quasars. When a single Gaussian is used to model each line, the peak flux ratio of the [O III] 4960 Å and 5008 Å components are fixed at the expected 1:3 ratio and the width and velocity offsets are set to be equal. In the double Gaussian fit, the peak flux ratio of the second components is again fixed at 1:3, and the width and velocity offsets are again set to be equal.

In six quasars a significantly better fit was obtained by allowing the flux ratio between the two components to vary. In these quasars the best-fitting peak ratio varies from 0.50 to 0.84, with mean 0.70. **Check ICA component fits to see if this looks real.**

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 telluric absorption or of low S/N were masked out.

However, the fits aren't always good. How to be quantative about this? Chi-squared values indicating bad fits? Also the fact that we sometimes model Fe as OIII. Discuss the limitations of using the Boroson & Greene template? Even if don't use results, should discuss Gaussian fits and their limitations, because that's what most people will be using.

4.3.1 Derived Parameters

All [O III] line properties are derived from the [O III]5008 emission, but, as described below, the kinematics of the peak at 4960 Å are constrained by our fitting routine to be identical.

We do not to attach any physical meaning to the individual Gaussian components used in the model. While it is true that in some quasars the [O III] emission can be clearly separated in to a narrow component at the systemic redshift and a lower-amplitude, blueshifted broad component (e.g. Shen, 2016), often this decomposition is highly

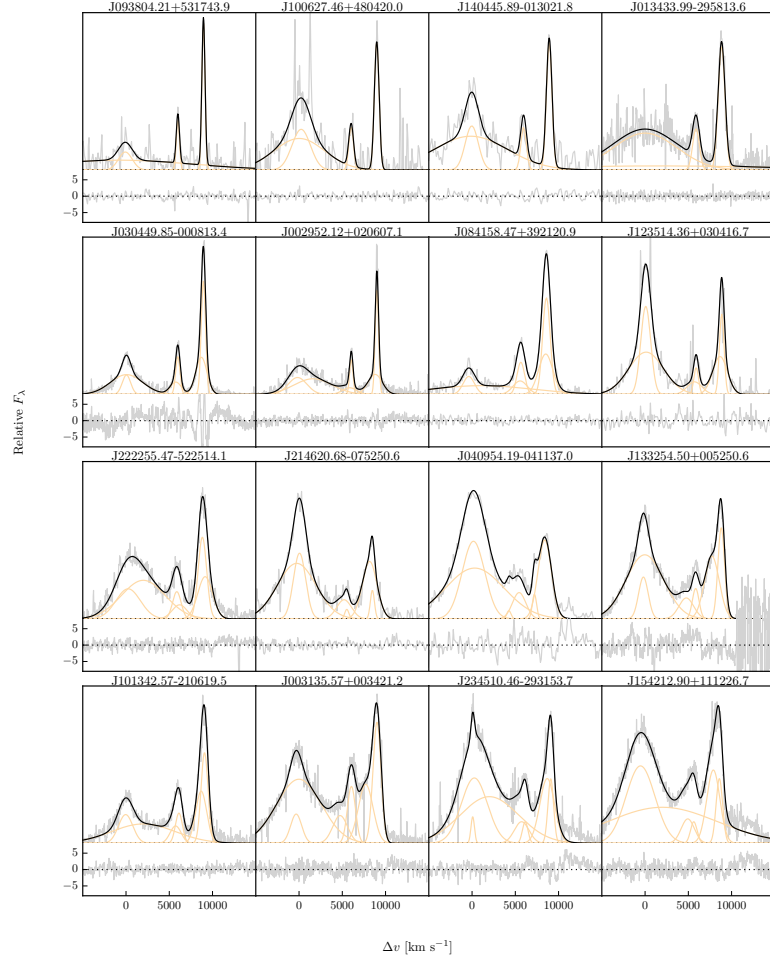


Figure 4.2: Multi-component Gaussian fits to the continuum-subtracted $H\beta/[O\text{III}]$ emission in 16 quasars, chosen to be the representative of the wide range of $[O\text{III}]$ line widths we measure in our sample. The data is shown in grey, the best-fitting parametric model in black, and the individual model components in orange. The broad $H\beta$ centroid is used to measure the systemic redshift, and Δv is the velocity shift from the line rest-frame transition wavelength for $H\beta$. Below each fit we plot the data minus model residuals, scaled by the errors on the fluxes. **Resample model at higher resolution.**

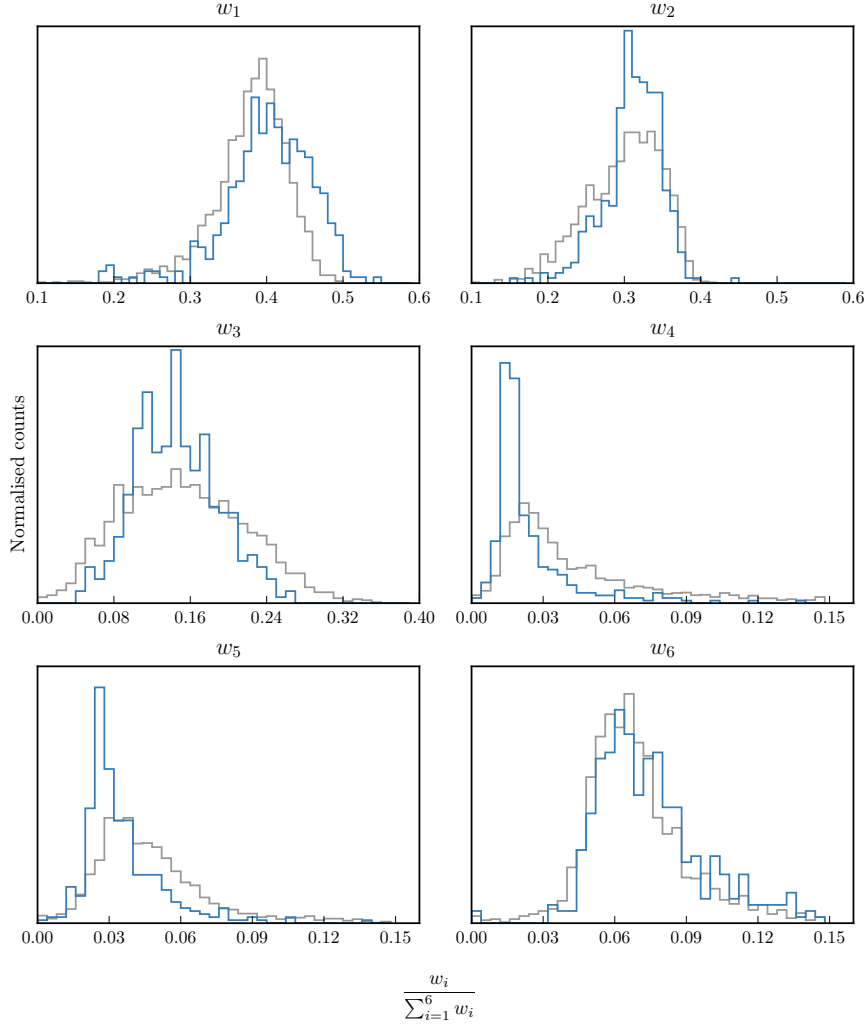


Figure 4.3: The relative weight in each of the six positive ICA components for the high-luminosity (blue) and low luminosity samples (grey). In the high-luminosity sample Fe II emission is stronger (component w_1). The core [O III] emission is weaker (components w_4 , w_5) but the strength of the blueshifted wing is the same (w_6).

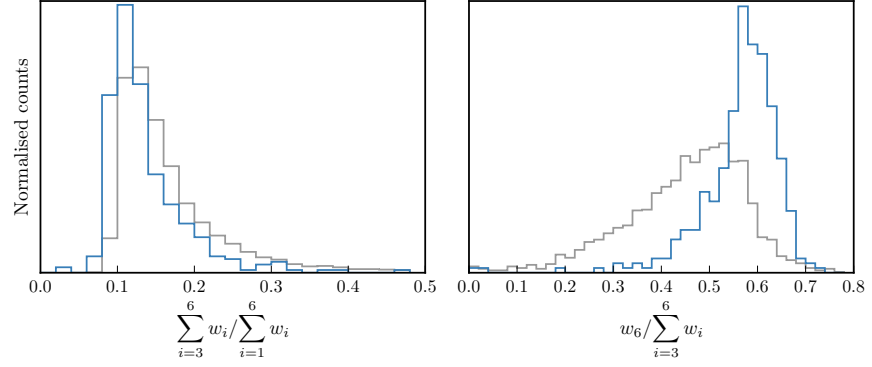


Figure 4.4: The relative weight in the three ICA components corresponding to [O III] emission (*left*) and the relative weight of the component most closely related to blueshifted [O III] emission relative to all three [O III] components (*right*). [O III] emission is weaker in the high-luminosity sample, but the relative contribution but the fractional contribution from the blueshifted component to the total [O III] emission is higher. Hence [O III] is weaker, broader, and more asymmetric in the high-luminosity sample. See Zakamska discussion.

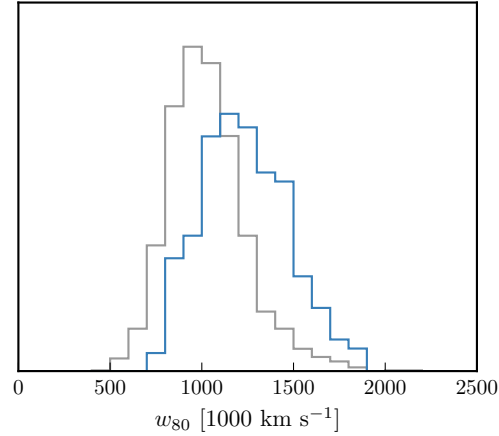


Figure 4.5: Comparison of [O III] velocity-widths in the high and low luminosity samples using the ICA component fits. If keep this need to explain in text how w_{80} is calculated from ICA component fits.

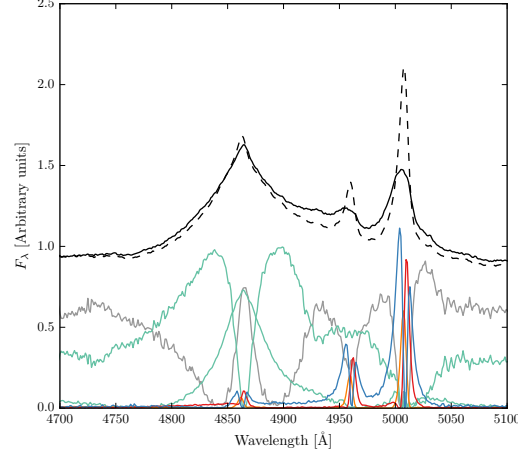


Figure 4.6: Black solid line is the median from the ICA fits to the high-luminosity sample. Black dashed line shows the median from the low-luminosity sample. The six positive ICA components are also shown.

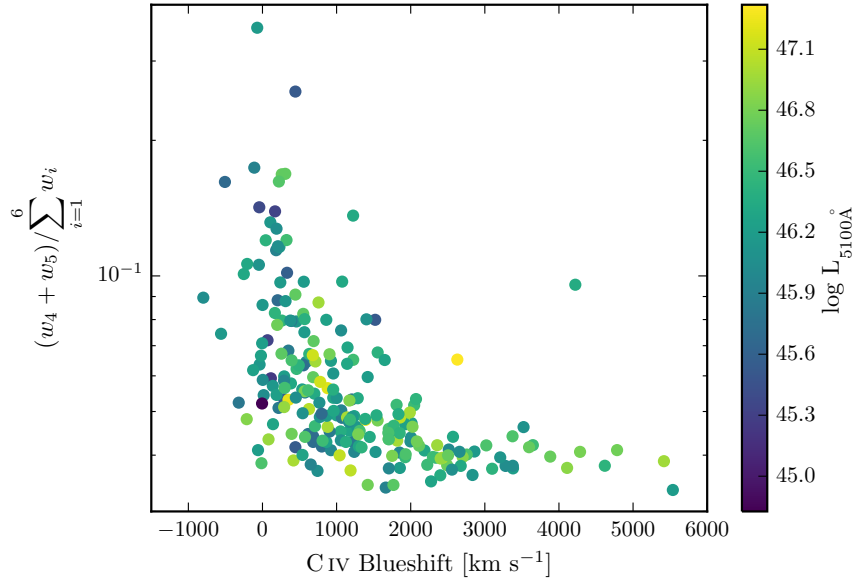


Figure 4.7: [O III] strength decreases as the C IV blueshift increases, I run in to problems comparing the C IV blueshift to the [O III] blueshift / velocity-width. See similar thing if I use [O III] EQW instead. Need to fix y ticks. Only showing the core components here. The C IV blueshift is now measured relative to the NIR ICA redshift. I think this trend is mostly being driven by the Eigenvector 1 correlations: as the blueshift increases the Fe II strength increases and the [O III] strength decreases. Doesn't appear to be driven by the luminosity. Is this tighter than EV1 trend shown with Fe/OIII strength by other authors? Is the AGN NLR absent in objects where outflows have reached kiloparsec scales, sweeping up the low-density material responsible for the [OIII]-emission?

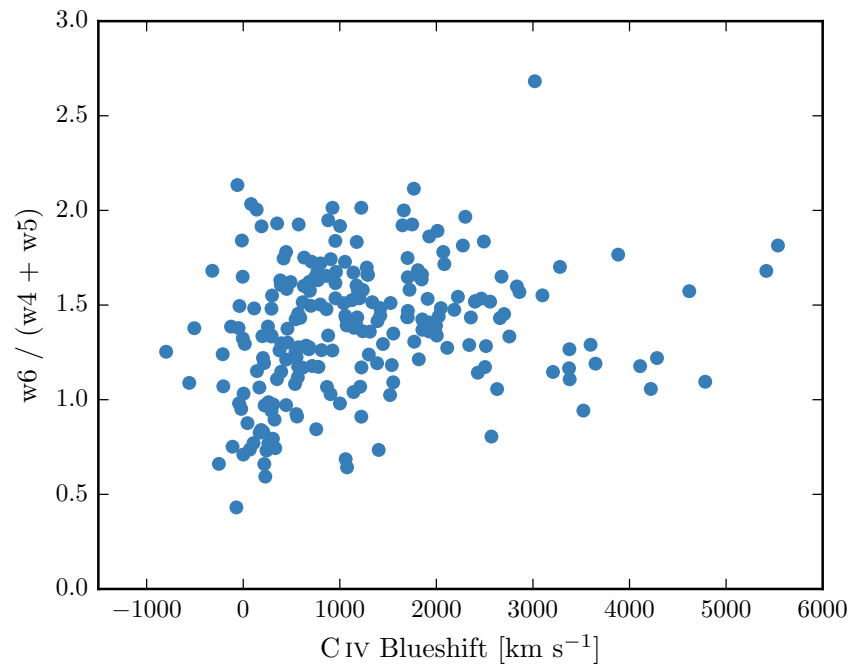


Figure 4.8: I think there is a trend here but at high blueshifts the OIII is undetected / very low S/N. Need to determine when we can believe OIII parameters. Why at low C IV blueshift is there a much bigger dynamic range than in [O III] blueshifts in Fig. 13. Is it just because we have more objects?

