

The correlation of line strength with luminosity and redshift from composite quasi-stellar object spectra

S. M. Croom,^{1*} K. Rhook,¹ E. A. Corbett,¹ B. J. Boyle,¹ H. Netzer,²
N. S. Loaring,³ L. Miller,³ P. J. Outram,⁴ T. Shanks⁴ and R. J. Smith⁵

¹Anglo-Australian Observatory, PO Box 296, Epping, NSW 1710, Australia

²School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel

³Department of Physics, Oxford University, Keble Road, Oxford OX1 3RH

⁴Physics Department, University of Durham, South Road, Durham DH1 3LE

⁵Astrophysics Research Institute, Liverpool John Moores University, 12 Quays House, Egerton Wharf, Birkenhead CH41 1TD

Accepted 2002 July 24. Received 2002 July 23; in original form 2002 June 27

ABSTRACT

We have generated a series of composite quasi-stellar object (QSO) spectra using over 22 000 individual low-resolution ($\sim 8\text{-}\text{\AA}$) QSO spectra obtained from the 2dF ($18.25 < b_j < 20.85$) and 6dF ($16 < b_j \leq 18.25$) QSO Redshift Surveys. The large size of the catalogue has enabled us to construct composite spectra in relatively narrow redshift ($\Delta z = 0.25$) and absolute magnitude ($\Delta M_B = 0.5$) bins. The median number of QSOs in each composite spectrum is ~ 200 , yielding typical signal-to-noise ratios of ~ 100 . For a given redshift interval, the composite spectra cover a factor of over 25 in luminosity. For a given luminosity, many of the major QSO emission lines (e.g. Mg II $\lambda 2798$, [O II] $\lambda 3727$) can be observed over a redshift range of 1 or greater.

Using the composite spectra we have measured the line strengths (equivalent widths) of the major broad and narrow emission lines. We have also measured the equivalent width of the Ca II $\lambda 3933$ K absorption feature caused by the host galaxy of the active galactic nuclei (AGN). Under the assumption of a fixed host galaxy spectral energy distribution (SED), the correlation seen between Ca II K equivalent width and source luminosity implies $L_{\text{gal}} \propto L_{\text{QSO}}^{0.42 \pm 0.05}$. We find strong anticorrelations with luminosity for the equivalent widths of [O II] $\lambda 3727$ and [Ne V] $\lambda 3426$. These provide hints to the general fading of the NLR in high-luminosity sources, which we attribute to the NLR dimensions becoming larger than the host galaxy. This could have important implications for the search for type 2 AGN at high redshifts. If average AGN host galaxies have SEDs similar to average galaxies, then the observed narrow [O II] emission could be solely a result of the host galaxy at low luminosities ($M_B \sim -20$). This suggests that the [O II] line observed in high-luminosity AGN may be emitted, to a large part, by intense star-forming regions. The AGN contribution to this line could be weaker than previously assumed.

We measure highly significant Baldwin effects for most broad emission lines (C IV $\lambda 1549$, C III] $\lambda 1909$, Mg II $\lambda 2798$, H γ , H β) and show that they are predominantly caused by correlations with luminosity, not redshift. We find that the H β and H γ Balmer lines show an *inverse Baldwin effect* and are positively correlated with luminosity, unlike the broad ultraviolet lines. We postulate that this previously unknown effect is caused by a luminosity-dependent change in the ratio of disc to non-disc continuum components.

Key words: galaxies: active – quasars: emission lines – quasars: general – galaxies: stellar content.

1 INTRODUCTION

The correlation of quasi-stellar object (QSO) emission-line properties with luminosity is a straightforward yet potentially highly informative test of standard physical models for active galactic

*E-mail: scroom@aoepp.aao.gov.au

nuclei (AGN). Since the discovery of an anticorrelation between the equivalent width (W_λ) of the C IV $\lambda 1549$ emission line and the continuum luminosity (L) by Baldwin (1977), a significant amount of effort has been expended to quantify this relationship (hereinafter referred to as the Baldwin effect), and investigating similar correlations with other QSO emission lines (Baldwin et al. 1989; Zamorani et al. 1992; Green, Forster & Kuraszkiewicz 2001). The results have revealed that the anticorrelation with luminosity is relatively weak, typically $W_\lambda \propto L^\beta$, with $\beta = -0.2$ and a large scatter. Similar correlations have been seen in most other broad emission lines including Mg II $\lambda 2798$, C III] $\lambda 1909$, Si IV+O IV] $\lambda 1400$ and Ly α with $-0.4 < \beta < -0.1$ (Green et al. 2001). It has also been claimed (Green et al. 2001) that the Baldwin effect may be dominated by an even stronger anticorrelation with redshift. However, in the magnitude-limited QSO samples that have been studied to date, it is extremely difficult to disentangle the effects of redshift and luminosity. It is typically only possible to access 1–1.5 mag at any given redshift, given the steep slope of the QSO luminosity function for magnitude-limited samples with $B < 19.5$ (Boyle, Shanks & Peterson 1988).

A further limitation of existing studies is that it is difficult to study the correlation with luminosity and/or redshift for weaker lines, in particular the narrow-line region (NLR). The spectra used in such analyses are typically ‘survey’ quality, i.e. relatively low signal-to-noise (S/N) ratio ($S/N \sim 5$ –10) and thus narrow emission lines can be difficult to detect in individual spectra.

Composite QSO spectra have been generated from most large QSO surveys over the past decade (Boyle 1990; Francis et al. 1991), providing a detailed picture of the ensemble average spectral properties of the QSO sample. Typical S/N ratios in these spectra approach or even exceed 100, with even relatively weak emission lines (e.g. [Ne V] $\lambda 3426$) easily detectable. However, previous surveys have been too small (comprising 1000 QSOs or less) to generate composite spectra as a function of both luminosity and redshift with which to examine correlations. With the recent advent of much larger QSO surveys such as the 2dF QSO Redshift Survey (2QZ, Croom et al. 2001) and Sloan Digital Sky Survey (SDSS, Vanden Berk et al. 2001; Schneider et al. 2002) we may now use composite, rather than individual spectra, to investigate the correlation of QSO spectral properties with luminosity and redshift in much greater detail that has hitherto been possible.

In this paper we describe the result of an analysis of composite QSO spectra based on the almost 22 000 QSOs observed to date (2002 January) in the 2QZ. The bulk of these objects lie around the break in the luminosity function (LF), thus providing a better sampling in luminosity at any given redshift than QSO surveys at the bright end of the LF (e.g. the Large Bright Quasar Survey; Hewett, Foltz & Chaffee 1995). Moreover, we have also included a few hundred brighter QSOs observed with the new 6° field (6dF) multi-object spectrographic facility on the UK Schmidt Telescope (Croom et al., in preparation) to increase the luminosity range studied at any given redshift to typically 3–4 mag. As well as providing a wide baseline over which to study correlations such as the Baldwin effect, this sampling of the QSO (L, z) plane provides an opportunity to disentangle the effects of luminosity and redshift.

In Section 2 we describe the data used in our analysis, while in Section 3 we discuss the methods used to generate the composite spectra and measure the spectral line equivalent widths. In Section 4 we present the results of our analysis, we then discuss these in the context of theoretical models in Section 5.

2 DATA

2.1 Generation of composite spectra

The data used in our analysis is taken from the 2dF and 6dF QSO Redshift Surveys (Croom et al. 2001; 6QZ Croom et al., in preparation). QSO candidates were selected for observation based on their stellar appearance and blue colours found from automated plate measurements (APM) of UK Schmidt Telescope (UKST) photographic plates and films in the u , b_j and r bands. The 2QZ/6QZ area comprises 30 UKST fields arranged in two 75×5 deg² declination strips centred on $\delta = -30^\circ$ and 0° . The $\delta = -30^\circ$ strip extends from $\alpha = 21^{\text{h}}40^{\text{m}}$ to $3^{\text{h}}15^{\text{m}}$ in the South Galactic Cap and the equatorial strip from $\alpha = 9^{\text{h}}50^{\text{m}}$ to $14^{\text{h}}50^{\text{m}}$ in the North Galactic Cap. The 2QZ and 6QZ sources were selected from the same photometric data, the only difference being their ranges in apparent magnitude: $18.25 < b_j < 20.85$ (2QZ) and $16.0 < b_j \leq 18.25$ (6QZ). The combined data sets thus produce a uniform QSO sample over a wide range in luminosity. Details of the candidate selection can be found in Smith et al. (2002).

The 2QZ objects were observed over the period 1997 October to 2002 January using the 2dF instrument at the Anglo-Australian Telescope. Observations were made with the low-dispersion 300B grating, providing a dispersion of $178.8 \text{ \AA mm}^{-1}$ ($4.3 \text{ \AA pixel}^{-1}$) and a resolution of $\simeq 8.6 \text{ \AA}$ over the range 3700–7900 \AA . Typical integration times were 55 min, in a range of observing conditions (1–2.5 arcsec seeing) resulting in median $S/N \sim 5 \text{ pixel}^{-1}$. The brighter 6QZ objects used in the present paper were observed in 2001 September using the 6dF facility at the UKST. A low-dispersion 250B grating was used to provide a dispersion of 286 \AA mm^{-1} ($3.6 \text{ \AA pixel}^{-1}$) and a resolution of $\simeq 11.3 \text{ \AA}$ over the range 3900–7600 \AA . Exposure times were typically 100 min resulting in median $S/N \sim 15 \text{ pixel}^{-1}$.