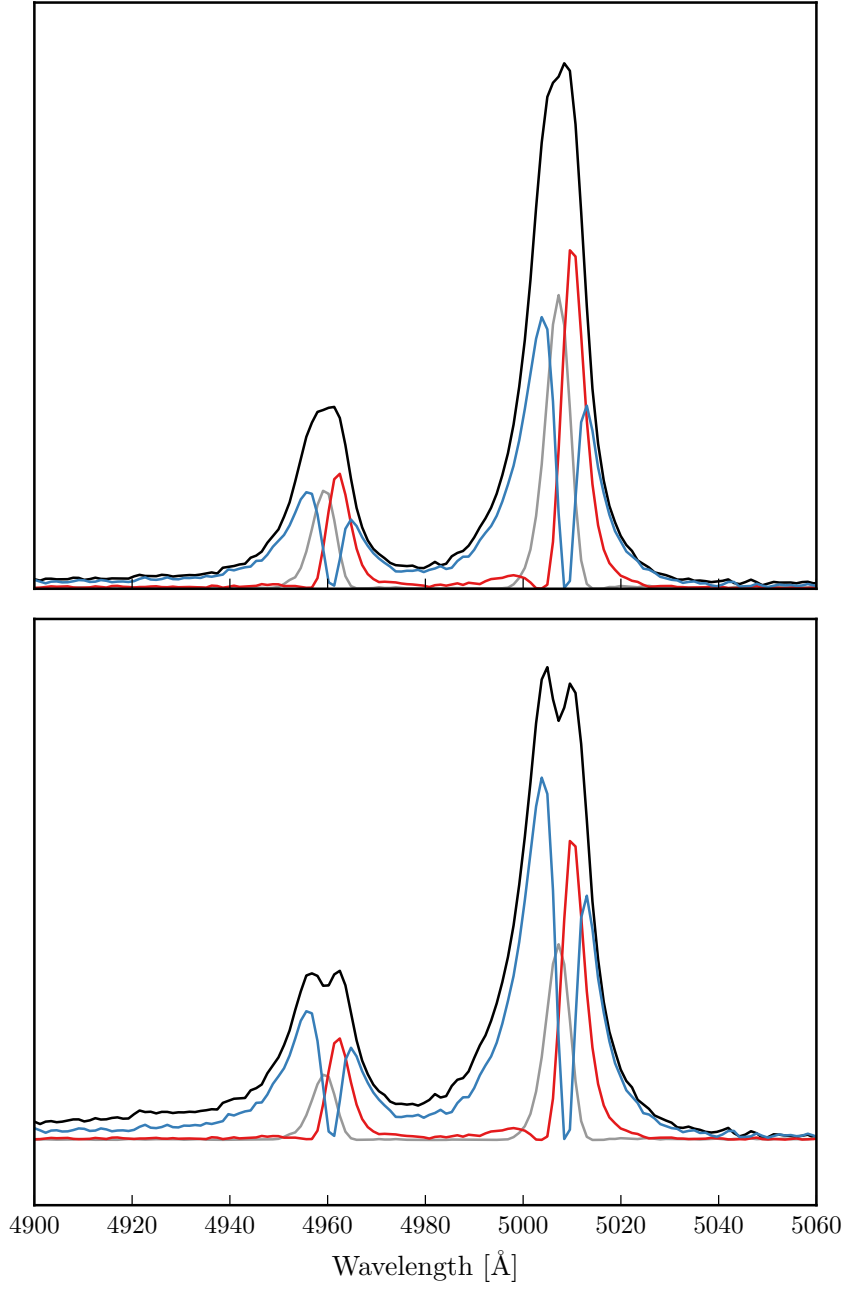


Figure 4.9: Comparison of median [O III] profiles from ICA fits to low- and high-luminosity samples.

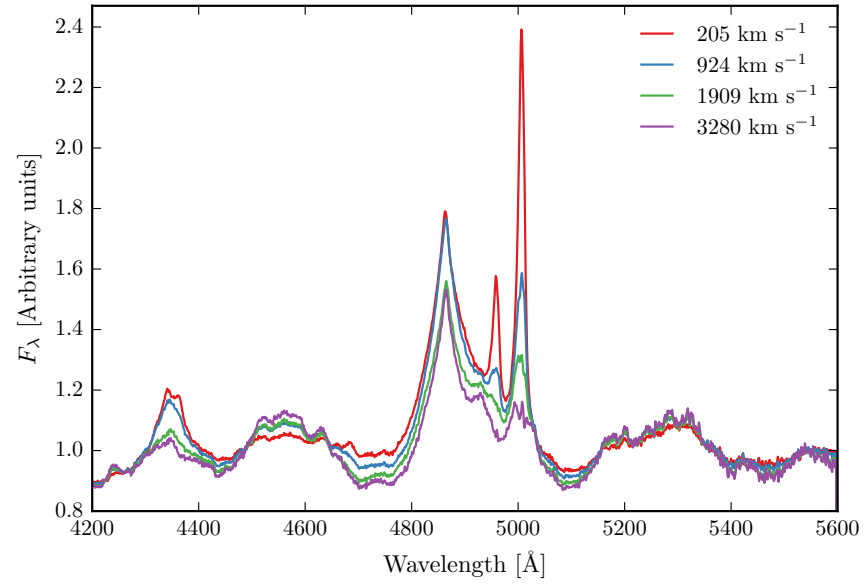


Figure 4.10: ICA median weights as a function of the CIV blueshift.

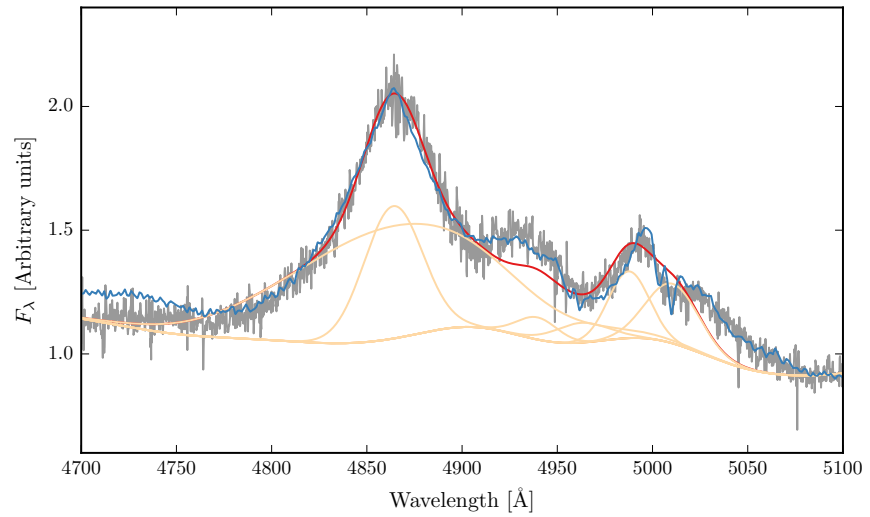


Figure 4.11: Example where poorly subtracting the iron can be confused with [O III].

uncertain. Furthermore, often the S/N is not sufficient to statistically justify the addition of a second Gaussian component. Instead, we characterize the [O III] line profile using a number of non-parametric measures, which are commonly used in the literature (e.g. Zakamska and Greene, 2014; Zakamska et al., 2016). A normalised cumulative velocity distribution is constructed from the best-fitting model, from which the velocities below which 5, 10, 25, 50, 75, 90, and 95 per cent of the total flux accumulates can be read off. The width of the emission line can then be defined, for example, using $w_{80} = v_{90} - v_{10}$. The absolute asymmetry in the line profile A is defined as $((v_{95} - v_{50}) - (v_{50} - v_5)) / (v_{95} - v_5)$ (Zakamska and Greene, 2014).

We also define the blueshift of the [O III] emission, which is a measure of the velocity shift of the profile from the expected position. This requires a measure of the observed line position, and an accurate measurement of the quasar systemic redshift. We use v_{10} to measure the location of the [O III] emission. Note that v_{50} is not suitable because when [O III] is low S/N we fit with single Gaussian.

The line width measures are not at present corrected for instrumental broadening, but this can easily be done.

4.3.2 Other flags

4.3.2.1 Flag 2

Low S/N. Includes some of the strong iron emitters.

4.3.2.2 Flag 3

74 objects where [O III] is undetected, although I don't have a rigorous definition of what this means. Merge with 2?

4.4 ICA COMPONENT FITS

What is the segue here? Sometimes Gaussians give poor fit? Not clear what is Hb, what is OIII, what is Fe? Show some examples of when the multi-Gaussian fits fail. Need to describe sample used in ICA component decomposition and briefly describe method (or refer to Allen & Hewett).

4.5 MEASURING THE QUASAR SYSTEMIC REDSHIFT

Explain flags Need to send paul redshifts

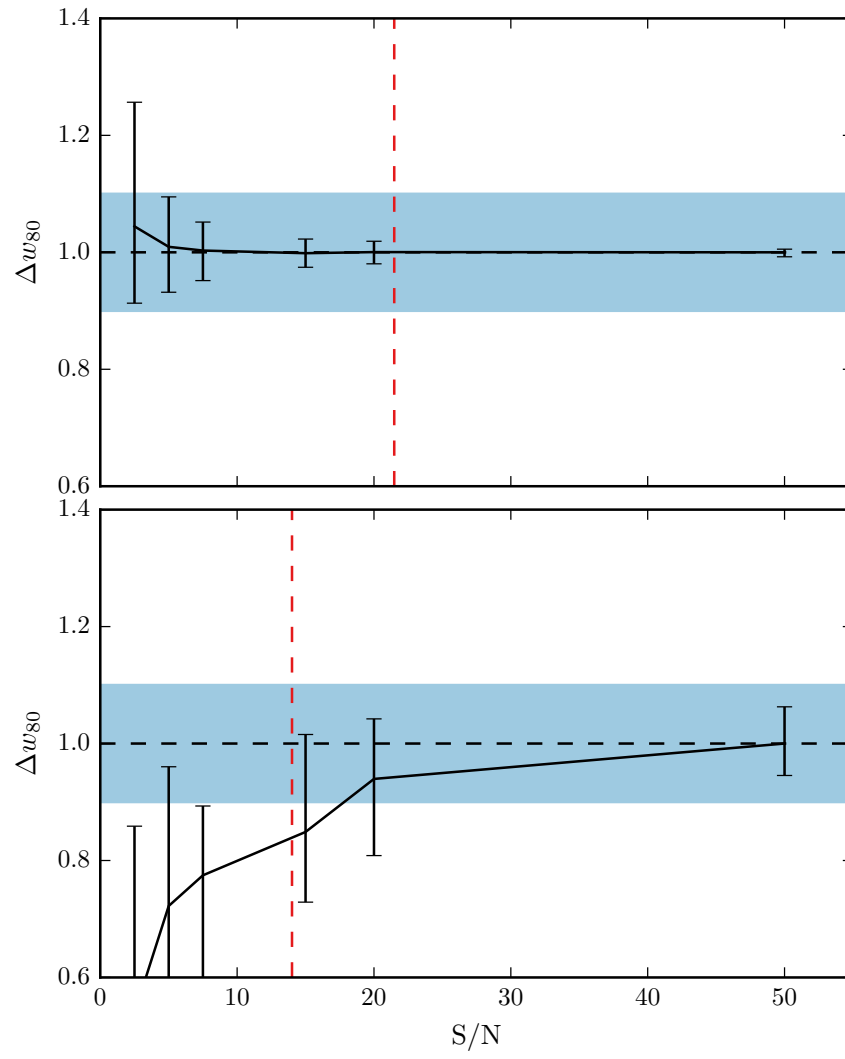


Figure 4.12

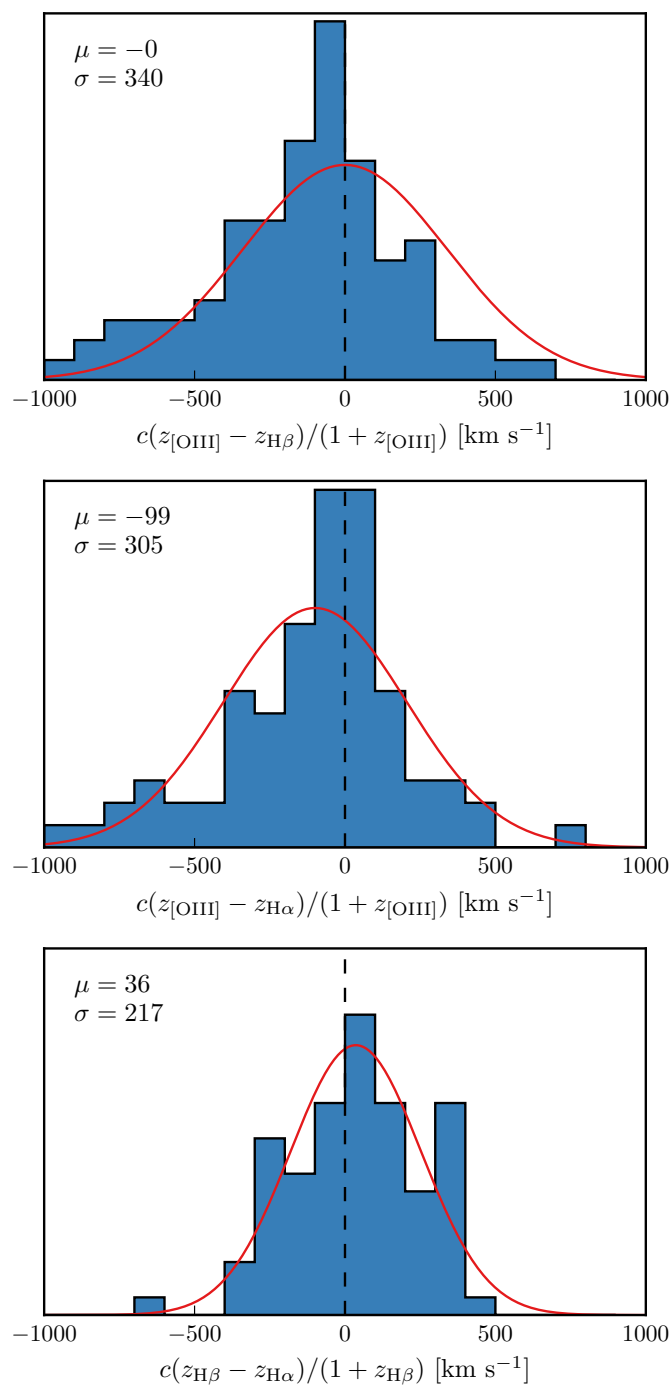


Figure 4.13: Redshift comparisons. Lots have been excluded from Ha/Hb so need to look at flags greater than one. What is the big peak? Gaussian fit to the first one has failed. Find out why these plots look different to ones in paper.

4.5.1 $H\alpha$

There are 224 quasars in our sample with spectra covering the $H\alpha$ emission line. We discard seven of these from our sample because of very low S/N (<2.5 measured in the $H\alpha$ line), leaving 217. To measure the position of the line we fit a parameteric model, which is very similar to the model described in Paper I. The continuum emission is first modeled and subtracted using the procedure described in Paper I. We then test five different models with increasing degrees of freedom to model the $H\alpha$ emission. The model we select is the simplest model for which the fractional change in the reduced chi-squared from the model with the lowest reduced chi-squared is less than ten per cent.

The models we test are: (1) a single broad Gaussian; (2) two broad Gaussians with identical velocity centroids; (3) two broad Gaussians with different velocity centroids; (4) two broad Gaussians with identical velocity centroids, and additional narrower Gaussians to model the narrow $H\alpha$ emission, and the narrow components of $[N\text{ II}]\lambda\lambda 6548, 6584$ and $[S\text{ II}]\lambda\lambda 6717, 6731$; (5) two broad Gaussians with different velocity centroids, and additional narrower Gaussians. If used, the width and velocity of all narrow components are set to be equal in the fit, and the relative flux ratio of the two $[N\text{ II}]$ components is fixed at the expected value of 2.96. The number of quasars fit by each model is: model 1 - 10; model 2 - 71; model 3 - 32; model 4 - 51; model 5 - 53. The redshift is then measured at the peak flux of the $H\alpha$ model, including both the broad and narrow components of $H\alpha$ if appropriate.

4.5.2 ICA

The only sensible way to measure the systemic redshift is using the NIR ICA fit. $H\alpha$ and $H\beta$ seem to give no systematic offset but large scatter, and they are often asymmetric, so should only use peak. And $H\alpha$ isn't always available, and have other narrow components need to decompose. $O\text{ III}$ peak is mostly fine to use, but then there are some objects where the whole line is blueshifted. And critically, at large CIV blueshifts the $[O\text{ III}]$ emission is often undetected. Can show comparison of NIR ICA redshifts to... Optical ICA? Hewett & Wild?

Should emphasise that most people use $[O\text{ III}]$ to get the most reliable systemic redshift. While this is fine at low luminosities, at high luminosities this can result in large errors (profile can be dominated by blueshifted component, Fe emission can be improperly subtracted, or $[O\text{ III}]$ might not be detected at all. Publish ICA components with this paper?

Can also describe what I found trying to get redshifts from broad $H\alpha$, $H\beta$? (Narrow components generally very weak at these luminosities so can't be used.) Generally find no systematic errors but large

($\sim 1000 \text{ km s}^{-1}$ scatter). Comparing NIR ICA to [O III] for the [O III] with high S/N I find small (few hundred km s^{-1}) scatter.

Should publish [O III] redshifts with this paper for people to use.

4.5.3 *Parameter uncertainties and upper limits*

Describe how uncertainties on best-fitting parameters were calculated.

In 78 quasars, or approximately 25 per cent of or sample, the [O III] is undetected, or detected with very low S/N. In this section we describe how upper limits on the [O III] equivalent width were calculated. Firstly, the best-fitting model comprising the continuum, Fe II, and H β emission is subtracted from the spectra, leaving behind only emission due to [O III]. From this spectra we generate 100 mock spectra, where the flux at each wavelength is randomly drawn from a Normal distribution with a mean equal to the flux convolved with a Gaussian of width 200 km s^{-1} and a width equal to the known error. We then perform an error-weighted linear least-squares regression with an [O III] template derived from a fit to a very high S/N low redshift SDSS composite spectra. The equivalent width of the best-fitting model is recorded for each of the 100 realisations of the spectra. The error in the equivalent width is defined as the root-mean-square of these values.

Calculated uncertainties using Monte Carlo. Uncertainties on v_{10} are very large, which I think makes sense since the wing Gaussian will be appearing and disappearing, giving a large dispersion in v_{10} . Or regardless v_{10} is just very sensitive to the noise. Maybe I should be using v_{25} instead?

4.5.4 *Absolute flux calibration of spectra and continuum luminosities*

Relative flux-calibration of the infrared spectra as a function of wavelength has been achieved, to $\simeq 10$ per cent, through observations of appropriate flux standards. The absolute flux levels, however, can be in error by large factors due to variable atmospheric conditions combined with the narrow slit widths. For the majority of the quasars we have, therefore, established the absolute flux scale for each near-infrared spectrum using the same quasar SED-model fitting scheme employed in Paper I. Briefly, the SED-model was fit, with the normalisation and $E(B-V)$ as free variables, to optical/infrared magnitudes, or SDSS/BOSS spectra (check order I do this.) This allows us to extrapolate from the optical when we do not have photometric data in the near-infrared. The spectra were then normalised to the SED model using a linear error-weighted least-squares regression in the the regions of the spectra covered by the H/K bands. The monochromatic continuum luminosity at 5100 \AA was calculated directly from the normalised SED-model. **If this sounds strange can also calculate from fit to nor-**

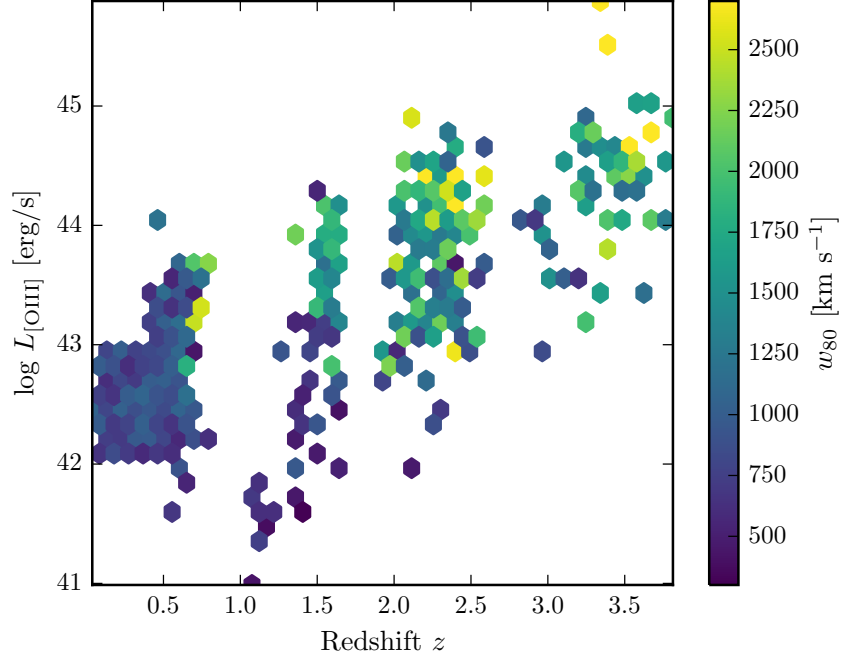


Figure 4.14: The [O III] velocity-width, characterised by w_{80} , as a function the [O III] luminosity and the quasar redshift. The color of each hexagon denotes the mean w_{80} for the objects in that luminosity-redshift bin. We have supplemented our sample with low- z objects from Zakamska and Greene, (2014) and medium ($z \sim 1.5$) redshift objects from Harrison et al., (2016). If I keep this plot make sure its clear which points belong to which sample.

maised spectra. Check if any missing normalisation / monochromatic luminosities.

4.6 RESULTS FROM ICA FITS

Need to convince the reader that the ICA components approximately correspond to real components. Explain how non-parameteric measures derived from ICA reconstructions.

We find there is a decreasing symmetric component at high luminosities. Relates directly to Shen and Ho, (2014). A stable narrow line region is removed by the outflowing material. Shen and Ho, (2014) showed that the strength of the core [O III] component decreases with quasar luminosity and optical Fe II strength faster than the wing component, leading to overall broader and more blueshifted profiles as luminosity and Fe II strength (or C IV blueshift) increases.

4.7 LUMINOSITY/REDSHIFT-EVOLUTION OF [O III] PROPERTIES

In this section we look for any luminosity/redshift dependent changes in the [O III] line properties. To do this we extend the dynamic range of our samples in terms of both luminosity and redshift by supplementing our sample with quasars presented by Zakamska and Greene, (2014) and Harrison et al., (2016).

The Zakamska and Greene, (2014) objects are a sample of 568 obscured luminous quasars selected from SDSS (Reyes et al., 2008; Yuan, Strauss, and Zakamska, 2016). They are selected to have [O III] luminosities above $10^{8.5} L_{\odot}$ and have a median redshift $z = 0.397$.

We also include 40 quasars at redshifts $1.1 \leq z \leq 1.7$ from the KMOS AGN Survey at High redshift (KASHz) with [O III] line measurements.

We also have the same information for $\sim 20\,000$ SDSS spectra from Mullaney et al., (2013).

In Figure 4.14 we show the [O III] velocity width as a function of the [O III] luminosity and the quasar redshift. The lack of any redshift-evolution between $z = 0$ and $z = 1.5$ was reported by Harrison et al., (2016). Our additional data suggests that this continues to $z \sim 2.5$. On the other hand, at fixed redshift, we see a significant correlation between the [O III] velocity width and the luminosity.

The fact that we don't see many broad lines in the Zakamska and Greene, (2014) objects even at luminosities $> 43 \text{ erg/s}$ could be due to the fact that these are all type II quasars, whereas the sample presented in this paper are all type I. Mullaney et al., (2013) showed that the [O III] lines of type I quasars are typically broader than in type II quasars.

Looking at the [O III] velocity width as a function of luminosity tells us about the physical drivers of the outflows observed in [O III]. The correlation with luminosity suggests that the highest velocity outflows are associated with the most luminous AGN. This has been reported for low-redshift AGN, for both ionized and molecular outflows (e.g. Westmoquette et al. 2012; Veilleux et al. 2013; Arribas et al. 2014; Ciccone et al. 2014; Hill & Zakamska 2014).

This suggests that the outflows are driven by radiative forces. On the other hand, Mullaney et al., (2013) find that once the correlation between the [O III] luminosity and the radio luminosity has been taken in to account, the [O III] velocity width is more strongly related to the radio luminosity of the AGN.

4.8 EQUIVALENT WIDTH

In Fig. 4.15 we show the [O III]5008 EW as a function of the quasar bolometric luminosity. Bolometric luminosity is estimated from the monochromatic continuum luminosity at 5100 \AA using the correction

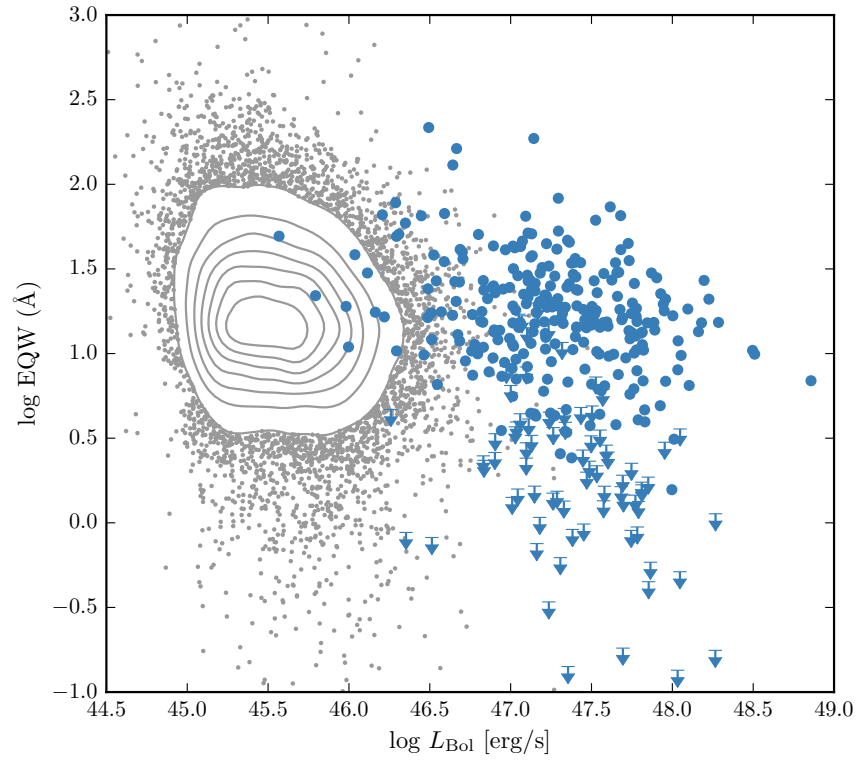


Figure 4.15: The [O III] EW as a function of the quasar bolometric luminosity for the sample presented in this paper (blue circles) and the low- z SDSS sample (grey points and contours). Upper limits are denoted by the downward arrows.

factor given by Richards et al., (2006a). For comparison, we also show the low- z sample from Shen et al., (2011).

The equivalent width of [O III] has been found to strongly decrease as a function of redshift and/or luminosity (e.g. Brotherton, 1996; Netzer et al., 2004; Sulentic et al., 2004; Baskin and Laor, 2005b).

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

Netzer et al., (2004) found 1/3 of their high luminosity sample had very weak [O III], whereas quasars with weak [O III] are very rare for nearby AGN. We find a very similar fraction. Netzer et al., (2004) claim that for the population of strong [O III] emitters there is no reduction of EW with increasing source luminosity. On the other hand, there are many weak or no [O III] emitters at high luminosity that could give the impression that the line EQW decreases with increasing source luminosity.

4.8.1 OIII outflows

Our best-fitting profiles show a strongly blue-asymmetric profile (Fig. 4.16), with a significant fraction of the total emission in a blue wing. The luminous blueshifted broad wing and the extremely broad profile reveals high-velocity outflowing ionized gas. This can be explained if the far-side of any outflowing gas, that is moving away from the line of sight, is obscured by dust in the host galaxies (e.g. Heckman et al. 1981; Vrtilek 1985). Observations at both low and high redshifts commonly observe this blueshifted component. Our results, and those of other authors, suggest that kpc-scale outflows in ionized gas are common among the most luminous high-redshift actively accreting SMBHs.

The situation is very different in nearby AGN, where the [O III] velocity width is dominated by the galactic potential and correlates well with the stellar velocity dispersion. H I, CO and absorption line measures of the host galaxy rest frame suggest that [O III] usually gives consistent results within 200 km/s (de Robertis 1985; Whittle 1985; Wilson & Heckman 1985; Condon et al. 1985; Stripe 1990; Alloin et al. 1992; Evans et al. 2001).

We see a correlation between the [O III] velocity width and blueshift. As the blueshift of the line increases it gets broader. This is consistent with Shen and Ho, (2014), where the strength of the narrow core is decreasing, leading to a broader and more blueshifted profile.

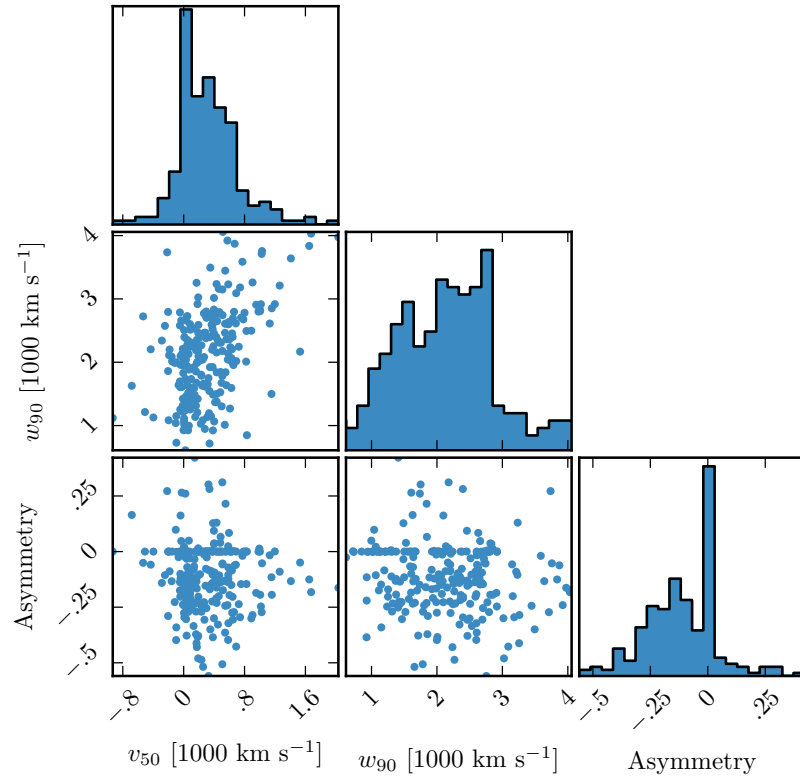


Figure 4.16: The distributions of and correlations between a subset of the non-parameteric measures we made of the best-fitting [O III] models.

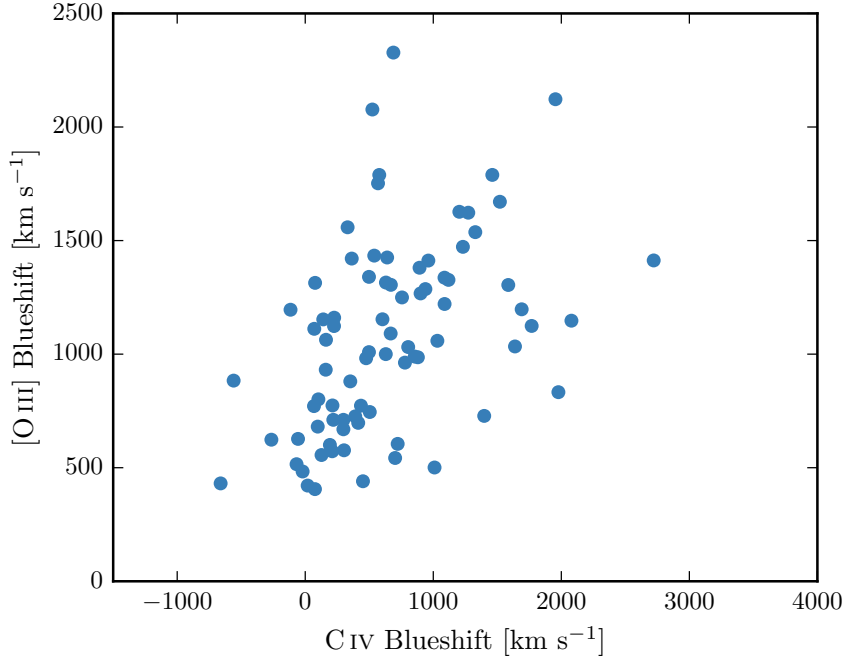


Figure 4.17: The relation between the blueshifts of C IV and [O III]. Equivalent to Fig 8. We use the H β peak in this figure, which I think is responsible for some of the trend. However, we do see a correlation (albeit noisier) using the NIR ICA redshifts. Not a sensible to use the [O III] redshifts, since these become much more unreliable at the high C IV blueshift end (when [O III] is weaker: figure 7. Note that we are using v_{10} for the [O III] position and v_{50} for the C IV position. We can't use v_{50} for [O III] because sometimes we are using a single Gaussian, especially if the [O III] is weaker and we miss the broad component. Need to remake this plot / don't use at all because I don't believe some of the Gaussian fits to [O III], especially at high C IV blueshifts when [O III] is weak and Fe II is strong.) Only objects where fit with two components.

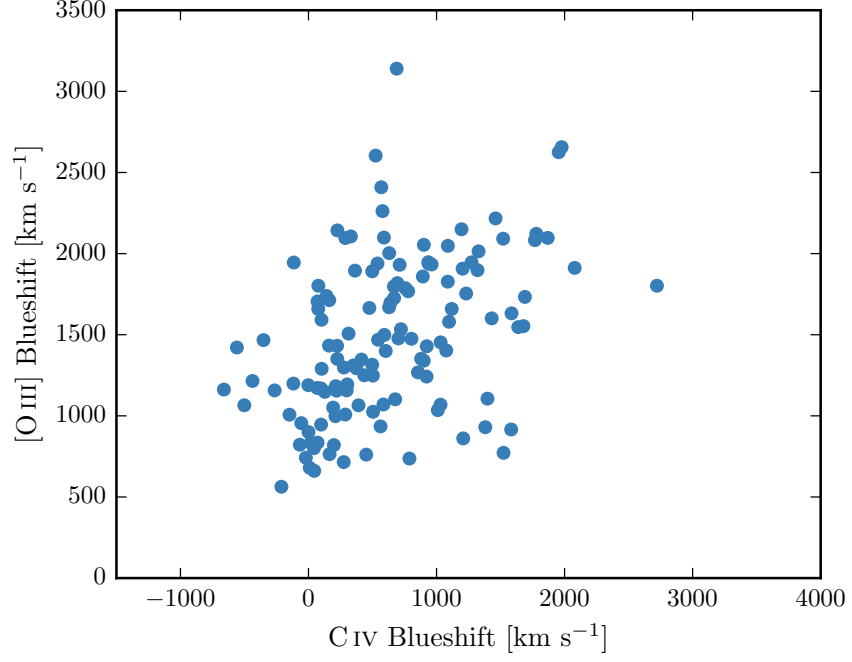


Figure 4.18

4.8.2 $[O\text{ III}]$ and $C\text{ IV}$ outflows are linked

As described in Paper I, we have searched for optical counterparts to our near-infrared spectra. Optical spectra are available for XXX quasars in our catalogue, and cover the broad $C\text{ IV}$ doublet. As we described in Paper I and Coatman et al., (2016), $C\text{ IV}$ is often blueshifted, which almost certainly signal the presence of strong outflows, most likely originating in a disc wind. In Paper I we demonstrated that the quasars in our sample cover the full range of $C\text{ IV}$ blueshifts seen in the SDSS quasar population, which makes our sample unique in that it allows us to study properties of the quasar across the full parameter range.