Data from both 2dF and 6dF were reduced using the pipeline data reduction system 2DFDR (Bailey et al. 2002). Identification of spectra and the determination of redshifts was carried out by a automated program, AUTOZ (Croom et al. 2001; Miller et al., in preparation). Each spectrum was checked by eye by two members of the team. In our analysis below we only include quality Class 1 identifications (96 per cent reliable identification), these being the best quality spectra. We also only take the best spectrum (based on quality class and then the S/N ratio) of each object in the case where there is more than one spectrum available. The combined 2QZ/6QZ data set provides us with 22 041 independent QSO spectra.

Typical redshift errors are $\sigma_z = 0.003$ and photometric errors in the b_j band are ~ 0.1 mag. Absolute magnitudes were computed from the observed photographic b_j magnitude, after correction for Galactic extinction (Schlegel, Finkbeiner & Davis 1998), using the K -corrections found by Cristiani & Vio (1990). Throughout we assume a flat cosmological world model with $\Omega_0 = 0.3$, $\lambda_0 = 0.7$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

3 METHOD

We have generated composite QSO spectra in discrete absolute magnitude ($\Delta M_B = 0.5$ mag) and redshift ($\Delta z = 0.25$) bins. The bin widths were chosen to give good resolution in luminosity and redshift, whilst typically retaining over 100 QSOs in at least $5M_B$ bins (a factor of 10 in luminosity) at each redshift (see Table 1). Once QSOs identified as broad absorption line (BAL) QSOs had been removed, there remained a total of 21 102 QSOs with which to

Table 1. The number of QSOs in each of our absolute magnitude–redshift (M_B – z) bins. For each bin the central redshift and absolute magnitude is displayed. The last column shows the total number of QSOs in each magnitude interval over all redshifts. In some M_B – z intervals there are only a small number of QSOs. In these cases the spectra in adjacent M_B intervals were combined together, an * or † indicates where this has been done. For example, in the $z = 0.375$ interval, QSOs in the $M_B = -24.75, -24.25$ and -23.75 bins were combined together.

M_B	Redshift												All
	0.125	0.375	0.625	0.875	1.125	1.375	1.625	1.875	2.125	2.375	2.625	2.875	
–29.25	–	–	–	–	–	–	–	–	–	1*	–	–	1*
–28.75	–	–	–	–	–	–	–	–	–	–	1*	–	1*
–28.25	–	–	–	–	–	–	–	–	1*	1*	2*	3*	7*
–27.75	–	–	–	–	–	–	2*	1*	8*	13*	15*	4*	43*
–27.25	–	–	–	–	1*	3*	5*	23*	48*	48	37*	13*	180
–26.75	–	–	–	1*	5*	10*	54	100	152	116	101	37*	585
–26.25	–	–	–	3*	13*	80	183	259	316	289	211	71	1432
–25.75	–	–	3*	12*	76	194	370	517	545	439	277	50	2483
–25.25	–	–	4*	77	218	434	667	788	791	446	99	–	3524
–24.75	–	2*	27*	178	441	766	956	942	341	25	–	–	3678
–24.25	–	7*	85	360	703	974	779	92	–	–	–	–	3000
–23.75	–	14*	217	546	778	456	11	–	–	–	–	–	2022
–23.25	2*	66	334	692	359	–	–	–	–	–	–	–	1453
–22.75	–	101	410	420	2	–	–	–	–	–	–	–	933
–22.25	2*	158	528	51	–	–	–	–	–	–	–	–	739
–21.75	8*	220	274	–	–	–	–	–	–	–	–	–	502
–21.25	15†	239	38	–	–	–	–	–	–	–	–	–	292
–20.75	22†	117	–	–	–	–	–	–	–	–	–	–	139
–20.25	22†	27*	–	–	–	–	–	–	–	–	–	–	49
–19.75	16*	7*	–	–	–	–	–	–	–	–	–	–	23*
–19.25	11*	–	–	–	–	–	–	–	–	–	–	–	11*
–18.75	4*	–	–	–	–	–	–	–	–	–	–	–	4*
–18.25	–	–	–	–	–	–	–	–	–	–	–	–	–
–17.75	1*	–	–	–	–	–	–	–	–	–	–	–	1*

generate the composite spectra. The most important issue relating to the construction of the composites was that the spectra were not flux calibrated. The effects of differential atmospheric refraction, corrector chromatic aberration and fibre positioning errors makes obtaining even a relative flux calibration for sources extremely challenging. We therefore chose not to attempt flux calibration of our spectra. We did, however, correct for absorption owing to the atmospheric telluric bands (the optical fibres also provide some absorption in these same bands). We summed all the spectra in a single observation in order to obtain a mean absorption correction, which was then applied to the data. Also, pixels which had anomalously high variance owing to residuals of night sky emission lines were flagged as bad and discarded from our analysis.

As the spectra were not flux calibrated we decided to normalize each spectrum to a continuum level as a function of wavelength. This allows us to measure equivalent widths, linewidths and line centres, however, we lose any information concerning continuum shape and absolute line strengths. Fitting the continuum relies on defining line-free parts of the spectrum. This is not always possible, particularly in regions of the spectrum dominated by weak Fe II emission. Our approach, therefore, was to remove all strong emission-line features, interpolating linearly between pseudo-continuum bands defined on each side of the line. The strong features removed, and the continuum bands defined are listed in Table 2. After removing these strong lines, a fourth-order polynomial was fitted to each spectrum, which was then used to divide the spectra, providing an approximate continuum normalization. In a second step to remove residual large-scale features in the spectrum, each spectrum was divided by a median filtered version using a wide box-car filter of width 201 pixels (each spectrum containing 1024 pixels or 1032 pixels for 2dF

Table 2. List of strong spectral features removed before continuum fitting. A simple linear interpolation is made between two ‘continuum’ bands defined on either side of the feature.

Feature	Blue cont. band (Å)	Red cont. band (Å)
Ly α +N v	1130–1155	1280–1290
Si iv+O iv]	1350–1360	1445–1470
C iv+He ii	1445–1470	1685–1705
C iii]+Al iii	1800–1830	1985–2020
Mg ii+Fe ii	2650–2685	3025–3065
[O ii]	3675–3705	3745–3785
[Ne iii]	3845–3855	3905–3920
H δ	4020–4050	4165–4200
H γ	4220–4270	4430–4460
Fe ii	4430–4460	4710–4760
H β + [O iii]	4710–4760	5080–5105
Fe ii	5080–5105	5450–5500
He i	5740–5790	5940–5980
H α	6320–6380	6745–6805

and 6dF data, respectively). At the edges of the spectrum the filter was reduced in size to a minimum half-width of 5 pixels.

The above processing was all carried out in the observed frame. After continuum normalization the spectra were shifted to the rest frame, interpolating linearly on to a uniform scale of 1 \AA pixel^{-1} . Finally, the composite spectra were produced by taking the median value of each pixel. For each pixel the median z and M_B of the contributing QSOs was also determined. We can then determine