The $C\text{ IV}$ velocity centroid measurements are taken directly from paper I. We define the ‘location’ of the $[O\text{ III}]$ emission using v_{10} , although the results are the same if v_{20} , v_{50} etc. are used instead.

In Figure 4.17 we show the $C\text{ IV}$ blueshifts against the $[O\text{ III}]$ blueshifts. This comparison is done for a sub-sample of 146 objects where we have good measurements of the $C\text{ IV}$, $[O\text{ III}]$, and $H\beta$ (to measure the systemic redshift) profiles. Objects with $S/N > 3$ are shown as blue filled circles and objects with $S/N < 3$ as open grey circles. We calculated the median S/N per pixel in the best-fitting model for the $[O\text{ III}]5008$ emission.

There is a clear and strong correlation. Similar correlations have been tentatively found in lower redshift quasars and AGN (Zamanov et al., 2002).

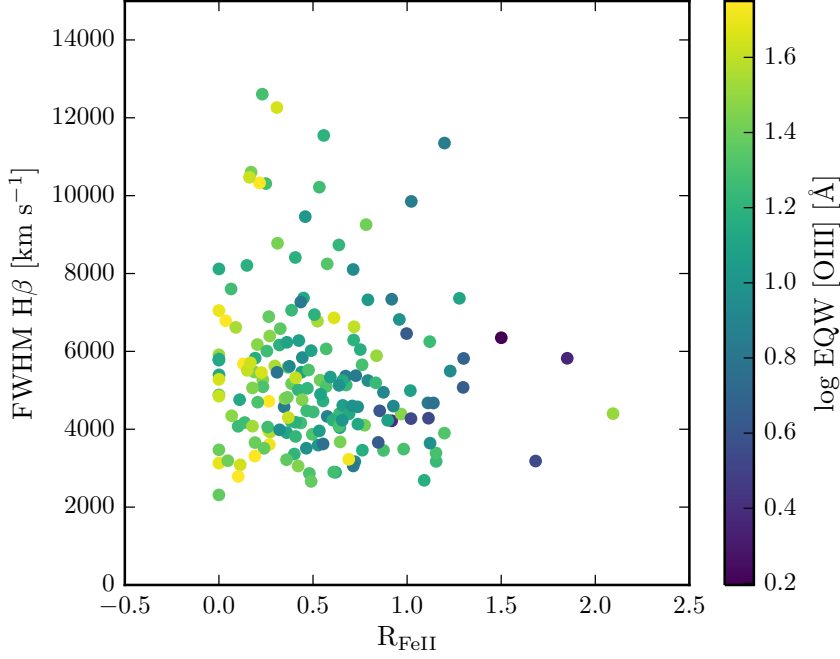


Figure 4.19: The [O III] EQW as a function of the H β FWHM and the optical Fe II strength (EQW Fe II / EQW H β).

The blueshifting of C IV is known to correlate with luminosity (Richards et al., 2011). In [O III], the blueshifted wing becomes relatively more prominent as the luminosity of the quasar increases (Shen and Ho, 2014). Therefore, it is plausible that the correlation between the C IV and [O III] blueshifts is a secondary effect that is driven by the correlation of each with the luminosity. In Figure ?? we color the data points by the luminosity, and no luminosity-dependent trends are apparent. We find that both the [O III] and C IV blueshifts are correlated with the luminosity, but that these correlations are much weaker than the correlation between the [O III] and C IV blueshifts.

4.9 EIGENVECTOR ONE CORRELATIONS

In Figure 4.19 we show the [O III] EQW as a function of the H β FWHM and the optical Fe II strength. The optical Fe II strength is defined as the ratio of the Fe II and H β EQW, where the Fe II EQW is measured between 4434 and 4684Å. These parameters form part of ‘eigenvector 1’ (EV₁), the first eigenvector in a principal component analysis which originated from the work of Boroson and Green, (1992). In our sample, these parameters follow very similar correlations to what is observed at low- z (e.g. Shen and Ho, 2014). In particular, the anti-correlation between the [O III] and Fe II EQWs.

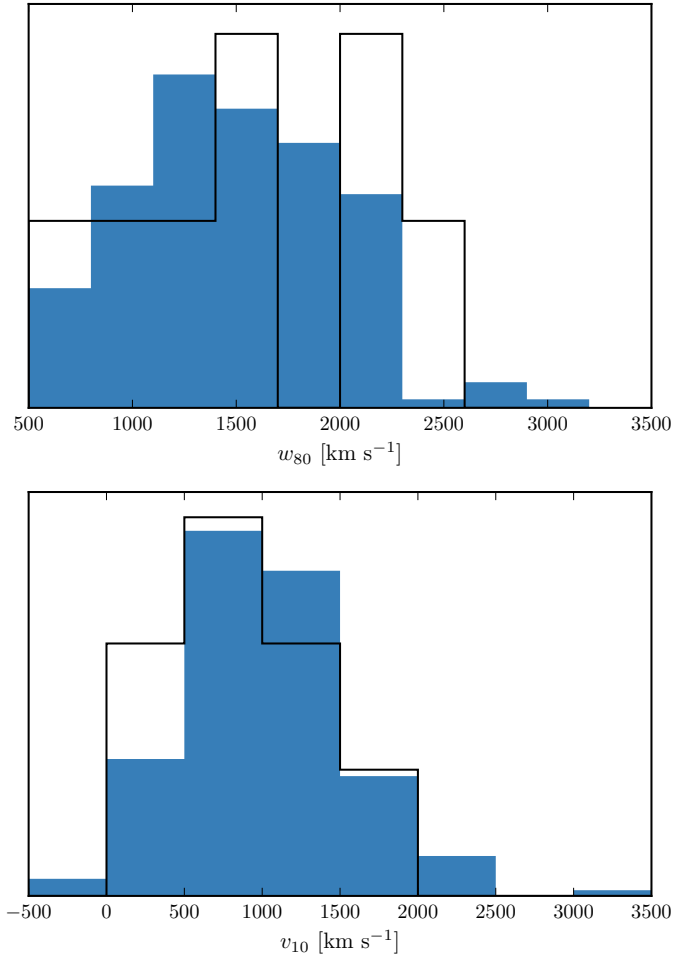


Figure 4.20: The distribution of w_{80} and v_{10} for the 19 BALs are compared to the distribution for the non-BALs. These look rubbish. Cumulative distributions instead? Try doing gaussian kernel density estimator

Same as Shen, (2016), we confirm that the EV1 correlations hold at high luminosities/redshifts. See also Sulentic et al. 2004, 2006; Runnoe et al. 2013. Make sure it's clear that Shen, (2016) quasars make up a significant chunk of our sample.

4.10 MAPPING EV1 TO CIV BLUESHIFT AND EQW

4.11 SIGNAL TO NOISE TESTS

4.12 BROAD ABSORPTION LINE QUASARS

19 quasars in our catalogue are classified as broad absorption line (BAL) quasars, using either the SDSS classification flags or the Allen et al., (2011) catalogue. We find that the BAL quasars have typ-

ically broader [O III] than the rest of the sample. Note that in the Zakamska et al., (2016) sample of very red quasars, the incidence of BALs is very high, and these objects have extremely broad [O III] profiles. A two-sided Kolmogorov-Smirnov statistic on the w_{80} distributions returned a p-value of 0.10. What does this mean? Try with different parameters? Histograms look rubbish so maybe just give the numbers.

4.13 DISCUSSION

4.13.1 *Type II quasars*

Implications of our findings on searches for high-redshift type 2 quasars. It could be that type II quasars exist. If you look at CIV/MgII the narrow line components are very weak. So the contribution from the narrow line region is very weak in luminous quasars, and you just won't see it even if the broad line region is obscured. Findings in this paper seem to suggest that the startic narrow line region is very weak in luminous quasars.

SED PROPERTIES

While many authors have focused on studies of specific sub-sets of active galactic nuclei (AGN) with extreme observational properties, what is missing is an understanding of how these extreme subsets relate to the population as a whole. I have addressed this problem using multi-wavelength spectral energy distributions (SEDs) of large samples of quasars. I have constructed an SED model which is able to reproduce the average optical to near-infrared (NIR) colours of 10,000s of AGNs spanning a broad range in redshift and luminosity.

5.1 DATA

The systematic study of the dependence of the SED shape on physical parameters has, until very recently, been limited by the difficulty in obtaining a large sample of quasars with good multi-wavelength coverage and large dynamic range in luminosity and redshift. In this work, we take advantage of a number of recent, sensitive, wide-field surveys, covering the UV to mid-IR spectral region.

5.1.1 *The Sloan Digital Sky Survey*

We use the Seventh Data Release (DR7) of the Sloan Digital Sky Survey (SDSS; York et al., 2000) spectroscopic quasar catalogue (Schneider et al., 2010), which includes 105,783 objects across 9380 deg². The SDSS obtained images in five broad optical passbands: u ($\lambda_{\text{eff}} = 3543\text{\AA}$), g ($\lambda_{\text{eff}} = 4770\text{\AA}$), r ($\lambda_{\text{eff}} = 6231\text{\AA}$), i ($\lambda_{\text{eff}} = 7625\text{\AA}$), and z ($\lambda_{\text{eff}} = 9134\text{\AA}$). We use BEST point-spread function (PSF) magnitudes, correcting for Galactic extinction using the maps of Schlegel, Finkbeiner, and Davis, (1998), assuming a Milky Way (MW) extinction curve (Pei, 1992) and an extinction to reddening ratio $A(V)/E(B - V) = 3.1$. Although the SDSS asinh magnitude system is intended to be on the AB system (Oke and Gunn, 1983), the photometric zero-points are known to be slightly off the AB standard. To account for this we add 0.03 mag to the u, g, r and i magnitudes, and 0.05 mag to the z magnitude.

DR7Q quasar targets were primarily selected to have $i \leq 19.1$ if the colours were consistent with being at redshift $z < 3$, and $i \leq 20.2$ if consistent with $z > 3$ (Richards et al., 2002). The survey is sensitive to the most luminous quasars at a given redshift. The large number of objects at $z < 3$ with $i > 19.1$ were selected by algorithms other than the main quasar selection. For example, quasar targets were also

Where did these numbers come from?

selected if they matched within $2''$ of an object in the Faint Images of the Radio Sky at Twenty-cm (FIRST) catalogue of radio sources (Becker, White, and Helfand, 1995).

5.1.2 *UKIDSS Large Area Survey*

We use the UKIRT Infrared Deep Sky Survey (UKIDSS; Lawrence et al., 2007) Large Area Survey (ULAS) which has observed $\sim 3,200 \text{ deg}^2$ in four near-IR passbands: Y ($\lambda_{\text{eff}} = 1.0305 \mu\text{m}$), J ($\lambda_{\text{eff}} = 1.2483 \mu\text{m}$), H ($\lambda_{\text{eff}} = 1.6313 \mu\text{m}$), and K ($\lambda_{\text{eff}} = 2.2010 \mu\text{m}$). We used the ninth data release (DR9) of the ULAS. Cross-matching (with a $2''$ radius and picking only the nearest neighbour) the SDSS DR7Q catalogue with the ULAS catalogue, which covers only $\sim 38\%$ of the SDSS footprint, resulted in 37,893 matches. The ULAS magnitudes are aperture corrected magnitudes in a $2''$ diameter aperture and are not corrected for Galactic extinction.

5.1.3 *WISE All-WISE Survey*

The Wide-field Infrared Explorer (WISE; Wright et al., 2010) mapped almost the sky in four mid-IR band-passes: W1 ($\lambda_{\text{eff}} = 3.4 \mu\text{m}$), W2 ($\lambda_{\text{eff}} = 4.6 \mu\text{m}$), W3 ($\lambda_{\text{eff}} = 12 \mu\text{m}$), and W4 ($\lambda_{\text{eff}} = 22 \mu\text{m}$). The WISE AllWISE Data Release ('AllWISE') combines data from the nine month cryogenic phase of the mission that led to the 'AllSky' data release with data from the NEOWISE program (Mainzer et al., 2011). Cross-referencing the SDSS DR7Q catalogue with the AllWISE catalogue resulted in 102,734 matches. Two objects were matched to multiple AllWISE objects, and were discarded from the sample. Vega to AB conversion factors for WISE photometry are given in the WISE Explanatory Supplement (Cutri et al., 2013)

5.1.4 *Completeness of Photometry*

Objects which are faint in the SDSS *i* band-passes are more likely to have magnitudes which fall below the limiting magnitudes of the UKIDSS and WISE band-passes at longer wavelengths. For a given *i* magnitude, a quasar with a blue spectrum is more likely to be undetected at longer wavelengths than a quasar with a red spectrum. Therefore, as we allow fainter quasars in to our sample we will be biased towards objects with redder spectra. We impose an observed *i* magnitude lower limit of 19.1 mag, which is the magnitude limit of the main SDSS colour-selection algorithm. We verified that above this limit the DR7Q-matched sample is 95% complete in all band-passes with $S/N > 5$ (excluding WISE W3 and W4) and that this fraction is not changing rapidly with the brightness of the sample.

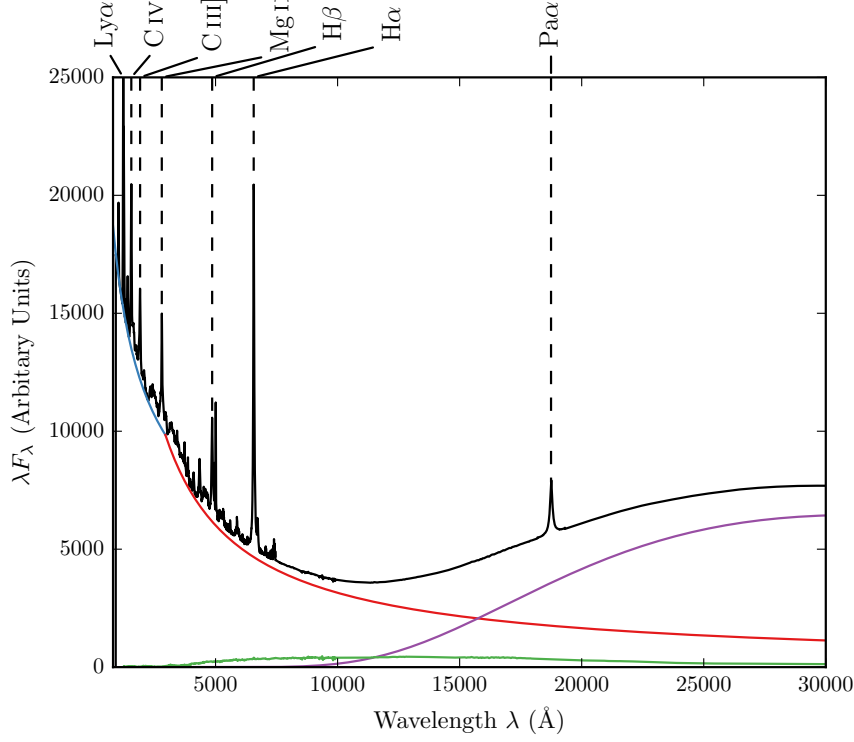


Figure 5.1: Model spectrum at $z = 1$, showing the contributions to the total flux from the blue power-law slope, red power-law slope, black-body and host galaxy. The locations of the most prominent emission lines in the spectrum are also indicated.

5.1.5 Final Sample

We exclude objects flagged as BALQSOs by Shen et al., (2011), since our model is unable to reproduce the broad absorption troughs that appear in the spectra of these objects. The final sample contains 61,411 objects in the redshift range $0.2 < z < 3.8$.

5.2 SED MODEL

I have constructed a new SED model which reproduces the SEDs of AGNs from the rest-frame UV ($\sim 0.1\mu\text{m}$) to the rest-frame near-IR ($\sim 3\mu\text{m}$). In this section, I will describe how I have modelled the emission from the various components contributing to the emission in this spectral region. The model spectrum is shown in Figure 5.1, with each of the main components indicated.

5.2.1 Accretion Disk

More than half the bolometric luminosity of an unobscured AGN is emitted in the Big Blue Bump, which extends from the near-IR at $1\ \mu\text{m}$ to past $0.1\ \mu\text{m}$ in the UV, and possibly all the way to the soft X-ray region. The Big Blue Bump emission is thought to arise from an accretion disc. In the $0.1 - 1\ \mu\text{m}$ region the spectrum is generally characterised by a power-law of the form $F_\nu = C\nu^{-\alpha}$ where α is the power-law index, C is a constant, and F_ν is the flux per unit frequency, usually measured in units of $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$. Equivalently this can be expressed as $F_\lambda = C'\lambda^{\alpha-2}$ where F_λ is the flux per unit wavelength, usually measured in units of $\text{s}^{-1} \text{cm}^{-2} \text{\AA}^{-1}$.

The value of the power-law index is uncertain. From a theoretical perspective, models of geometrically thin accretion discs (Shakura and Sunyaev, 1973) assume, in particular, that the disc is stationary, axisymmetric, and extends down to the innermost stable circular orbit, and that angular momentum is transported by local ‘viscous’ stresses that convert gravitational energy entirely into heat. This gives the dependence of the effective temperature on radius as $t_{\text{eff}} \propto r^{-3/4}$. A spectrum is then calculated by dividing the disc into concentric annuli, calculating the spectrum emitted by each annulus and then summing them all together. Assuming that each annulus radiates like a blackbody, the $r^{-3/4}$ effective temperature distribution gives $F_\nu \propto \nu^{1/3}$ (Peterson, 1997), although it is unclear whether this is consistent with observations.

In our model we characterised the Big Blue Bump from $\sim 0.1 - 1\ \mu\text{m}$ as a broken power-law with three free parameters: a break-wavelength λ_{break} , a blue power-law index α_{blue} for wavelengths shorter than the break wavelength, and a red power-law index α_{red} for wavelengths longer than the break wavelength.

5.2.2 Hot Dust

At wavelengths longer than $1\ \mu\text{m}$, emission from hot dust begins to dominate over emission from the accretion disc. The SED in this region is generally characterised either by a power-law ($\propto \lambda^{\beta_{\text{NIR}}}$), with $\beta \simeq 0.5$ (e.g. Richards et al., 2006a; Zhang et al., 2014), or by a blackbody at $\sim 1300\ \text{K}$, thus peaking in the near-IR (e.g. Leipski et al., 2014). We modelled the hot dust emission using a simple blackbody:

$$F_\lambda = \frac{2hc^2}{\lambda^5} \frac{1}{e^{\frac{hc}{\lambda k_B T}} - 1} \quad (5.1)$$

The blackbody component has two free parameters: the temperature of the blackbody T_{BB} and the overall normalisation.

5.2.3 Emission Lines

Hundreds of emission lines are present in a typical AGN spectra. Some of the most prominent lines are shown in Figure 5.1. The emission line spectrum is taken from Maddox and Hewett, (2006), who extend the composite of Francis et al., (1991) to include the $H\alpha$ (6560Å) and $Pa\alpha$ (18750Å) emission lines. A single parameter, EL_{scale} , scales the equivalent widths of all emission lines equally:

$$F_{\lambda} = EL_{\text{scale}} \times \frac{F_{\lambda,\text{el}}}{F_{\lambda,\text{cont}}} \times F_{\lambda} \quad (5.2)$$

where $F_{\lambda,\text{el}}$ is the line flux in the template, $F_{\lambda,\text{cont}}$ is the continuum flux in the template, and F_{λ} is the continuum flux in the model.

5.2.4 Host Galaxy

Emission from the host galaxy is important, particularly in the region around the $1\mu\text{m}$ inflection point in the quasar SED. While the host galaxies of bright quasars tend to be massive, bright ellipticals, the hosts of lower luminosity AGN can have disc components (e.g. Dunlop et al., 2003). Our model incorporates $z = 0$ Sa, Sb, Sc and elliptical-type templates from Mannucci et al., (2001), which for simplicity do not evolve with redshift. We characterise the relationship between the luminosity of the AGN L_{AGN} and the luminosity of the host galaxy L_{Gal} as a power-law

$$L_{\text{Gal}} = L_{\text{AGN}}^{\beta} \quad (5.3)$$

with power-law index $\beta = 0.42$ (Maddox and Hewett, 2006). Dividing both sides of Equations 5.3 by the luminosity of the AGN gives the luminosity of the host galaxy relative to the luminosity of the AGN

$$\frac{L_{\text{Gal}}}{L_{\text{AGN}}} = L_{\text{AGN}}^{\beta-1} \quad (5.4)$$

which for $\beta < 1$ decreases with increasing AGN luminosity. In a flux limited sample, the AGN luminosity will tend to increase with redshift and so the luminosity of the host galaxy relative to the luminosity of the quasar will decrease with increasing redshift. Hence, the contribution from the host galaxy to the total flux is important at low redshift, but becomes gradually less significant towards higher redshifts.

Since the contribution from the host galaxy to the flux changes as a function of AGN luminosity, and hence redshift, we choose a reference redshift z_{nrn} where we set the fractional contribution of the

host galaxy to the total flux, η . In an arbitrary region of the spectrum (we use $4000 - 5000 \text{ \AA}$) we calculate both the AGN continuum flux $F_{\text{AGN}}(z_{\text{nrn}})$ and the flux from our host galaxy template spectrum $F_{\text{Gal}}(z_{\text{nrn}})$. The fractional contribution from the host galaxy to the total flux is then:

$$\eta = \frac{CF_{\text{Gal}}(z_{\text{nrn}})}{F_{\text{AGN}}(z_{\text{nrn}}) + CF_{\text{Gal}}(z_{\text{nrn}})} \quad (5.5)$$

where the constant C is the factor by which we must multiply the unnormalised galaxy spectrum in order for Equation 5.5 to hold true. Rearranging for the constant C we find

$$C = \frac{\eta}{1 - \eta} \frac{F_{\text{AGN}}(z_{\text{nrn}})}{F_{\text{Gal}}(z_{\text{nrn}})} \quad (5.6)$$

Hence at redshift z_{nrn} the host galaxy flux we add to our rest frame quasar continuum is

$$F_{\lambda} = \frac{\eta}{1 - \eta} \frac{F_{\text{AGN}}(z_{\text{nrn}})}{F_{\text{Gal}}(z_{\text{nrn}})} F_{\lambda, \text{Gal}} \quad (5.7)$$

where $F_{\lambda, \text{Gal}}$ is our host galaxy template spectrum in the quasar rest frame. The contribution from the host galaxy at a different redshift z is given by

$$F_{\lambda} = \frac{\eta}{1 - \eta} \frac{F_{\text{AGN}}(z)}{F_{\text{Gal}}(z)} \frac{F_{\text{AGN}}(z_{\text{nrn}})}{F_{\text{Gal}}(z_{\text{nrn}})} \left(\frac{F_{\text{AGN}}(z)}{F_{\text{Gal}}(z)} \right)^{-1} F_{\lambda, \text{Gal}} \quad (5.8)$$

$$= \frac{\eta}{1 - \eta} \frac{F_{\text{AGN}}(z)}{F_{\text{Gal}}(z)} \frac{F_{\text{AGN}}(z_{\text{nrn}})}{F_{\text{Gal}}(z_{\text{nrn}})} \frac{F_{\text{Gal}}(z)}{F_{\text{AGN}}(z)} F_{\lambda, \text{Gal}} \quad (5.9)$$

$$= \frac{\eta}{1 - \eta} \frac{F_{\text{AGN}}(z)}{F_{\text{Gal}}(z)} \frac{L_{\text{AGN}}(z_{\text{nrn}})}{L_{\text{Gal}}(z_{\text{nrn}})} \frac{L_{\text{Gal}}(z)}{L_{\text{AGN}}(z)} F_{\lambda, \text{Gal}} \quad (5.10)$$

$$= \frac{\eta}{1 - \eta} \frac{F_{\text{AGN}}(z)}{F_{\text{Gal}}(z)} \frac{L_{\text{AGN}}(z_{\text{nrn}})}{L_{\text{AGN}}(z_{\text{nrn}})^{\beta}} \frac{L_{\text{AGN}}(z)^{\beta}}{L_{\text{AGN}}(z)} F_{\lambda, \text{Gal}} \quad (5.11)$$

$$= \frac{\eta}{1 - \eta} \frac{F_{\text{AGN}}(z)}{F_{\text{Gal}}(z)} \left(\frac{L_{\text{AGN}}(z)}{L_{\text{AGN}}(z_{\text{nrn}})} \right)^{\beta-1} F_{\lambda, \text{Gal}} \quad (5.12)$$

We need to know how the luminosity of the AGN depends on redshift. This is given by:

$$\frac{L_{\text{AGN}}(z)}{L_{\text{AGN}}(z_{\text{nrn}})} = 10^{-0.4(M_{\text{AGN}}(z) - M_{\text{AGN}}(z_{\text{nrn}}))} \quad (5.13)$$

where $M_{\text{AGN}}(z)$, the absolute magnitude of an AGN at redshift z , is given by

$$M(z) = m - 5(\log_{10} D_L(z) - 1) \quad (5.14)$$

and $D_L(z)$ is the luminosity distance to a source at redshift z in parsecs. Hence:

$$\frac{L_{\text{AGN}}(z)}{L_{\text{AGN}}(z_{\text{nm}})} = 10^{-0.4(M_{\text{AGN}}(z) - M_{\text{AGN}}(z_{\text{nm}}))} \quad (5.15)$$

$$= 10^{(\log_{10}(\frac{D_L(z)}{D_L(z_{\text{nm}})})^2)} \quad (5.16)$$

5.2.5 Lyman- α Forest Absorption

The optical spectra of high redshift quasars show hundreds of sharp absorption lines, which mostly correspond to the redshifted neutral hydrogen Ly α 1216Å transition. These absorption features are collectively referred to as the *Lyman- α forest*. To simulate the effect of Lyman- α forest absorption on our model SED we use the parametrisation of Becker et al., (2013), who derived an analytic function for the effective optical depth τ_{eff} over the redshift range $2 < z < 5$ made using 6065 quasar spectra from SDSS DR7. In their model the effective optical depth τ_{eff} is given by

As Paul how he implements this. My code for this can easily be sped up.

$$\tau_{\text{eff}} = \tau_0 \times \left(\left(\frac{1+z}{1+z_0} \right)^b + C \right) \quad (5.17)$$

where,

$$\begin{aligned} t_0 &= 0.751 \\ b &= 2.9 \\ C &= -0.132 \\ z_0 &= 3.5 \end{aligned}$$

The transmitted flux $F_{\lambda, \text{trans}}$ at redshift z is then given by

$$f_{\lambda, \text{trans}} = F_{\lambda} \times e^{-\tau_{\text{eff}}} \quad (5.18)$$

An absorption line at λ_{abs} in the rest-frame of an AGN at redshift z_{AGN} has wavelength

$$(1 + z_{\text{AGN}})\lambda_{\text{abs}} \quad (5.19)$$

in the rest frame of an observer on Earth. In the rest-frame of a cloud of neutral hydrogen at redshift z_{cloud} the absorption line has wavelength

$$\frac{(1 + z_{\text{AGN}})\lambda_{\text{abs}}}{(1 + z_{\text{cloud}})} \quad (5.20)$$

and so to absorb Lyman- α at 1216 \AA the gas cloud must be at a redshift

$$z_{\text{cloud}} = \frac{(1 + z_{\text{AGN}})\lambda_{\text{abs}}}{1216 \text{ \AA}} - 1 \quad (5.21)$$

For every wavelength $\lambda_{\text{abs}} < 1216 \text{ \AA}$ in the rest-frame of an AGN at redshift $z > 2$ we calculate z_{cloud} using Equation 5.21 and then calculate the transmitted flux at λ_{abs} by substituting z_{cloud} in to Equations 5.17 and 5.18.

5.2.6 Lyman-Limit Systems

Ask Paul if this
should be
implemented

Lyman-limit systems are clouds of HI which are optically thick at the Lyman limit (912 \AA), which generally implies a neutral hydrogen column density $N(\text{HI}) > 10^{17} \text{ cm}^{-2}$. Photons at wavelengths shorter than the Lyman-limit will be absorbed, which creates a sharp break in the observed continuum. We model the effect of a Lyman-limit system at the redshift of the quasar by setting the flux at wavelengths less than 912 \AA in the quasar rest frame to zero.

5.2.7 Dust Extinction

The selection criteria of the SDSS DR7Q catalogue, and particularly the DR10Q catalogue, are sensitive to quasars with moderate amounts of dust reddening (possibly as high as $E(B-V) \sim 0.5$; Richards et al., 2003) at the redshift of the quasar, and so we included the effect of dust extinction in our model. We considered four types of extinction curve: the Large Magellanic Cloud (LMC), Small Magellanic Cloud (SMC), Milky-Way (MW) extinction curves from Pei, (1992) and an extinction curve appropriate for the quasar population which has been derived by Paul Hewett. To derive the quasar extinction curve, UKIDSS photometry was used to provide an $E(B-V)$ estimate, via the magnitude displacement of each quasar from the locus of unreddened objects. At redshifts $2 < z < 3$ the reddening measure is made at rest-frame wavelengths $3500\text{--}7000 \text{ \AA}$, where Galaxy, LMC and SMC extinction curves are very similar. The SDSS spectra of the same objects are then employed to generate an empirical extinction curve in the ultraviolet, down to 1200 \AA . The resulting curve has no 2200 \AA feature and rises rapidly with decreasing wavelength but is not as steep as the SMC curve. The extinction curves give the colour excess $E(B - \lambda)$ relative to the colour excess $E(B - V)$ as a function of

wavelength λ . The colour excess $E(B - V)$ is related to the extinction in the V band, $A(V)$, via a parameter R ,

$$A(V) = R \times E(B - V) \quad (5.22)$$

where $R = 3.1$ in the MW and $R \simeq 3$ in the Magellanic Clouds. Hence the extinction at a wavelength λ $A(\lambda)$ is

$$A(\lambda) = E(B - V) \times \left[\frac{E(\lambda - V)}{E(B - V)} + R \right] \quad (5.23)$$

where the colour excess $E(B - V)$ is a free parameter in our model. The attenuation of the flux at a given wavelength is then:

$$F_\lambda = F_\lambda 10^{-A(\lambda)/2.5} \quad (5.24)$$

in the rest frame of the quasar.

5.3 THE 'STANDARD' SED MODEL

We will begin by deriving a 'standard' SED model by constraining a single set of parameters with a large sample of $0.2 < z < 4$ quasars encompassing a range of luminosities, accretion rates etc. The free parameters in our model are the blue power-law slope, the red power-law slope, the power-law break wavelength, the blackbody temperature, the blackbody normalisation, the emission line equivalent width scaling, and the fractional contribution from the host galaxy to the total flux. The reddening $E(B-V)$ is fixed to zero, since a large fraction of SDSS quasars have very small amounts of dust reddening (Richards et al., 2003). For the host galaxy we use a Sb-type template derived by Mannucci et al., (2001). With some choice of initial parameters, we generate a set of model observed spectra at redshifts from $z = 0.25$ to $z = 3.75$ in intervals of $\Delta z = 0.1$. We then transform our set of model spectra into a set of model ugrizYJHKW1W2 SEDs

The throughput functions of the SDSS ugriz, UKIDSS YJHK and WISE W1W2W3 band-passes are shown in Figure 5.2, along with our model AGN spectra at three different redshifts. The mean flux density in a band-pass P is given by

$$f_\lambda(P) = \frac{\int P(\lambda) f_\lambda(\lambda) \lambda d\lambda}{\int P(\lambda) \lambda d\lambda} \quad (5.25)$$

where $P(\lambda)$ is the dimensionless throughput function of the band-pass. The corresponding magnitude, $m_\lambda(P)$, is then

$$m_\lambda(P) = -2.5 \log(f_\lambda(P)) - m_0(P) \quad (5.26)$$

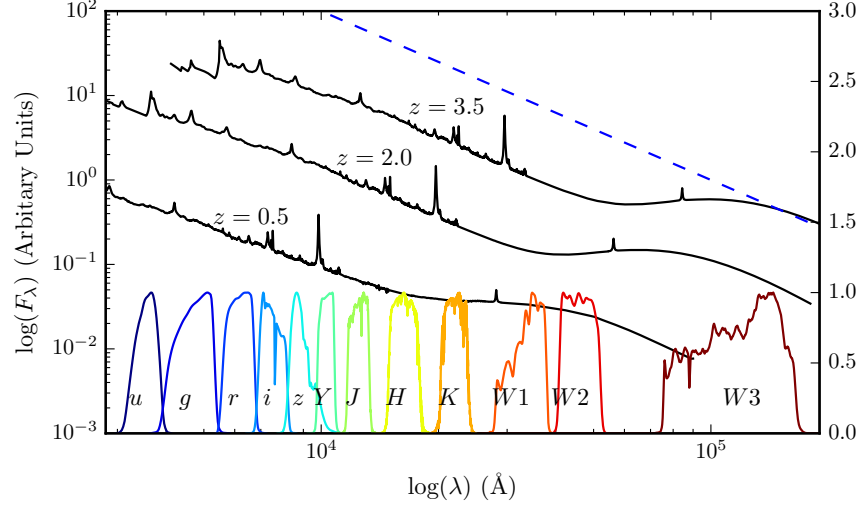


Figure 5.2: Model spectrum at three different redshifts (each arbitrarily scaled), and throughput functions for SDSS, UKIDSS and WISE band-passes (scaled so that the peak transmission is equal to one.) The dashed line indicates the slope of the AB magnitude system zero point.

where $m_0(P)$ is the zero-point magnitude of band P . In the AB magnitude system, the zero-point flux per unit wavelength is

$$\frac{f_\lambda(\lambda)}{\text{erg cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1}} = 0.1087 \left(\frac{\lambda}{\text{\AA}} \right)^{-2}. \quad (5.27)$$

This is substituted into Equation 5.25 to give a zero-point mean flux density which is then converted into a corresponding magnitude.