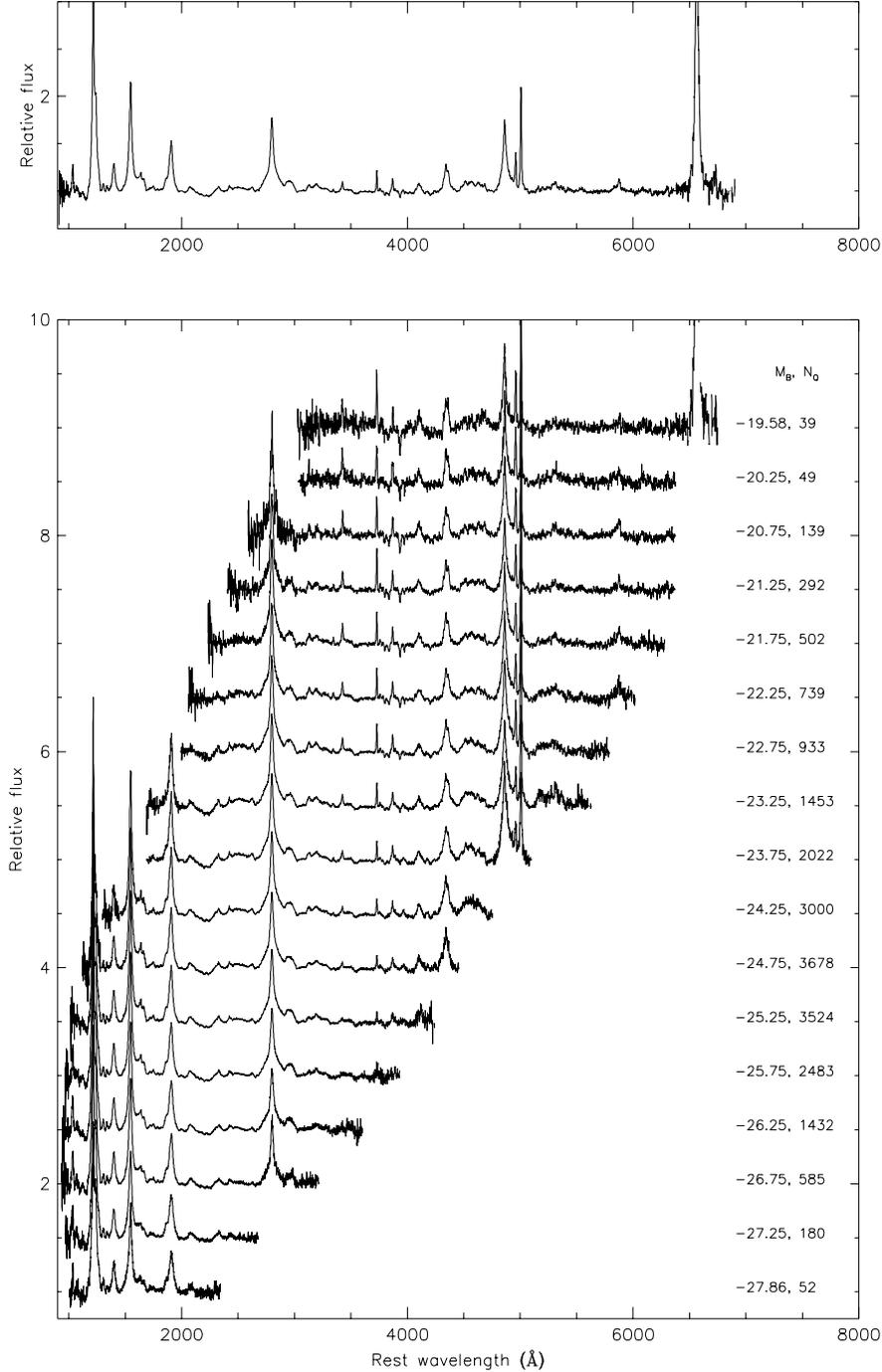


Figure 1. Composite QSO spectra. Top: all spectra combined into one composite. Bottom: composites computed in absolute magnitude intervals ($\Delta M_B = 0.5$) with no redshift binning. The brightest and faintest bins have been made wider to include a sufficient number of spectra. All the spectra have a continuum level of one, but have been offset for clarity.

appropriate values for each feature, and not just each composite. The values of z and M_B assigned to each feature are the average, over the wavelength range of the feature, of the pixel median z and M_B values. We derived errors for each composite by looking at the distribution of values to be medianed for each pixel. The 1σ errors were taken to be the 68 per cent semi-interquartile range of the pixel values divided by the square root of the number of objects contributing. We have constructed composites in z - M_B and also composites binned in absolute magnitude only. One final composite was made from all

the spectra (see Fig. 1). The resulting composites are normalized to the pseudo-continuum over most parts of the spectrum, except for the 2000–3500 Å region where our procedure treats the Fe II emission bands as though they were a continuum. Therefore, in our subsequent analysis of these spectra we are unable to deduce any results concerning these broad Fe II features.

The composites in Fig. 1 (and subsequent figures) are plotted when at least 10 individual QSOs contribute to the spectrum. It can be seen that as the number of QSOs is reduced the S/N

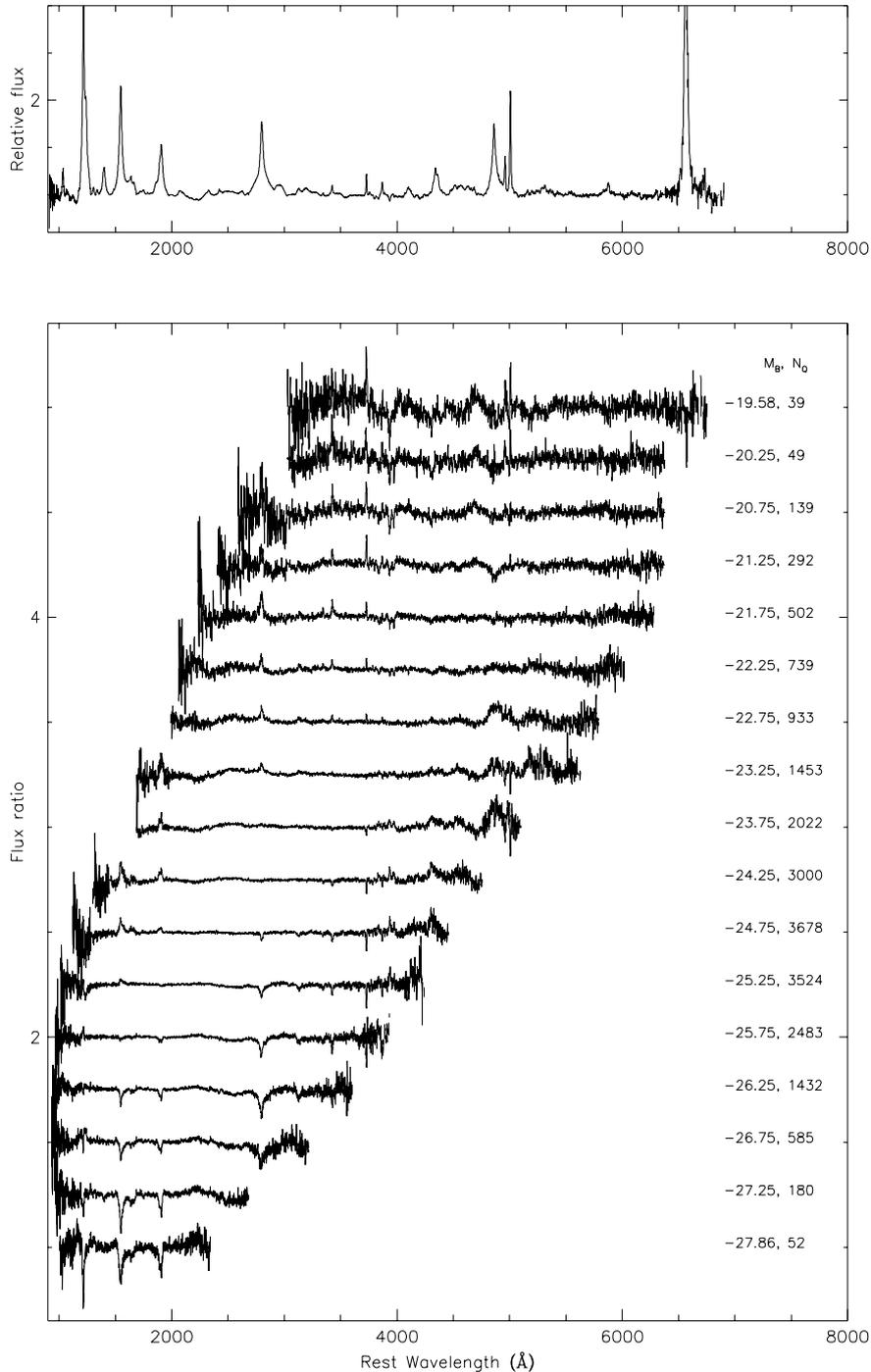


Figure 2. The luminosity-segregated composite spectra, as shown in Fig. 1, divided by the average composite (top). From this apparent correlations can be seen between luminosity and line strength in a number of lines, including [O II], [Ne V], Mg II, C III] and C IV. These correlations are discussed in the text. Again, the mean flux ratio in each spectrum is one, but the spectra have been offset for clarity.

ratio declines. From this plot a number of trends can already be seen, with the narrow [Ne V], [O II] and [Ne III] showing an anticorrelation of line strength with luminosity. The broad emission lines of C IV, C III] and Mg II also show a similar correlation, appearing to confirm previous detections of the Baldwin effect. A further graphical representation of these (and other) correlations is shown in Fig. 2, which shows the luminosity-segregated composites divided by the mean composite. This confirms that anticorrelations with luminosity are seen for a wide variety of emission lines. An anticorrelation

is also seen between the strength of the Ca II H and K absorption lines and QSO luminosity, consistent with a picture where the host galaxy luminosity of QSOs is only weakly correlated with QSO luminosity. Finally, we note that the Balmer series (in particular H β and H γ) appears to show a positive correlation with luminosity, in contrast to the other emission lines. We will analyse these apparent correlations in a quantitative manner below.

In Fig. 3 we show examples of the composites divided into absolute magnitude and redshift bins. These allow us to decouple the

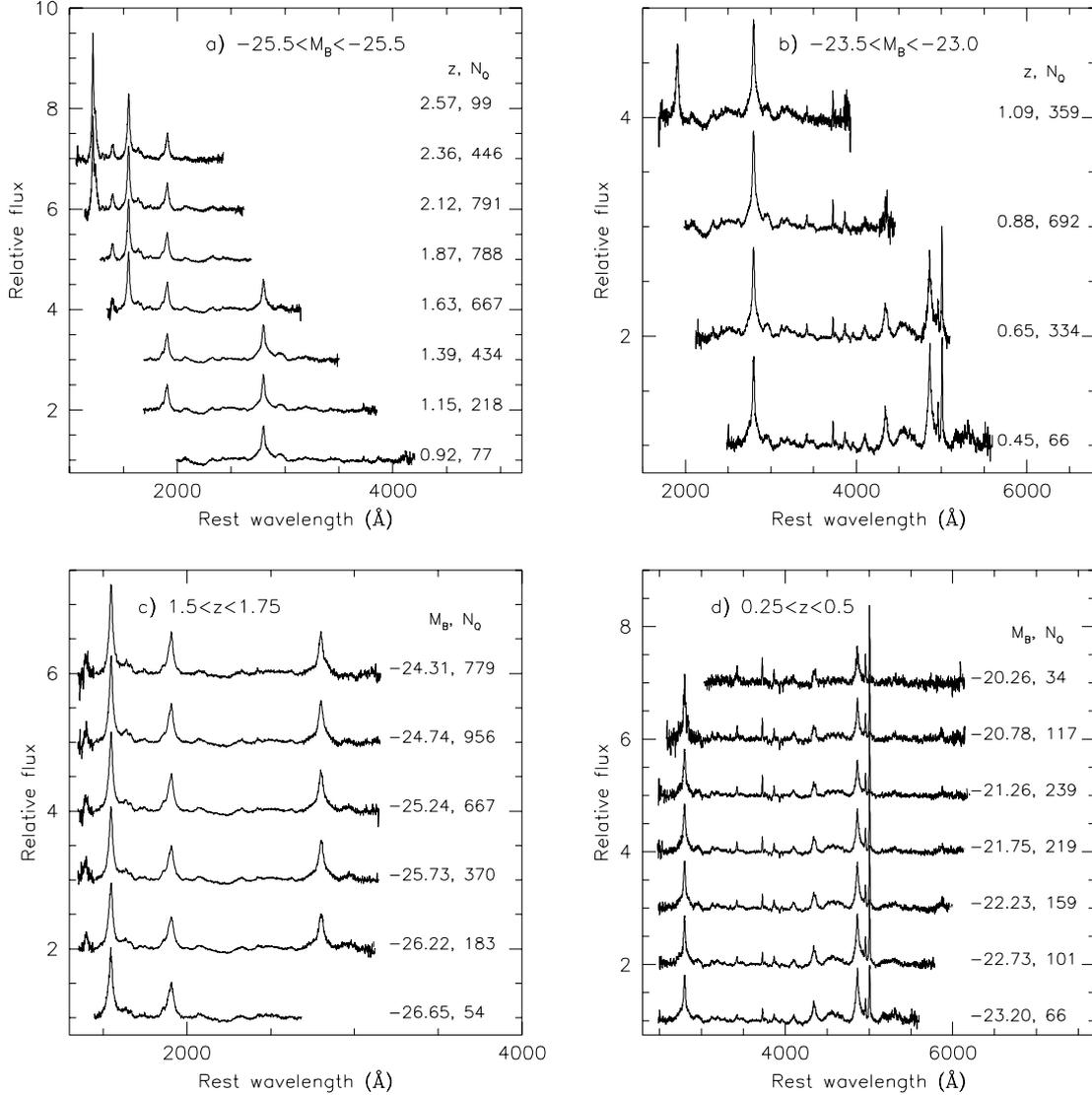


Figure 3. Examples of the QSO composites generated in absolute magnitude–redshift intervals. Top: constant luminosity examples ($-25.5 < M_B < -25.0$ and $-23.5 < M_B < -23.0$) over a range of redshifts. Bottom: constant redshift intervals as a function of luminosity for $1.5 < z < 1.75$ and $0.25 < z < 0.5$.

effects of redshift and luminosity. Figs 3(a) and (b) show composite spectra with a fixed luminosity over a range of redshifts. There is no obvious evidence for emission features varying with redshift in these plots. Figs 3(c) and (d) show composites in a fixed redshift interval over a range in luminosity. In this case we do see an apparent correlation between luminosity and some lines (C IV, [Ne v], [O II], [Ne III], Ca II K), with the lines in question becoming weaker with increasing luminosity. To investigate the nature of these correlations, in particular whether they are primarily a function of luminosity or redshift, we will carry out detailed fitting of the spectral features, followed by a correlation analysis.

3.1 Line fitting procedure

The composite spectra exhibit a number of spectral features, both in emission and absorption, which their high S/N ratio allow us to fit. Twelve of these features, including three narrow (forbidden) lines, seven broad (permitted) emission lines, one semiforbidden line (C III) and one absorption feature (Ca II K), were selected for

detailed study. These features were chosen because they exhibit large equivalent widths (e.g. Ly α , C IV) and, in the case of the narrow emission lines, are relatively free from contamination by other emission lines.

The local pseudo-continuum on either side of each spectral feature was fitted with a straight line, using a linear least-squares method, and subtracted from the spectrum. This continuum was by no means the ‘true’ continuum as the emission lines in QSOs often lie on top of other emission lines, in particular broad Fe II features. It was, however, relatively flat and close to the feature of interest. The majority of the strong emission lines in QSO spectra are blended with other, weaker, emission lines, usually from different elements. Additionally, permitted emission lines such as the Balmer series often exhibit both a broad and a narrow component, which are emitted from physically distinct regions. It is therefore necessary to model and remove the contribution from these different lines to obtain an accurate measurement of the linewidths and equivalent widths.

The overlapping lines contributing to each spectral feature were modelled using multicomponent Gaussian fits as listed in Table 3.

Table 3. Spectral features studied. The first column gives the principal emission line in each feature for which line equivalent widths were measured. Columns 2 and 3 give the regions of the spectrum used in the continuum fit. Columns 4 to 10 give the properties of the individual fitted components. Column 4 lists the component number, column 5 the element/elements causing the emission, column 6 the laboratory (vacuum) wavelengths of the components and column 7 indicates whether the emission is narrow or broad. Columns 8, 9 and 10 list the fitted parameters, showing which parameters were tied together.

Principal line	Blue cont. (Å)	Red cont. (Å)	Component number	Emission source	λ_{lab} (Å)	Emission type	λ_c	Amp.	σ
Ly α	1135–1155	1320–1340	1	Ly α	1215.67	Broad	λ_1	a_1	σ_1
			2	Ly α	1215.67	Broad	λ_2	a_2	σ_2
			3	N v	1240.14	Broad	λ_1	a_3	σ_1
			4	N v	1240.14	Broad	λ_1	a_4	σ_2
Si IV+O IV]	1350–1365	1440–1455	1	Si IV ¹	1396.76	Narrow	λ_1	a_1	σ_1
			2	Si IV ¹	1396.76	Broad	λ_2	a_2	σ_2
			3	Si IV ¹	1396.76	Broad	λ_2	a_3	σ_3
C IV	1440–1460	1690–1710	1	C IV ¹	1549.06	Narrow	λ_1	a_1	σ_1
			2	C IV ¹	1549.06	Broad	λ_2	a_2	σ_2
			3	C IV ¹	1549.06	Broad	λ_2	a_3	σ_3
			4	He II ²	1640.42	Broad	λ_4	a_4	σ_4
C III]	1800–1820	1975–1995	1	C III]	1908.73	Narrow	λ_1	a_1	σ_1
			2	C III]	1908.73	Broad	λ_2	a_2	σ_2
			3	C III]	1908.73	Broad	λ_2	a_3	σ_3
			4	Al III	1857.40	Broad	λ_2	a_4	σ_2
			5	Si III]	1892.03	Broad	λ_2	a_5	σ_2
Mg II	2640–2660	3030–3050	1	Mg II ¹	2798.75	Broad	λ_1	a_1	σ_1
			2	Mg II ¹	2798.75	Broad	λ_2	a_2	σ_2
			3	Fe II blend	2965	Broad	λ_3	a_3	σ_3
[Ne v]	3360–3380	3450–3470	1	[Ne v]	3426.84	Narrow	λ_1	a_1	σ_1
			2	Fe II?	3415	Narrow	λ_1	a_1	σ_2
[O II]	3700–3710	3742–3752	1	[O II] ¹	3728.48	Narrow	λ_1	a_1	σ_1
[Ne III]	3845–3850	3910–3915	1	[Ne III]	3869.85	Narrow	λ_1	a_1	σ_1
			2	He I	3889.74	Broad	λ_2	a_2	σ_2
Ca II K	3900–3910	4010–4020	1	Ca II K	3934.78	Broad	λ_1	a_1	σ_1
			2	Ca II H	3969.59	Broad	λ_1	a_1	σ_1
			3	[Ne III] ³	3968.58	Narrow	λ_3	a_3	σ_3
H δ	4000–4020	4200–4220	1	H δ	4102.89	Narrow	λ_1	a_1	σ_1
			2	H δ	4102.89	Broad	λ_2	a_2	σ_2
H γ	4200–4220	4440–4460	1	H γ	4341.68	Narrow	λ_1	a_1	σ_1^{\dagger}
			2	H γ	4341.68	Broad	λ_2	a_2	σ_2
			3	[O III]+[Fe II]	4361.62	Narrow	λ_2	a_3	σ_1^{\dagger}
H β	4740–4760	5070–5090	1	H β	4862.68	Narrow	λ_1	a_1	σ_1
			2	H β	4862.68	Broad	λ_2	a_2	σ_2
			3	[O III]	4960.30	Narrow	λ_1	a_3	σ_1
			4	[O III]	5008.24	Narrow	λ_1	a_4	σ_1

Notes: 1. Unresolved doublets or multiplets, mean wavelength quoted. 2. Although this feature is identified here as He II it is actually a blend of several lines including Fe II and O III] and is therefore relatively broad. 3. The [Ne III] λ 3968 feature is also contaminated by the He I λ 3970 which may be present in emission or absorption. 4. When both H γ and H β were present in the spectrum, the velocity width of these components were fixed to that measured for the [O III] λ 5007 emission line.