The model SEDs are normalised such that the i magnitude of each model SED is 18.0 mag. This gives us an array of model magnitudes as a function of redshift and band-pass. We generate an equivalent data array by dividing our quasar sample into redshift bins from $z = 0.2$ to $z = 3.8$ with bin width $\Delta z = 0.1$. We normalise the individual quasar SEDs such that the observed i magnitude is equal to 18.0 mag, and then calculate a median SED in each redshift bin.

To fit the model to the data we minimise the sum of the squares of the differences between the elements in the model magnitude array and the elements in the data magnitude array. The minimisation is done using the ‘nelder-mead’ algorithm. Our SED model is valid only up to $\lambda \sim 3\mu\text{m}$ in the quasar rest frame (the approximate wavelength of the peak in hot dust emission); beyond this additional contributions to the total flux from cooler dust will become significant. This prevents us from using the two highest wavelength WISE bands in the fit. We also exclude the SDSS u and g band-passes from the fit at $z > 2.7$ and $z > 3.7$ respectively, where absorption in the Lyman α forest becomes large.

Parameter	Symbol	Before Correction	After Correction
Blue power-law index	α_{blue}	0.58	0.58
Red power-law index	α_{red}	-0.04	-0.05
Power-law break	λ_{break}	2945	2957
Blackbody temperature	T_{BB}	1216 K	1186 K
Blackbody normalisation	C_{BB}	0.22	0.21
Emission line scaling	C_{EL}	0.63	0.73
Galaxy fraction	η	0.29	0.28
E(B-V)	E(B-V)	0.00	0.00

Table 5.1: Best-fitting parameters from fit to DR7Q-matched sample. **Only give best-fit values after correction.**

The best-fitting parameters from the fit are shown in Table 5.1. The colours ($u - g$, $g - r$, etc.) of the median SED, the individual quasars, and the best-fitting model are plotted as a function of redshift in Figs. 5.3 and 5.4. Most of the large variations that can be seen in the median colours of the quasars as a function of redshift are due to strong emission lines being redshifted in to and out of the bandpasses of the band-passes being used.

Re-do fit

5.4 DISCUSSION OF FIT

In Figure 5.5 we show the difference between the magnitudes from the best-fitting model and the median magnitudes from the sample. We have transformed the effective wavelengths of the band-passes to the rest frame of the quasars in each redshift bin, to give to the residuals as a function of rest-frame wavelength. We represent the residuals measured in each band-pass using a different coloured line. Differences between residuals from different band-passes at the same rest-frame wavelength could indicate redshift evolution of the typical quasar SED.

The residuals indicate that over a large redshift range the model does a fairly good at reproducing the median observed colours of the DR7Q-matched sample. Most discrepancies are at the < 0.1 mag level. It is remarkable that a single model is so effective; the properties of a typical quasar to not change significantly over a wide range of redshifts and luminosities. On the other hand, for the individual objects there is a significant scatter about the mean. In general, our goal is to use this intrinsic spread in SED properties in order to understand the diversity in physical quasar properties.

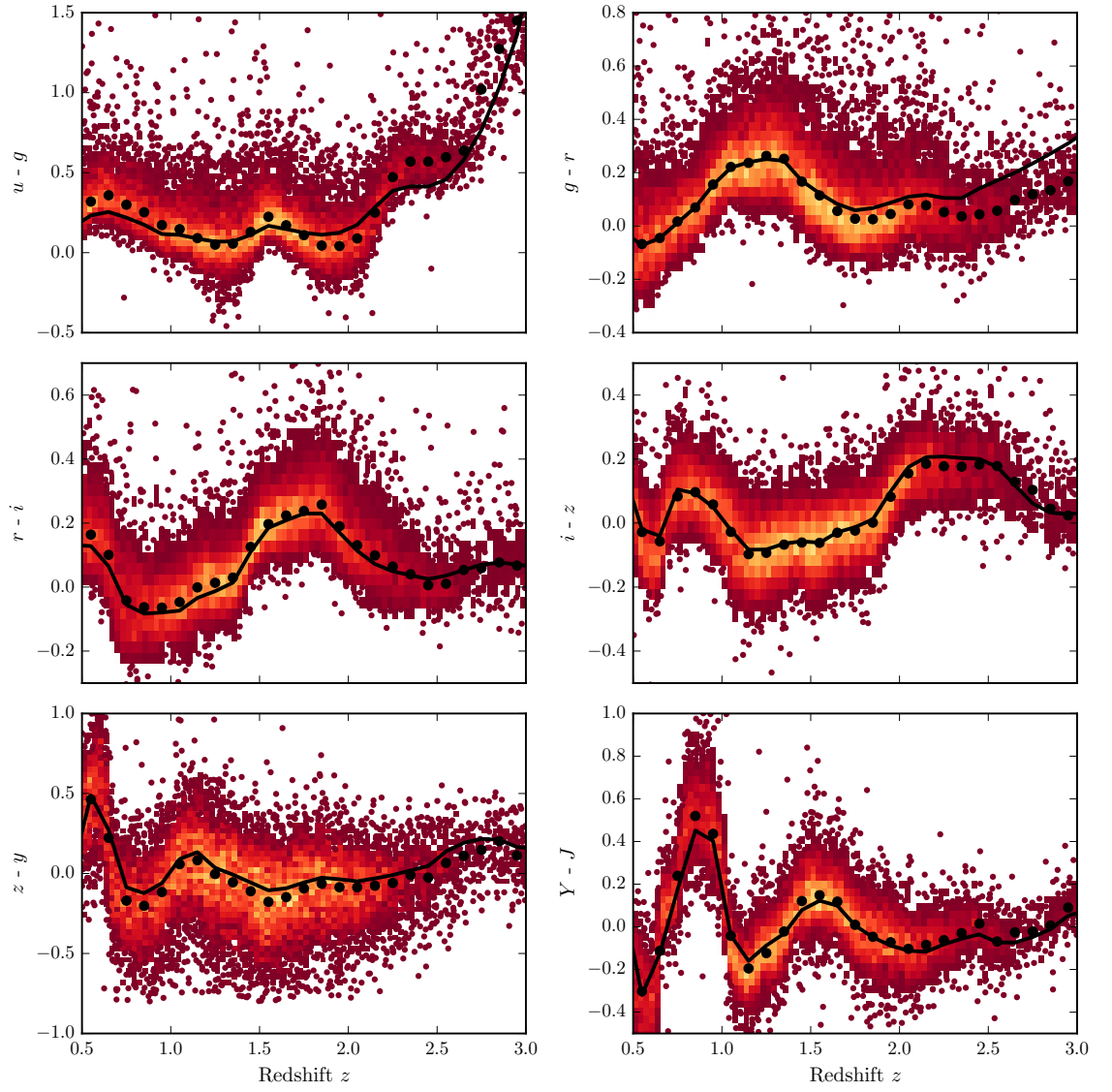


Figure 5.3: Colours of median SED (black circles), individual objects (grey points), best-fitting model (black line) as a function of redshift.

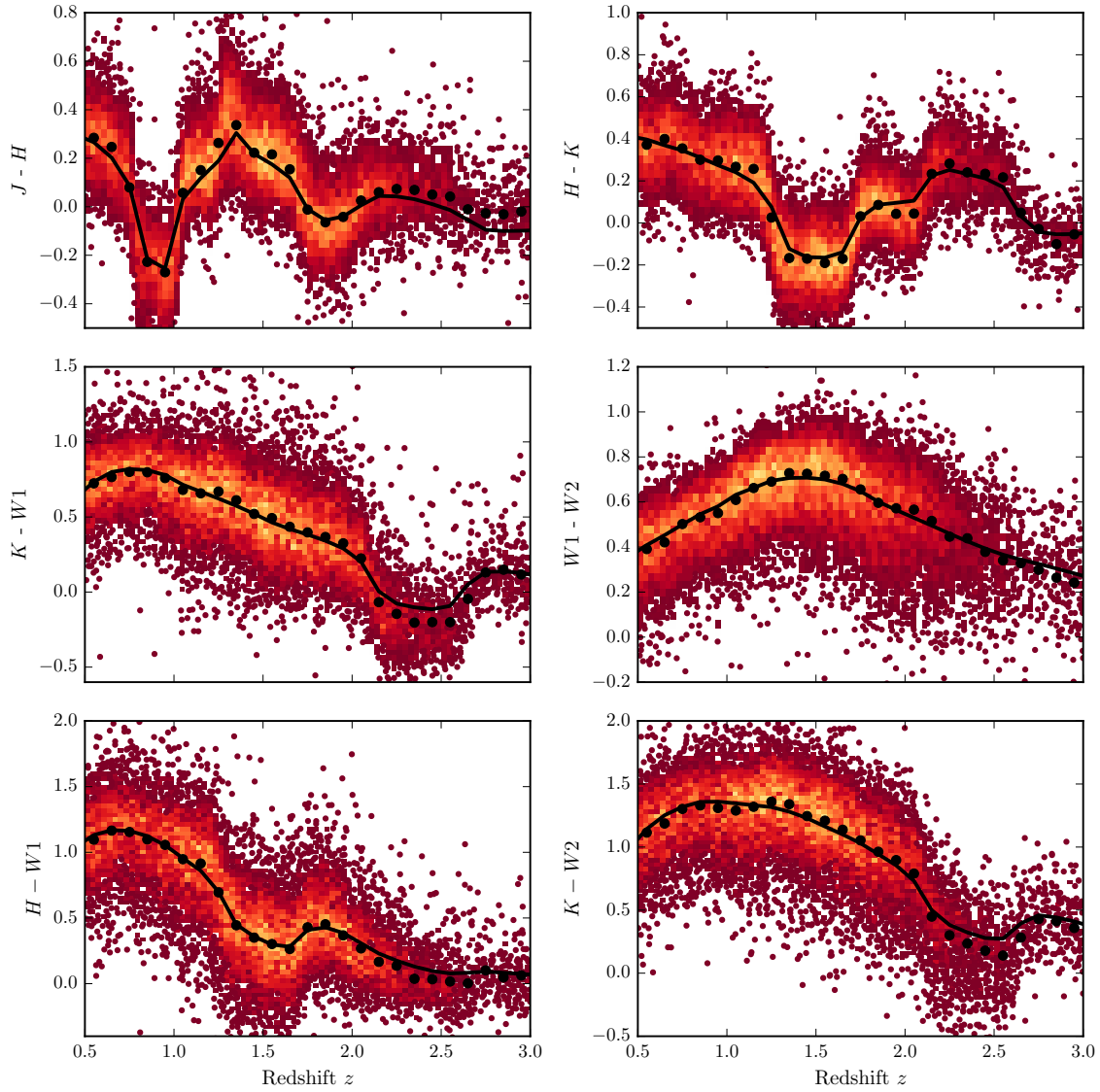


Figure 5.4: Colours of median SED (*black circles*), individual objects (*grey points*), best-fitting model (*black line*) as a function of redshift.

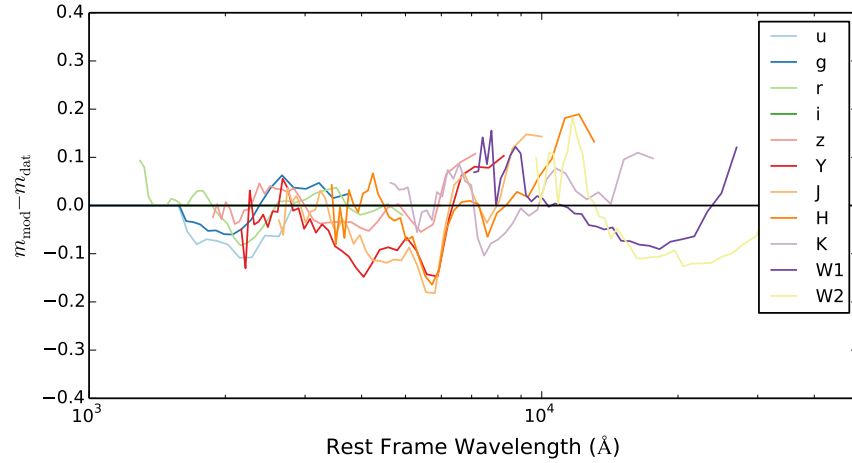


Figure 5.5: Residuals from fit to DR7Q-matched catalogue as a function of rest-frame wavelength.

5.4.1 Flux Correction

5.5 HOT DUST

Including a black-body with $T \sim 1250\text{K}$, a simple parametric model matches the ugrizYJHKW1W2 (SDSS+UKIDSS+WISE) median colours of luminous quasars at redshifts $0.2 < z < 0.4$ extraordinarily well. The spread in the KW1W2 colours (Figure 5.8), probing the rest-frame $\sim 1\text{-}2$ micron region, is significant and strongly suggests presence of real variation in the "hot dust" temperature and luminosity among the quasars. Several other investigations have drawn attention to the rest-frame near-infrared SEDs, with populations of 'dust free' objects postulated (Hao et al., 2010; Hao et al., 2011; Jiang et al., 2010; Mor and Trakhtenbrot, 2011)

Reverberation measurements of nearby AGNs suggest that hot dust very close to the central source (few tens of light days; e.g. Minezaki et al., 2004; Suganuma et al., 2006). The hot dust signature could contain information about inner face of an obscuring torus structure and/or constrain the dust content of an accretion disk wind. Several studies have shown that the luminosity of the NIR excess emission correlates with that of the central engine with a slope close to unity e.g. Gallagher et al., 2007, suggesting that the dust is reprocessing radiation from the accretion disc.

Outflows may emerge from the outer region of the accretion disc or even the innermost region of the torus, in which the gas clouds are dusty and relatively cold. Indeed, there is observational evidence for dusty outflows close to the central engine (e.g. Bowler et al., 2014). The dust is heated by the central engine, and radiates in the near-infrared band. Wang et al., (2013), fitting the NIR emission with a

*I have text on
empirical correction.
Re-do once I am
happy with SED
model*

single power-law, found that objects with strong outflow signatures (blue-shifted C IV) have more hot dust emission relative to the accretion disc emission in a large sample of $z \sim 2$ non-BAL quasars. It could be that this correlation is induced by a third factor that simultaneously affects outflows and dust emission, for instance the inclination angle or metallicity. Alternatively the dust could be intrinsic to outflows and may have a non-trivial contribution to the outflow acceleration.

5.5.1 *Parameterising the hot dust emission*

We characterise the hot dust properties of our sample in terms of the temperature and luminosity of a blackbody. We choose to parameterise the luminosity in terms of the NIR to UV luminosity ratio (which is proportional to the covering factor of hot dust ($L_{\text{NIR}}/L_{\text{Bol}}$) used in other studies (Roseboom et al., 2013). The UV and NIR luminosity are calculated between 2000 and 9000 Å and 1 and 3 μm respectively.

In Figure 5.6 we see that the two parameters are clearly correlated. For a lower temperature black-body the NIR to UV luminosity ratio is larger. Such a correlation is to be expected: as the black-body temperature is lowered, the peak shifts to longer-wavelengths (following Wien’s displacement law). Because of this degeneracy we need to be very careful to separate out real trends of $R_{\text{NIR/UV}}$ with other quasar properties from indirect trends resulting from a mutual dependence on T_{BB} .

Some previous studies (e.g. Wang et al., 2013; Zhang et al., 2014) have instead parameterised the near-IR emission using a power-law. We tested this parameterisation, and evaluated its effectiveness relative to using a black-body. The power-law is normalised at 9000 Å, where its flux is set equal to the flux of the UV/optical model. The NIR power-law slope is fit between ~ 1 and $2.4 \mu\text{m}$ (with the exact wavelength region being fit depending on the redshift of the quasar). We found large residuals in the best-fitting model which varied systematically as a function of $\lambda_{\text{eff}}/(1+z)$. This suggests that the power-law model is a poor fit to the shape of the near-IR emission. One needs to take care in looking at trends with luminosity given the observed-frame passband information on the rest-frame SED can produce some strong systematics with redshift, particularly if the SED-model is not a good fit to the actual SED. A similar conclusion was reached by Gallagher et al.

5.5.2 *Sample*

Our goal is to determine the temperature and abundance of the hot dust component in individual quasars. These properties will be mea-

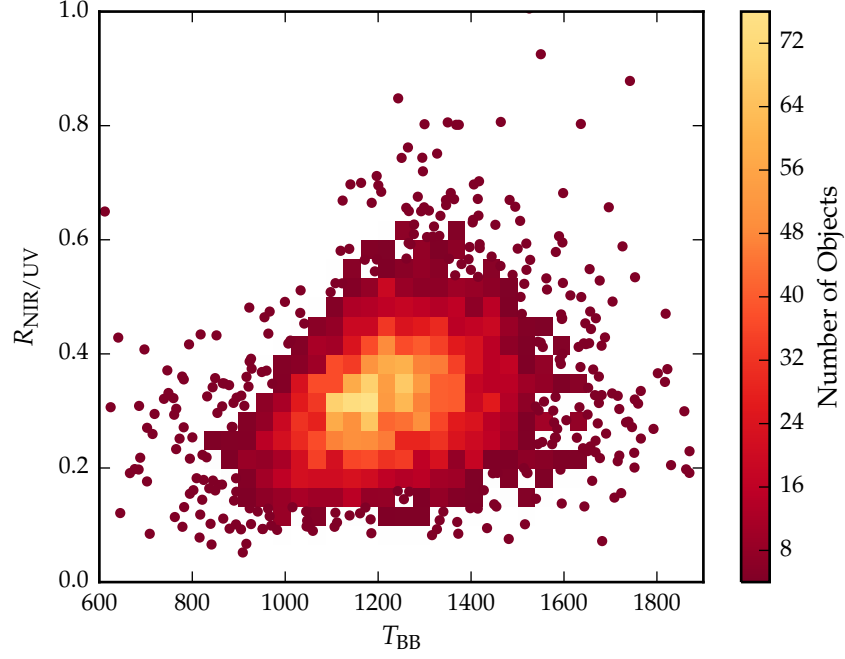


Figure 5.6: Ratio of NIR to UV luminosity ($R_{\text{NIR/UV}}$) against temperature (T_{BB}) for low- z sample. The density of points is shown in more dense regions of the space, and individual objects in less dense regions.

sured by fitting a model to the SDSS-UKIDSS-WISE photometry. Constraining a $T \sim 1200\text{K}$ blackbody component in the SED model requires photometric data covering $\sim 1\text{--}3\mu\text{m}$ in the rest-frame of the quasar.

The observed-frame wavelength coverage of the available passbands limits the redshift range of the quasars which can be used. One does need to take care in looking at trends with luminosity given the observed-frame passband information on the rest-frame SED can produce some strong systematics with redshift, particularly if the SED-model is not a good fit to the actual SED. We consider only quasars at redshifts $z > 1$ where the relative host galaxy contribution to the SED is negligible. At redshifts $1 \lesssim z \lesssim 1.5$ the available $ugrizYJHKW_1W_2$ photometry provides good coverage of the rest-frame SED up to $\sim 2\mu\text{m}$. At $z \sim 1.5$ the W_2 passband is shifted to $\sim 1.8\mu\text{m}$; at higher redshifts the wavelength coverage of the W_2 band becomes much less than the peak wavelength of a $T \sim 1200\text{K}$ blackbody and experiments showed that such a component can not be adequately constrained by the available photometry. Quasars in the redshift interval $1.5 < z < 2$ are therefore excluded from our sample.

For the quasars at $z \sim 1$ the WISE W_3 band is probing rest-frame wavelengths of $\sim 5\text{--}6\mu\text{m}$. This region of the SED is dominated by emission from cooler, more distant dust, which is not accounted for in our model. However, at redshifts $z \gtrsim 2$ the WISE W_3 passband probes

sufficiently short wavelengths to be useful in constraining the shape of the hot blackbody component. Therefore for quasars at redshifts $z > 2$ we again have sufficient constraints from the $ugrizYJHKW_1W_2W_3$ photometry to determine the temperature and normalisation of the blackbody component. There are few objects in our sample with redshifts $z > 2.7$, and so we set this as an upper limit on the redshift of our sample. Because of these constraints, our sample is divided into two parts: one at low redshifts ($1 < z < 1.5$) and the other at higher redshifts ($2 < z < 2.7$).

We include only quasars with observed magnitudes brighter than 19.1 in the i band-pass, i.e. the quasars selected by the main SDSS quasar selection algorithm (70,214 quasars). Cross-matching (with a $2''$ radius and picking only the nearest neighbour) the SDSS DR7Q catalogue with the ULAS catalogue, which covers only $\sim 38\%$ of the SDSS foot-print, resulted in 37,886 matches. Of these 36,628 have been detected in one or more of the WISE band-passes. We exclude quasars flagged as broad-absorption line quasars from the sample (leaving 35,272 quasars). We impose a lower-limit signal-to-noise ratio (S/N) > 5 magnitudes in the K , W_1 and W_2 band-passes for the low- z sample and $S/N > 5$ in the W_1 , W_2 , and W_3 band-passes for the high- z sample to ensure reliable photometry. This gives us 5,910 quasars in our low- z sample and 1,989 quasars in our high- z sample.

We will hold most parameters fixed, and vary only those we are interested in, i.e. the blackbody parameters which parameterise the NIR emission. Therefore we need to define a sub-sample of objects which we know are well fit by our standard SED model in the UV/optical region. This means excluding objects with extreme emission line equivalent widths or significant dust extinction. We use the $i - K$ colours of the quasars as a measure of the overall colour of the quasars as it provides the longest baseline in wavelength without being affected by absorption in the $Ly\alpha$ forest at high redshifts. We discarded from our sample quasars with $i - K$ colors redder than our standard model with dust reddening $E(B-V) = 0.075$ and bluer than $E(B-V) = -0.075$ (Figure 5.7). Following this cut we are left with 4,615 quasars in our low- z sample and 1,692 quasars in our high- z sample.

5.5.3 Diversity in hot dust properties

In Figure 5.8 we plot the $W_1 - W_2$ colors of the DR7Q-matched sample as a function of redshift at $z < 3$. In this redshift range the W_1 and W_2 band-passes are probing the $1.2 - 2.8\mu\text{m}$ and $1.6 - 3.8\mu\text{m}$ region of the rest frame SED respectively. For reference, the peak wavelength is at $2.4\mu\text{m}$ for a black-body radiating at 1200K (close to the sublimation temperature of dust grains). At any given redshift we see a ~ 0.5 mag dispersion in the $W_1 - W_2$ colors.

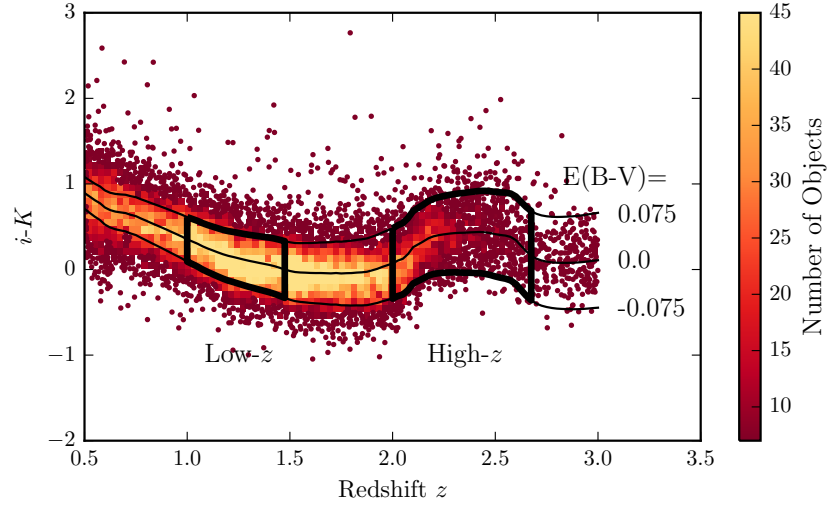


Figure 5.7: $i - K$ vs z . Demonstrates how sample was defined. The grey points show, as a function of redshift, the $i - K$ colours of all DR7Q quasars which are not classified as broad-absorption line quasars by Shen et al. and i magnitude > 19.1 . The black line shows the $i - K$ colour of our standard, unreddened SED model as a function of redshift. The red and blue lines show the $i - K$ colours of our SED model with dust reddening $E(B-V) = 0.075$ and $E(B-V) = -0.075$ respectively. A significant amount of this reddening can be attributed to intrinsic variations in the UV power-law slopes of the individual quasars, which is why we allow a negative reddening. However, there is a clear ‘red tail’ to the colour distribution which can be explained by dust reddening at the redshift of the quasar. We defined two samples, at low ($0.5 < z < 1.5$) and high ($2 < z < 2.7$) redshift, which are shown in the figure.

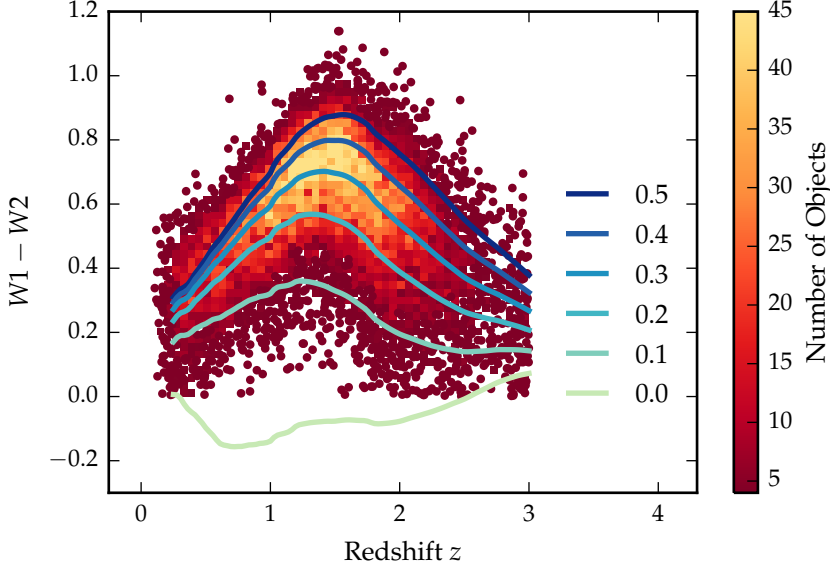


Figure 5.8: $W1 - W2$ colours of DR7 sample as a function of redshift. Above a certain density threshold points are represented by a density plot. On top we plot the colours of our standard SED model, with a fixed temperature and a varying NIR ($1 - 3 \mu\text{m}$) to UV ratio.

On the same axes in the Figure we have plotted the $W1 - W2$ colours derived from our SED model with a fixed blackbody temperature (1216K) and a ratio of NIR to UV luminosity ranging from 0.0 to 1.0, with the other model parameters held constant. We conclude that even with the sample restricted to be fairly uniform in its UV/optical properties, we still get an interesting spread in $W1-W2$ colors, which we can use to learn about the diversity of NIR properties in our sample. In the rest of this chapter we will characterise the hot dust properties of our sample, and test its relation to quasar properties such as luminosity, black-hole mass and normalised accretion rate, and outflow-properties.

In Figure 5.6 we show that there is quite a range of temperature and normalisation present in our sample. However, we need to check how much of this is due simply to uncertainties in the fits stemming from uncertainties in the photometry. In order to achieve this we took our standard SED model with a single temperature and normalisation black-body component, and generated 200 mock SEDs with a brightness distribution similar to that of our real sample. We estimated the mean uncertainty of the magnitudes in the K, $W1$, and $W2$ band-passes as a function of apparent brightness. We then sampled the K, $W1$, and $W2$ magnitudes from Gaussian distributions, with a mean equal to the magnitude of the model SED, and the width equal to the mean uncertainty at the appropriate brightness. Finally,

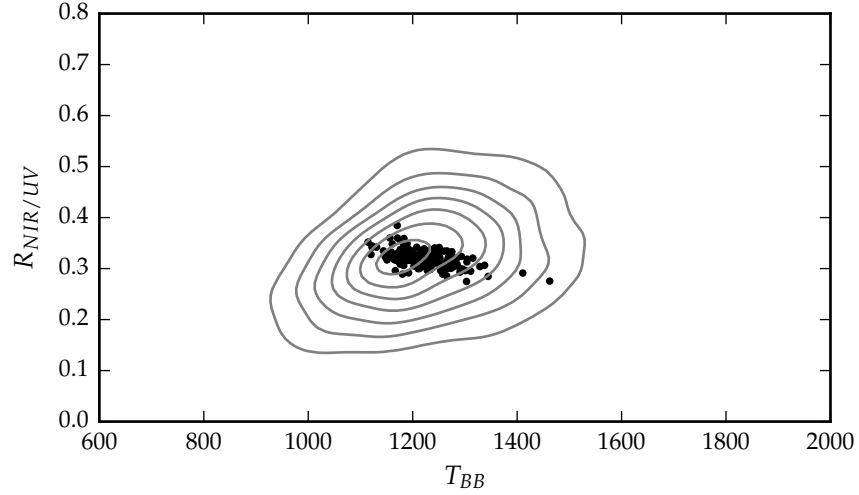


Figure 5.9: Ratio of NIR to UV luminosity ($R_{\text{NIR/UV}}$) against temperature (T_{BB}). The grey contours show equally-spaced lines of constant probability density generated using a Gaussian kernel-density estimator on our data sample. The black points are for our mock data.

we fit these mock SEDs using our standard fitting procedure. The results are shown in the Figure below, on top of the results from our real sample (shown as grey contours). We can see that uncertainty in the photometry introduces a significant scatter to the temperature, but that this scatter is less than the intrinsic scatter in the data. This demonstrates that there is a real distribution of hot dust temperatures and luminosities in our sample.

5.6 FITTING PROCEDURE

We will fit a model to the individual quasar SEDs, allowing the temperature and normalisation of the black body component to vary. The model spectrum is redshifted to the redshift of the quasar being fit and is then multiplied by the `ugrizYJHMW1W2W3` throughput functions and normalised appropriately to give AB magnitudes. To fit the model to the data we minimise the sum of the squares of the differences between the elements in the model magnitude array and the elements in the data magnitude array. To avoid significant absorption in the Lyman- α forest at high- z , we restrict our fitting to wavelengths greater than 2000Å; when the effective wavelength of a band-pass falls below this limit the band-pass is excluded from the fit. The minimisation is done using the ‘nelder-mead’ method, as implemented in the `minimize` function from the Python module `scipy`.

*2000Å is quite large
given the
Lyman-alpha forest
impacts from 1216Å.*

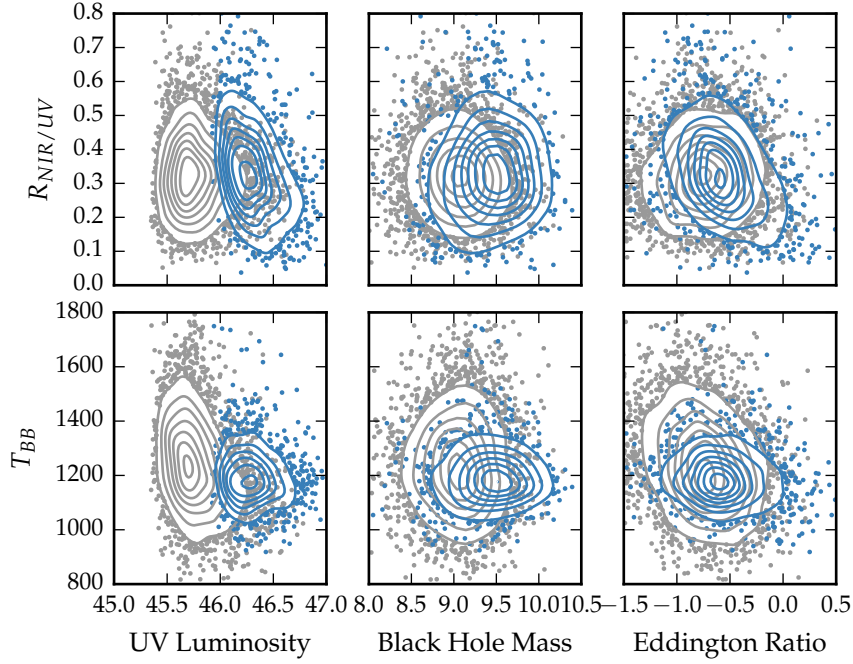


Figure 5.10: Best-fit black-body temperature against UV luminosity (left), black-hole mass (center) and Eddington ratio (right) for $1 < z < 1.5$ sample (black) and $2 < z < 2.7$ sample (blue). In region of high-density we represent the density with contours generated using a Gaussian kernel density estimation. **Needs re-making with new BH masses.**

5.7 RESULTS

5.7.1 Correlations with quasar properties

We now look for correlations between the properties of the black-bodies we have fitted to the hot dust emission and other properties of the quasar such as redshift, black-hole mass, and normalised accretion rate (Eddington ratio).