We note, however, that assuming a Gaussian form for the features in our spectra may be a gross oversimplification, and future work will endeavour to define non-parametric measurements of line properties as well as these Gaussian fits. Each component was fitted with a Gaussian of the form

$$F(\lambda) = a \exp\left[-\frac{(\lambda_c - \lambda)^2}{\sigma^2}\right], \quad (1)$$

where a is the peak emission, λ_c is the wavelength of the peak emission and σ is the width of the line. When possible, the number of independent parameters in the model was reduced by linking some of them together. For example, since the [O III] λ 5007, 4959 and narrow H β emission arises from the same region of the QSO (the narrow-line region) it is reasonable to assume that the emitting

gas will have similar velocity shifts and dispersions. The central wavelengths and linewidths of the [O III] λ 4959 and narrow H β emission were therefore tied to those of the [O III] λ 5007. Columns 8–10 in Table 3 show how features were tied together. For example, all the components of the broad H β line are free, while the line centres (λ_c) and widths (σ) for the two [O III] lines and the narrow H β line are tied together.

The narrow emission lines were modelled as single Gaussians and were restricted to velocity widths $< 1500 \text{ km s}^{-1}$. Adequate fits to the Ca II absorption feature and the broad Balmer emission lines (H β , H γ and H δ) were also obtained using a single Gaussian, although there is some evidence (see Fig. 4) from these high S/N ratio spectra that the broad H β has an asymmetric non-Gaussian profile. The broad ultraviolet (UV) lines, i.e. from Mg II λ 2798 blueward,

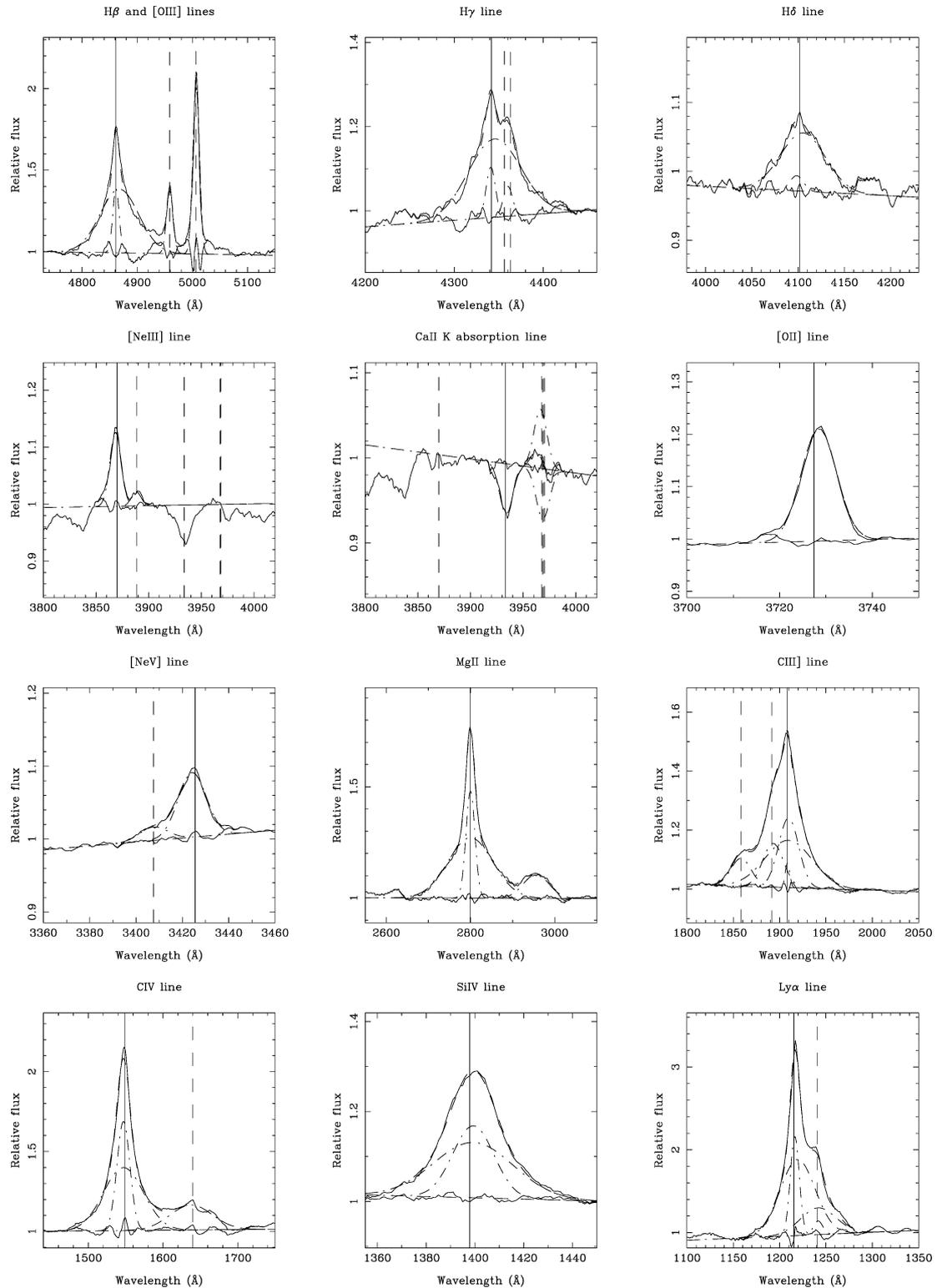


Figure 4. Example line fits. We plot each line or complex of lines fitted in our analysis for the case of the total composite spectrum. Shown is the data (solid line), the individual Gaussian components (dot-dashed lines), the sum of the Gaussian components (dashed line) and the residuals after subtracting the fit (solid line). In most cases the dashed line denoting the total fit is hardly visible over the data. The vertical lines indicate the wavelengths of the primary component (solid line) and secondary contaminating components (dashed lines).

display emission-line profiles with very broad bases, which cannot be adequately modelled by a single Gaussian. They were therefore fitted with two components; a very broad Gaussian (FWHM ~ 10000 km s $^{-1}$) and a narrower component (FWHM ~ 2000 – 4000 km s $^{-1}$). This narrower component is not believed to be emission from the narrow-line region since its velocity dispersion is much larger than that measured in the narrow lines (e.g. [O II] $\lambda 3727$ and [O III] $\lambda 5007$) which is typically ~ 800 km s $^{-1}$. When two broad components were fitted to a broad emission line, the central wavelengths of the Gaussians were tied together as we found no evidence for a velocity shift between the components. Systematic shifts between different components, have been seen by other authors (Brotherton et al. 1994b), and potential line shifts will be investigated in detail by Corbett et al., in preparation. The only line in which the broad components were not tied was Ly α , because absorption to the blue side of the line resulted in an asymmetric profile and it was necessary to allow a velocity shift between the two components to fit the line profile. Previous studies (Wills et al. 1993; Brotherton et al. 1994a,b) have highlighted the fact that the broad-line region can be well described by two components, often described as the intermediate line region and the very broad-line region. It is clear that the broad UV lines in our composite spectra show these two components, however, we reserve a detailed discussion of line shapes for the forthcoming paper, Corbett et al. In all cases the best fit to the spectral feature was found using χ^2 minimization techniques.

The broad Balmer emission line H γ proved difficult to de-blend as it is contaminated by emission from both [Fe II] $\lambda 4358$ and [O III] $\lambda 4363$ as well as narrow H γ emission. Since the [Fe II] and [O III] emission are within 6 Å of each other they are not resolved in the 2dF spectra and were therefore modelled as single narrow component centred between the two lines. The fit was further constrained by fixing the velocity width of the narrow H γ and the combined [Fe II] and [O III] lines to that obtained for the [O III] $\lambda 5007$ emission. It was not possible to de-blend the O IV] $\sim \lambda 1402$ multiplet emission from the Si IV $\lambda \lambda 1393, 1402$ emission and hence the equivalent width calculated for Si IV also contains emission from O IV].

Once the spectral feature had been modelled, the fits to the contaminating line emission were subtracted, leaving only the line of interest. The total flux in the line was measured by integrating the flux over a wavelength range defined as $\lambda_c \pm 1.5 \times \text{FWHM}$, where λ_c is the central wavelength and the FWHM is that of the broadest Gaussian component fitted to the line. The equivalent width of the emission (or absorption) was defined as

$$W_\lambda = \frac{F_{\text{line}}}{F_{\text{cont}}} \text{Å}, \quad (2)$$

where F_{line} is the integrated flux in the emission/absorption line and F_{cont} is the continuum flux measured in a 1-Å bin about the central wavelength of the fit to the line. By using the integrated residual flux in the spectral feature rather than the Gaussian fit to calculate F_{line} , we avoid introducing errors caused by the fact that the line emission may not be perfectly fitted by a Gaussian (e.g. H β). There is, however, an uncertainty in F_{line} caused by the modelling and subtraction of the contaminating line emission and continuum, which we have taken into account when calculating the errors in W_λ . In general, the multicomponent fits to the C III] line were degenerate, and so we also calculate the equivalent width of the total spectral feature, which is used in the analysis below.

3.2 Correlation analysis

Once the widths, equivalent widths and line centres were measured for all lines fitted above, we tested the data for correlations between

the line parameters and redshift and luminosity. In this paper we report the results for W_λ . Discussion of linewidths and centres will be reported elsewhere. We have carried out non-parametric rank correlation analysis, deriving the Spearman rank-order correlation coefficient, ρ . We select a priori 99 per cent to be the confidence level at which we will claim significant correlations. Specifically, we will test for an W_λ - z correlation by correlating $\log W_\lambda$ with $\log(1+z)$, as evolutionary parameters for QSOs are generally an approximate power law in $(1+z)$, in particular, QSO luminosity evolution (Boyle et al. 2000). In testing for W_λ - M_B correlations we will correlate $\log W_\lambda$ with M_B .

A particularly important issue is to deduce whether z or M_B is the primary parameter with which W_λ correlates. We approach this problem in two ways: the first is to carry out correlations in separate z or M_B intervals, removing any possible spurious correlations with the second independent variable. The second approach is to use partial Spearman rank correlation (e.g. Macklin 1982) to derive the correlation coefficient while holding one independent variable constant:

$$\rho_{AX,Y} = \frac{\rho_{AX} - \rho_{XY}\rho_{AY}}{\sqrt{(1 - \rho_{XY}^2)(1 - \rho_{AY}^2)}}, \quad (3)$$

where X and Y are two independent variables (e.g. z and M_B) and A is the dependent variable (e.g. W_λ). ρ_{AX} , ρ_{AY} and ρ_{XY} are the Spearman correlation coefficients for the separate correlations between two variables. The significance of $\rho_{AX,Y}$ is given by

$$D_{AX,Y} = \frac{\sqrt{N-4}}{2} \ln \left(\frac{1 + \rho_{AX,Y}}{1 - \rho_{AX,Y}} \right), \quad (4)$$

which is distributed normally about zero with unit variance (Macklin 1982), where N is the size of the sample. In using this partial rank correlation approach we are testing the null hypotheses that: (i) the W_λ - z correlation arises entirely from the W_λ - M_B and M_B - z correlations and (ii) the W_λ - M_B correlation arises entirely from the W_λ - z and M_B - z correlations. If the coefficients for the W_λ - z correlation are larger than those for the W_λ - M_B correlation, this would imply that W_λ is *primarily* correlated with z .

To determine the slope of any measured correlations we also carry out fits to the data using the non-linear Levenberg–Marquardt method. This method was used in order to fit a power law, while still properly taking into account the errors on the W_λ measurements (which is not possible in a standard linear least-squares approach). As will be seen below, the errors in the individual W_λ measurements were often much smaller than the scatter about the best-fitting line. This could reflect the fact that parameters other than those fitted are introducing extra dispersion and/or the power law is not an adequate fit to the data. To obtain a realistic error on the fitted parameters, we repeat the fitting procedure, rescaling the errors such that the reduced χ^2 is exactly one in each case, noting the specific cases where the data diverge significantly from a power law.

4 RESULTS

We derive both the bivariate and partial Spearman rank correlation coefficients for the correlation of W_λ with z and M_B . Our primary aims in doing this are to (i) test for the existence of any significant correlations and (ii) determine whether those correlations are primarily with z or M_B . This second point is important given the recent claim (Green et al. 2001) that the Baldwin effect is primarily a correlation with redshift.

4.1 Bivariate and partial Spearman rank correlation

We select the strongest and cleanest spectral features to test for correlations. In cases where there is significant contamination by other features the Gaussian fits for these have been subtracted off the summed flux to provide a clean estimate of line flux. This has not been done in a few cases, where the separate components cannot be clearly distinguished. In particular, we use the summed flux of all components in the Si IV+O IV complex, and do not subtract off narrow components from H δ or H γ (including [O III] λ 4363). Note, however, that we combine together broad and intermediate components of the same line (in particular, for all the broad UV lines).

We carry out the correlations first for $\log(1+z)$ and M_B . The resulting correlation coefficients are listed in Table 4. The number of points used in the correlations ranges from 17 to 49 with a median of 22. Table 4 first lists the full bivariate correlation coefficients and probabilities for the $\log W_\lambda$ correlations with M_B and $\log(1+z)$. These do not take into account any potential spurious correlation caused by the correlation of M_B and $\log(1+z)$. We find that a number of lines show significant ($P < 0.01$) correlations. The C IV, C III]+Al III, Mg II, [Ne V], [O II], Ca II K, H γ and H β lines all show significant correlations with M_B . We find that fewer lines, only Ly α , C III]+Al III, [Ne V] and [O II], show correlations with $\log(1+z)$. The data (filled and open circles) and best-fitting correlations (solid lines) are shown in Figs 5 and 6.

We then derive the partial Spearman rank correlation coefficients of $\log W_\lambda$ with M_B and $\log(1+z)$, which are listed in the last four columns of Table 4. In all but two cases (Ly α , C III]+Al III) the significance of the correlation is larger for M_B than $\log(1+z)$. The strongest partial correlations with $\log(1+z)$ are for [O II] and Ca II K, which are also the lines showing the steepest correlations with M_B . This is consistent with these $\log(1+z)$ correlations being a result of the luminosity distribution of the QSOs, *within* a given luminosity range, changing with redshift. Our partial correlation analysis demonstrates that the correlations seen are primarily with M_B rather than redshift, in disagreement with the previous results of Green et al. (2001).

4.2 The correlation of W_λ and redshift

To further investigate the finding that W_λ primarily correlates with M_B and not z we now use the separate luminosity intervals of width $\Delta M_B = 0.5$ to search for correlations with redshift. In a given luminosity interval there are up to eight $\Delta z = 0.25$ redshift intervals sampled. For each measured emission line we test for a correlation between W_λ and $\log(1+z)$ independently within each luminosity interval using Spearman rank correlation. In some luminosity intervals there may be only a small number of redshift intervals in which a particular line is present. This is particularly true for lines such as H β and [O III] near the edge of the spectrum, which are only present in three redshift intervals. In the Spearman rank correlation analysis we limit ourselves to examining luminosity intervals that contain at least five separate measurements. This is because the significance of ρ is derived from $t = \rho \sqrt{(N-2)/(1-\rho^2)}$, which is approximately distributed as a Student's distribution. However, this breaks down for small N , as it predicts zero probability for $\rho = \pm 1$, whereas the true likelihood of this occurring is $(2/N!)$. Only the Si IV, C IV, C III]+Al III and Mg II lines have five or more measured equivalent widths in a given luminosity interval, hence only these lines are sensitive to tests for $\log W_\lambda$ versus $\log(1+z)$ correlations in each luminosity interval. No significant correlations are found for any of these lines, supporting the above finding that the correlations are primarily driven by luminosity. Fig. 5 shows the distribution $\log W_\lambda$ versus $\log(1+z)$ for all the lines. The symbols (circles) are larger for brighter luminosity intervals, and in a number of cases (e.g. [O II]) we see that the correlation could potentially be a result of intrinsically fainter (small circles) sources having larger equivalent widths. The above partial correlation analysis confirms this impression.

4.3 The correlation of W_λ and M_B

We now correlate W_λ with M_B in separate redshift intervals of $\Delta z = 0.25$. Again, only intervals with five or more W_λ measurements are tested for correlations, however, each line has at least one

Table 4. Spearman rank correlation coefficients for correlations of $\log(W_\lambda)$ with M_B and $\log(1+z)$. For each line tested we give the number of points correlated, N , the Spearman rank coefficient, ρ and the probability of the null hypothesis, P . Full, bivariate coefficients and probabilities are given first for the correlations with M_B and $\log(1+z)$, then we list the partial correlation values.

Line	N	Bivariate correlations				Partial correlations			
		$\log W_\lambda$ versus M_B		$\log W_\lambda$ versus $\log(1+z)$		$\log W_\lambda$ versus M_B		$\log W_\lambda$ versus $\log(1+z)$	
Ly α	19	-0.476	3.956E-02	0.709	6.757E-04	-0.209	4.121E-01	0.621	4.908E-03
N V	19	-0.119	6.265E-01	-0.059	8.110E-01	-0.170	5.055E-01	-0.135	5.975E-01
Si IV	30	0.184	3.303E-01	-0.134	4.811E-01	0.140	4.715E-01	-0.059	7.642E-01
C IV	34	0.816	4.058E-09	-0.281	1.080E-01	0.813	4.746E-10	0.258	1.482E-01
C III]+Al III	49	0.574	1.650E-05	-0.646	5.436E-07	0.313	2.987E-02	-0.465	7.267E-04
Mg II	35	0.493	2.598E-03	-0.178	3.058E-01	0.626	4.236E-05	0.471	4.369E-03
[Ne V]	23	0.899	5.625E-09	-0.720	1.060E-04	0.803	1.418E-06	0.330	1.355E-01
[O II]	30	0.913	2.153E-12	-0.634	1.670E-04	0.917	1.119E-15	0.658	5.756E-05
[Ne III]	24	0.334	1.112E-01	-0.049	8.196E-01	0.598	2.052E-03	0.528	8.692E-03
H δ	22	-0.240	2.821E-01	-0.008	9.702E-01	-0.504	1.862E-02	-0.457	3.648E-02
H γ	22	-0.634	1.529E-03	0.298	1.786E-01	-0.756	2.876E-05	-0.588	4.189E-03
H β	17	-0.711	1.382E-03	0.118	6.507E-01	-0.897	1.504E-07	-0.781	1.553E-04
[O III]	17	-0.179	4.920E-01	-0.012	9.628E-01	-0.265	3.280E-01	-0.199	4.675E-01
Ca II K	17	0.866	7.118E-06	-0.603	1.041E-02	0.939	4.146E-10	0.838	1.226E-05