5.7.2 Spectral properties

In the dusty wind model - first proposed by Konigl and Kartje, (1994) and later developed by, amongst others, Everett, (2005), Elitzur and Shlosman, (2006), Keating et al., (2012) - the ‘torus’ is the dusty part of a magneto-hydrodynamic wind beyond the dust sublimation radius. The MHD wind is roughly polar, and so the hot dust forms a vertical ‘wall’ around the accretion disk. UV photons from the accretion disk accelerate the wind via radiation line driving. That flattens the geometry of the wind and exposes more surface area that is viewable on a relatively face-on line of sight. The radiation pressure is increased

Calculate new BH masses and redo this section.

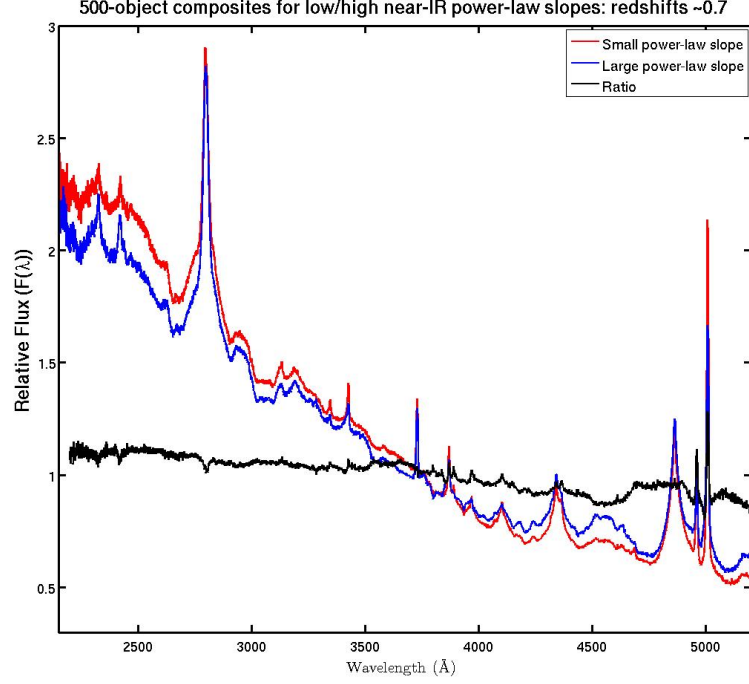


Figure 5.11: Composite SDSS spectra for objects at $z \sim 0.7$. We have divided sample into objects with objects best-fit by small (red line) and large (red line) values of β . **Change this to select by $R_{\text{NIR/UV}} / T_{\text{BB}}$. Label prominent emission lines.**

at higher luminosities and/or accretion rates. This can flatten the geometry of the wind, thereby increasing the range of angles for which the inner edge of the dusty wind - where dust is at its sublimation temperature - can be observed. A direct prediction is therefore that the in a quasars with high accretion rates and strong outflows, the emission from hot dust should be enhanced.

5.7.2.1 Low- z

The $z < 0.8$ SDSS spectrum composite comparison for the small and large β_{NIR} sub-samples is a very direct illustration of the Boroson and Green, (1992) Eigenvector 1 describing the spectral variation in the optical spectra of quasars; as Fe II EW increases the [O III] EW decreases. Hot dust emission increases with Fe II EW (Shen and Ho, 2014). We also note that the amount of hot dust correlates with the Si III/C III] emission ratios. The Si III/C III] ratio is generally considered to be a good indicator of density and is one of the primary EV1 correlates. The relative flux ratio of Si III to C III] increases when C IV is more blue-shifted (Richards et al., 2011). The Mg II emission line has exactly the same profile/shape for the two samples (apparent changes in Mg II seen in Fig. 5.11 are the result of changes in Fe II at

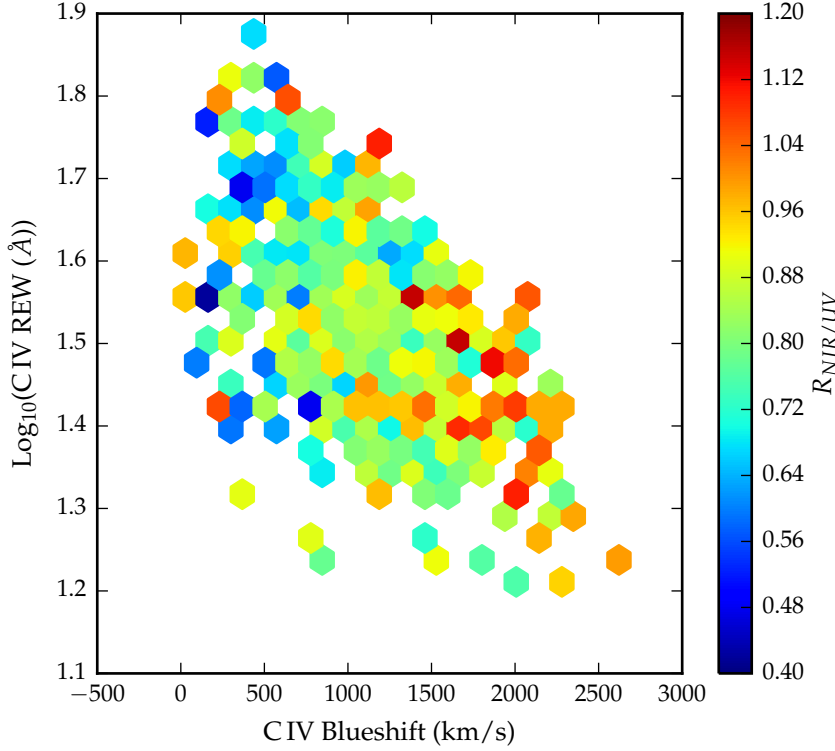


Figure 5.12: Rest-frame equivalent width and blueshift of the C iv line for 7,115 SDSS DR7 quasars. The colours of the hexagons denote the median hot dust ($T \simeq 1200$ K) abundance for all quasars at a given equivalent width and blueshift. Quasars with the most extreme outflow signatures are predominantly hot-dust rich. Only bins containing a minimum of two objects are plotted.

wavelengths just shortward of the line). Finally, we note that objects with more hot dust are slightly redder.

5.7.2.2 High- z

In Fig. 5.12 we show how the ratio of NIR to UV luminosity depends on the blueshift and rest-frame equivalent width of the C iv line. C iv blueshifts are calculated as in Section XX. We see that the NIR to UV luminosity ratio is strongly correlated with the blue-shift of the C iv emission line. A similar trend was noted by Wang et al., (2013). Interestingly, we note strong similarities to the object subsets selected according to their C iv-emission properties in Richards et al., (2011) (see Figures 11 & 12). We note that the correlation between the hot dust and the C iv emission properties will lead to apparent correlations between the host dust and the BH mass.

Need to re-do this and understand why beta-related trend is apparently stronger than with the blackbody parameters.

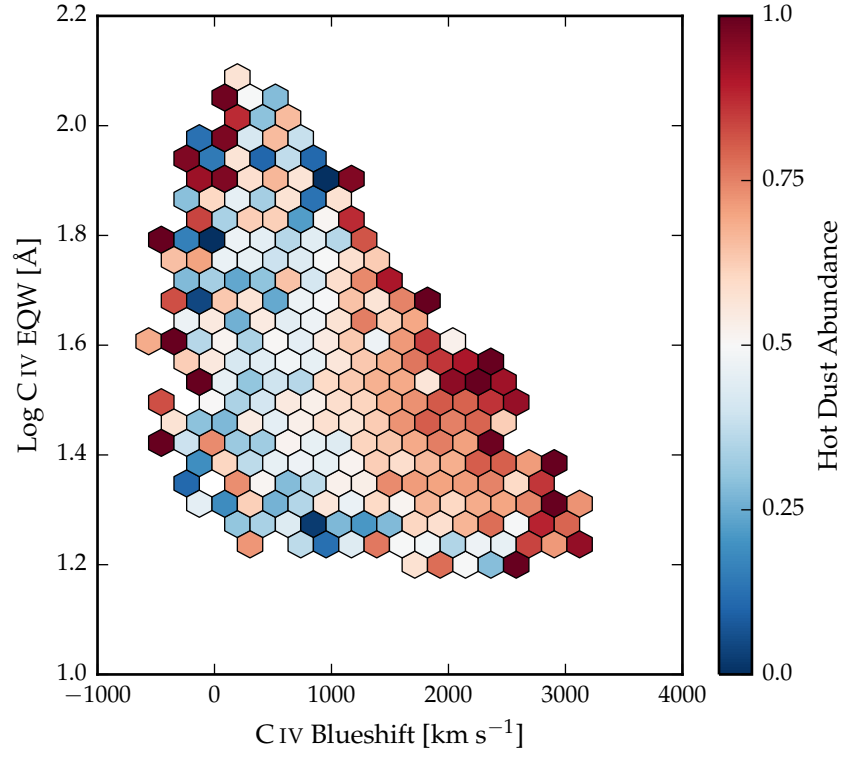


Figure 5.13

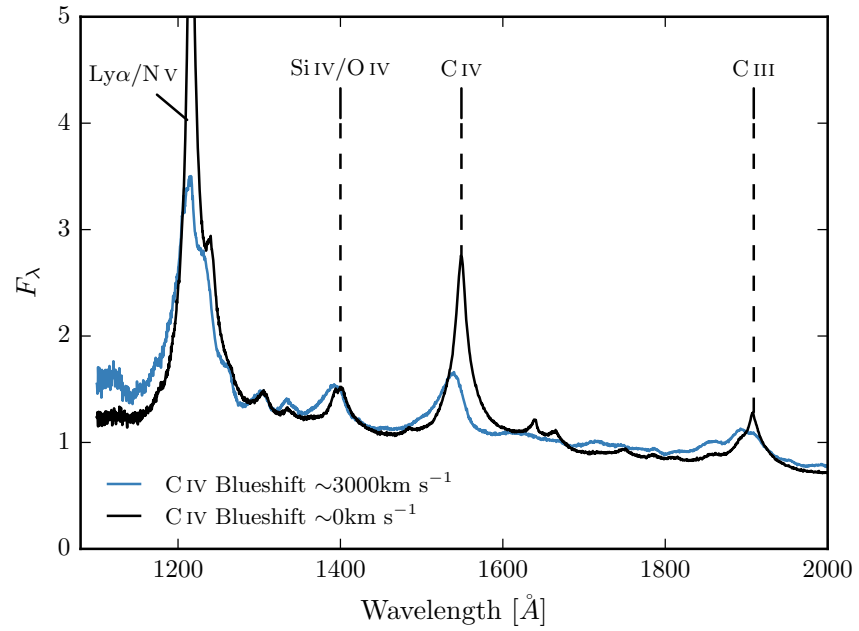


Figure 5.14

5.7.3 BALs and radio-loud/radio-quiet

In the spectra of about 20% of all quasars we observe broad ($> 2000\text{km/s}$) blue-shifted absorption troughs which are associated with quasar-driven out-flowing gas. BAL quasars in general have redder UV continua than non-BAL quasars, which is interpreted as the result of dust extinction. BAL quasars, on average, also have higher Eddington ratios and luminosities than non-BAL quasars.

We defined a sample of BAL quasars using the same method we used to define our sample of non-BAL quasars. At $1 < z < 1.5$ there are very few BAL quasars in our sample. In the $2 < z < 2.7$ redshift region we have 394 HiBAL quasars (the wavelength coverage of the SDSS spectra are not sensitive to LoBALs at these redshifts). Since BAL quasars are expected to suffer more from extinction due to dust, we have allowed $E(B-V)$ to vary for the BAL quasar sample.

We find that the black-body temperature distributions are consistent (median T_{BB} for both samples = 1180K), but the ratio of NIR to UV emission is higher in BALs ($R_{\text{NIR/UV}} = 0.92$ and 0.83 for BAL and non-BAL quasar sample respectively). This is qualitatively consistent with the results of Zhang et al., (2014).

It is well known that the blueshift of the C IV emission line in radio-quiet AGNs is, on average, stronger than in radio-loud AGNs (Marziani et al. 1996; Sulentic et al. 2000a; Richards et al. 2002, 2011). Statistically at least, the "radio-loud" objects are thought to have high black-hole masses and there is some form of radio-mode feedback (jet related) which is very different from the much more common (almost certainly wider opening-angle) outflow objects with large C IV-blueshifts.

So we find BALs have more hot dust, radio-loud have less. This is perfectly consistent with what we know about the positions of radio-loud objects and BALs in the C IV parameter space distribution (Richards et al., 2011).

Need to justify our use of model which has been optimised to fit colours of non-BAL quasars, when we know that BALs are typically redder.

What catalogue did we use to define quasar sample?

5.8 OTHER WORKS

Roseboom et al., (2013) studied a similar sample of luminous type 1 quasars. They, like us, modelled the NIR emission using a black-body and modelled the emission at longer wavelengths using a clumpy torus model. They find that while $L_{1-5\mu\text{m}}/L_{\text{IR}}$ appears relatively insensitive to L_{bol} and L_{IR} , a strong correlation appears between $L_{1-5\mu\text{m}}/L_{\text{IR}}$ and $L_{\text{IR}}/L_{\text{bol}}$ (i.e. the dust covering factor). As the covering factor decreases, the maximum inclination at which a type 1 quasar would be seen increases. An increase in the inclination will mean direct sight lines to more of the inner wall of obscuring material closest to the accretion disc.

Mor and Trakhtenbrot, (2011) also looked at the hot dust properties of a sample of $0.75 < z < 2$ quasars, with photometry from SDSS and WISE. They modelled the NIR emission with hot clouds of pure graphite dust. They reported an anti-correlation between the covering factor of hot dust clouds and the quasar bolometric luminosity. Like us, they neglect cooler dust components which will dominate the SED at longer wavelengths. As we have discovered (see Figure residual plot), the missing flux decreases with redshift because we observe shorter rest-frame wavelengths when the observed spectrum is redshifted to a greater degree. This will induce an anti-correlation between the luminosity of the hot dust component and the luminosity of the quasar (which is correlated with redshift). At $z=0.75$, the W3 band-pass (the longest in their fits) is sensitive to flux from $6.9\mu\text{m}$; at this wavelength we expect the contribution from cooler dust to dominate over the hot dust. It is possible that this effect could explain the tension with our own result that $R_{\text{NIR/UV}}$ does not depend on the quasar luminosity in our low- z sample.

Shen and Ho, (2014) quantify the relative torus emission using the $r - W1$ colour for a sample of $0.4 < z < 0.8$ SDSS quasars. At these redshifts W1 is observing between 1.9 and 2.4 microns in the rest-frame of the quasar, which suggests that they are sensitive to the same component of hot dust which we are investigating. They observe a mild trend of decreasing relative torus emission as the quasar luminosity increases. We note that their use of the $r - W1$ at much higher redshifts may be problematic, as the W1 flux will be increasingly dominated by direct emission from the accretion disc.

Gallagher et al., (2007) undertook a similar investigation for a much smaller sample of 234 radio-quiet quasars.

5.8.1 Eddington ratio

Wang et al., Zhang et al., and Mor & Trakhtenbrot find no significant dependence of the amount of hot dust on the Eddington ratio.

Shen & Ho find that torus emission is enhanced in quasars with larger R_{FeII} . They show how EW(OIII) and other high-ionisation lines (and to a lesser extent low-ionisation lines like MgII) anti correlate with R_{FeII} . The enhancement of torus emission relative to accretion disc emission at the high- R_{FeII} end of EV1 may be caused by more efficient disc winds that facilitate the formation of a dusty torus. From our $z \sim 0.8$ composite SDSS spectra, we observed that objects with large NIR to UV luminosity ratios on average have stronger FeII emission.

CONCLUSIONS / FUTURE WORK

Put some stuff from research proposals here

6.1 FUTURE: RED QUASARS

Punctuated fuelling episodes, e.g. driven by galaxy mergers, satellite accretion and even secular processes, almost certainly lead to AGN experiencing activity-, outflow- and obscuration-dominated cycles with some overlap between phases. However, quantitatively, it remains unclear how these phases relate to the fundamental properties of the accreting black-hole (e.g. mass (M_{BH}), bolometric luminosity (L_{bol}) and Eddington ratio (L/L_{Edd}) and the elements of the non-spherical geometry).