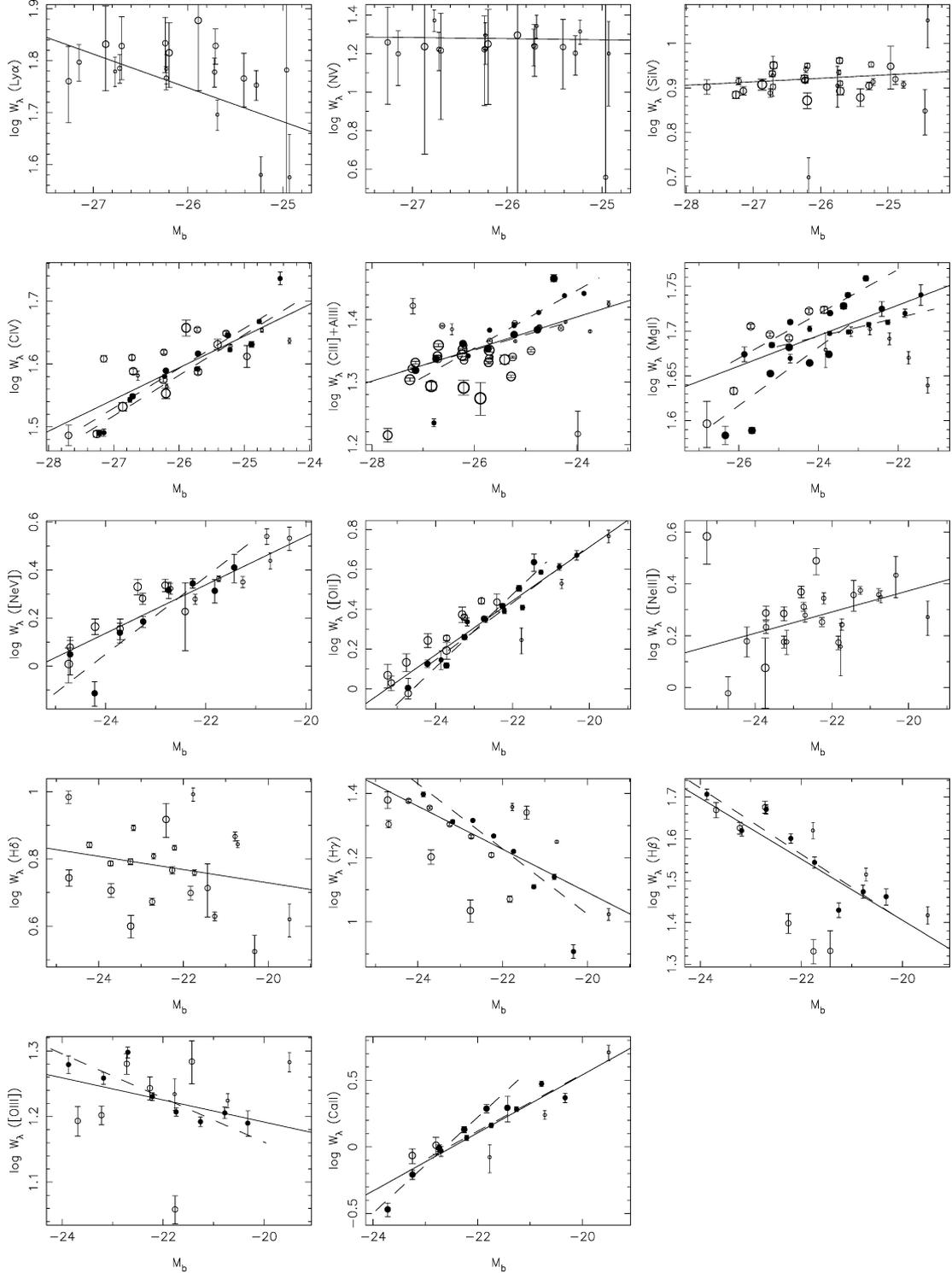


Figure 6. The correlations of $\log W_{\lambda}$ with M_B for each line. The circles denote the measured values of $\log W_{\lambda}$, with larger circles indicating higher-redshift intervals. In cases where an individual luminosity interval shows a significant correlation with redshift the circles are filled and the best-fitting power-law correlation is plotted (dashed lines). For every feature, regardless of whether a significant correlation is seen or not, the best power-law fit to all the data points is also shown (solid line).

particular redshift e.g. Mg II) would probably give rise to the opposite effect i.e. a tendency to overestimate the mean equivalent width at the faint magnitudes (lowest luminosities) in the sample.

Indeed, some Malmquist bias may be present in the data, although we have confirmed that the results presented here are robust against

the inclusion/exclusion of the faintest 0.5 mag interval in absolute magnitude. Moreover, the fact that a range of slopes are found for the correlation between equivalent width and M_B (both positive and negative) further suggests that any Malmquist bias plays a small role.

Table 5. Best-fitting linear correlations with Spearman rank correlation coefficients and probabilities. First we list the best fits for all redshift intervals combined. Then we list correlations from those lines which show significant correlations in individual redshift intervals. N is the number of points used in the correlation analysis. A and B are the intercept and gradient for the best-fitting line, such that $\log W_\lambda = A + BM_B$. δA and δB are the errors on the intercept and slope. ρ and P are the Spearman rank coefficient and probability, respectively. Probabilities marked by an asterisk (*) have $\rho = 1$ (a perfect correlation). These probabilities have been corrected to $P = (2/N)!$.

Line	N	Redshift	A	δA	B	δB	ρ	P
Ly α	19	0.00–4.00	0.059	0.531	−0.065	0.020	−0.476	3.956E−02
N v	19	0.00–4.00	1.124	1.004	−0.006	0.039	−0.119	6.265E−01
Si iv	30	0.00–4.00	1.127	0.173	0.008	0.007	0.184	3.303E−01
C iv	34	0.00–4.00	2.905	0.154	0.050	0.006	0.816	4.058E−09
C III]+Al III	49	0.00–4.00	2.021	0.097	0.026	0.004	0.574	1.650E−05
Mg II	35	0.00–4.00	2.103	0.109	0.017	0.005	0.493	2.598E−03
[Ne v]	23	0.00–4.00	2.557	0.263	0.101	0.012	0.899	5.625E−09
[O II]	30	0.00–4.00	3.409	0.222	0.135	0.010	0.913	2.153E−12
[Ne III]	24	0.00–4.00	1.192	0.325	0.041	0.015	0.334	1.112E−01
H δ	22	0.00–4.00	0.334	0.370	−0.020	0.016	−0.240	2.821E−01
H γ	22	0.00–4.00	−0.249	0.257	−0.067	0.011	−0.634	1.529E−03
H β	17	0.00–4.00	−0.053	0.353	−0.073	0.016	−0.711	1.382E−03
[O III]	17	0.00–4.00	0.854	0.235	−0.017	0.011	−0.179	4.920E−01
Ca II K	17	0.00–4.00	4.895	0.488	0.218	0.022	0.866	7.118E−06
C iv	7	1.75–2.00	3.260	0.196	0.065	0.008	1.000	3.968E−04*
C iv	6	2.00–2.25	3.240	0.248	0.063	0.010	0.943	4.805E−03
C III]+Al III	7	1.25–1.50	2.556	0.163	0.046	0.007	1.000	3.968E−04*
C III]+Al III	7	1.75–2.00	1.984	0.174	0.024	0.007	0.964	4.541E−04
Mg II	8	0.50–0.75	1.948	0.072	0.011	0.003	0.929	8.630E−04
Mg II	7	0.75–1.00	2.314	0.104	0.025	0.004	0.893	6.807E−03
Mg II	7	1.00–1.25	2.461	0.314	0.032	0.013	0.893	6.807E−03
[Ne v]	8	0.50–0.75	3.915	0.621	0.161	0.027	0.905	2.008E−03
[O II]	8	0.25–0.50	3.574	0.510	0.143	0.023	1.000	4.960E−05*
[O II]	8	0.50–0.75	4.500	0.329	0.183	0.014	0.976	3.314E−05
H γ	8	0.25–0.50	−1.006	0.343	−0.102	0.015	−0.952	2.604E−04
H β	8	0.25–0.50	−0.178	0.319	−0.079	0.014	−0.905	2.008E−03
[O III]	8	0.25–0.50	0.487	0.205	−0.034	0.009	−0.905	2.008E−03
Ca II K	6	0.25–0.50	4.823	0.794	0.214	0.037	0.943	4.805E−03
Ca II K	6	0.50–0.75	8.139	0.651	0.360	0.029	1.000	2.778E−03*

In Table 5 we list the parameters of all the significant correlations, including their significance and best-fitting parameters for the fit to $\log W_\lambda = A + BM_B$. From Table 5 we can see that there are significant differences between gradients of the different lines. The strongest correlation is found in the Ca II K line with a gradient of 0.218 ± 0.022 . The Balmer lines are the only ones to show a negative correlation with M_B (a positive correlation with luminosity), which confirms the visual impression gained from Fig. 2. We discuss the physical significance of these correlations below.

4.4 The correlation of W_λ with M_B in luminosity only divided composites

Given that the above analysis appears to suggest that the dominant correlation is with M_B we now derive correlations between $\log W_\lambda$ with M_B for the composites subdivided only on the basis of luminosity (Fig. 1). This can potentially reduce the noise and scatter in the correlations if W_λ is truly only correlated with M_B . The resulting correlations are shown in Fig. 7. The correlation coefficients and best-fitting parameters are listed in Table 6. We see evidence of significant correlations in many lines, with some exceptions. The lines that do not show correlations are Ly α , N v, Si iv, [Ne III], H δ and [O III]. We also note that in some cases, most notably C iv, C III]+Al III and Mg II, the dispersion about the best-fitting correlation is much larger than would be expected given the errors on

individual points. This suggests that other parameters may cause extra dispersion in the relation, or that a simple power-law fit is not actually a good description of the underlying physics.

5 DISCUSSION

We now attempt to understand the above measured correlations in the context of simple physical models for AGN emission and host galaxy properties. We will start by considering the host galaxy, and then look at the narrow- and broad-line regions in turn.

5.1 Host galaxy properties

The Ca II K absorption line is possibly the simplest to interpret, as it can only be caused by the stars present in the host galaxy of the QSO. We clearly see that as the AGN luminosity increases, the strength of the Ca II K declines, consistent with a picture in which the host galaxy does not increase in luminosity as fast as does the AGN. If the host galaxy was constant in luminosity, the slope of the correlation between $\log W_\lambda$ and M_B would be 0.4. Our best-fitting slope is 0.208 ± 0.023 , which is therefore consistent with the host galaxy luminosity increasing slowly. In this analysis we make the assumption that the average spectral properties of the AGN host galaxies do not change significantly with luminosity (or redshift). We also assume that there is no significant aperture effect introduced by the 2-arcsec

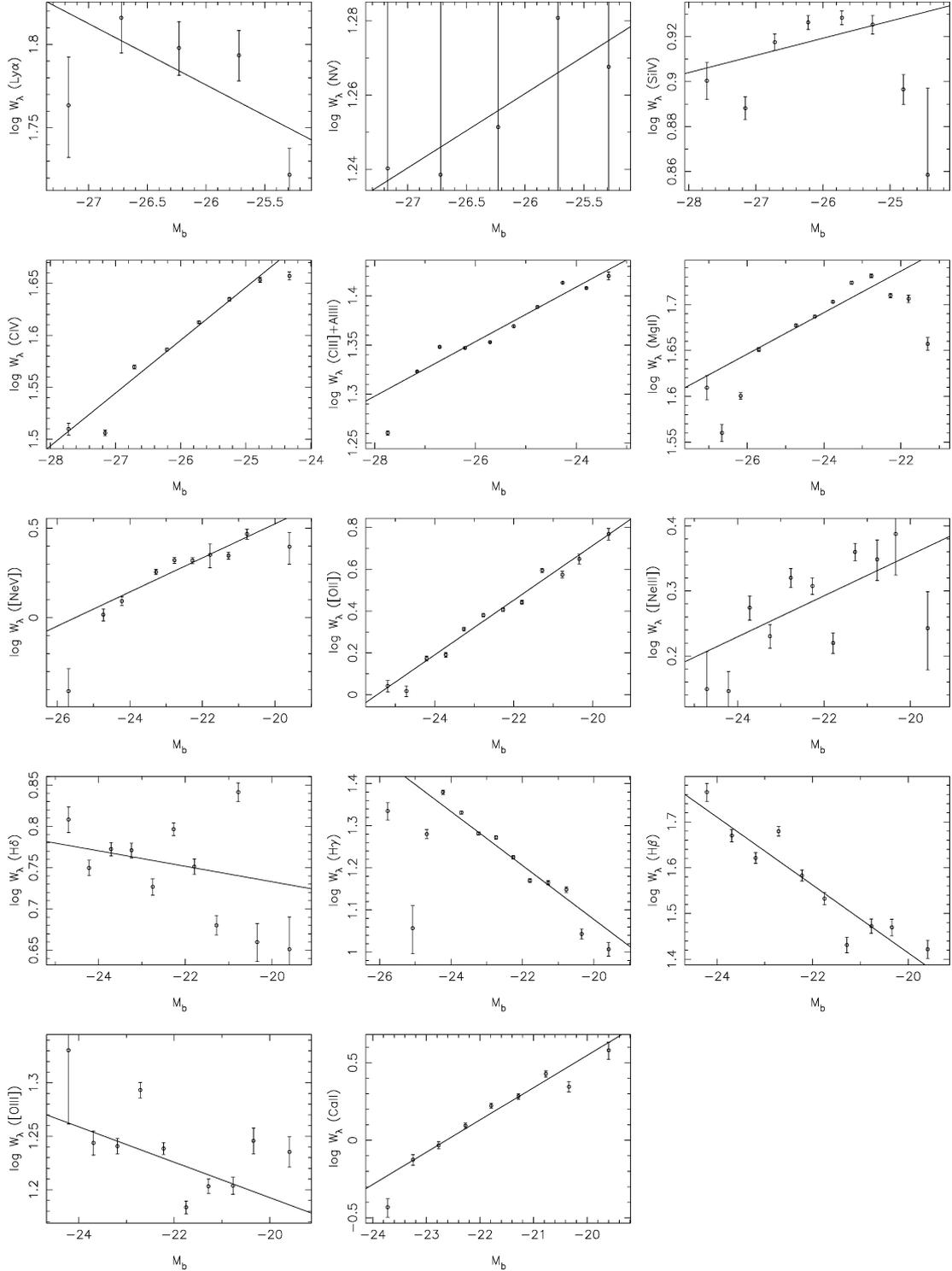


Figure 7. The correlations of $\log W_\lambda$ with M_B for each line, measured from composites subdivided by luminosity only. The best-fitting correlation is plotted in each case.

diameter of the 2dF fibres used to obtain the spectra. Based on the imaging results of Schade, Boyle & Letawsky (2000), we expect that galaxy hosts for QSOs in this luminosity range ($-21 > M_B > -24$) will be bulge-dominated with effective radii ranging from 1 to 2 kpc for QSOs at $z \sim 0.15$ and 3–6 kpc for QSOs with $z \sim 0.6$. In both cases the projected size of the bulges are approximately the same

size on the sky (~ 0.75 – 1.5 -arcsec diameter) and significantly less than the fibre diameter.

If we also assume that the majority of the continuum emission is caused by the QSO, we can derive a simple relation for the expected correlation between W_λ and M_B . We can set $L_{\text{line}} \propto L_{\text{gal}} \propto L_{\text{QSO}}^\alpha$, i.e. the QSO luminosity is proportional to the host galaxy luminosity to

Table 6. Best-fitting linear correlations with Spearman rank correlation coefficients and probabilities for composites subdivided by luminosity only. N is the number of points used in the correlation analysis. A and B are the intercept and gradient for the best-fitting line, such that $\log W_\lambda = A + BM_B$. δA and δB are the errors on the intercept and slope. ρ and P are the Spearman rank coefficient and probability, respectively.

Line	N	A	δA	B	δB	ρ	P
Ly α	5	0.822	0.665	-0.037	0.026	-0.400	5.046E-01
N v	5	1.781	1.118	0.020	0.043	0.800	1.041E-01
Si iv	8	1.118	0.181	0.008	0.007	-0.143	7.358E-01
C iv	8	2.917	0.143	0.051	0.006	0.976	3.314E-05
C III]+Al III	10	2.075	0.080	0.028	0.003	0.976	1.468E-06
Mg II	12	2.233	0.116	0.023	0.005	0.720	8.240E-03
[Ne v]	10	2.423	0.369	0.095	0.016	0.964	7.321E-06
[O II]	12	3.330	0.205	0.131	0.009	0.986	4.117E-09
[Ne III]	11	0.984	0.377	0.031	0.017	0.645	3.196E-02
H δ	11	0.546	0.265	-0.009	0.012	-0.491	1.252E-01
H γ	13	-0.203	0.175	-0.064	0.008	-0.703	7.319E-03
H β	10	-0.066	0.259	-0.074	0.012	-0.927	1.120E-04
[O III]	10	0.862	0.237	-0.017	0.011	-0.527	1.173E-01
Ca II K	9	4.703	0.492	0.208	0.023	0.983	1.936E-06

some power. If $\alpha = 0$ then the host galaxy has a constant luminosity. Converting to magnitude and $\log W_\lambda$ we then find

$$\log W_\lambda = 0.4(1 - \alpha)M_B + \text{constant}. \quad (5)$$

Taking our fitted value for the slope of the correlation we find that $\alpha = 0.48 \pm 0.06$. This then implies that QSO and host galaxy luminosity are correlated. This has, in fact, been found from direct imaging studies of QSO host galaxies. Schade et al. (2000) find $L_{\text{gal}} \propto L_{\text{QSO}}^{0.21}$ for a sample of low-redshift X-ray-selected AGN, with large scatter. This is a somewhat shallower slope than we find, however, we are in fact deriving the slope for the relation $L_{\text{gal}} \propto L_{\text{tot}}^\alpha$, not $L_{\text{gal}} \propto L_{\text{QSO}}$. We can instead construct a model that assumes an exact power-law correlation between L_{gal} and L_{QSO} . However, this does depend on us having some knowledge of the expected equivalent width for Ca II K in the host galaxy, without the AGN component. We cannot make an accurate assessment of the spectral properties of the host galaxy, as we have only fitted one feature. Instead we use a Ca II K line strength derived from the mean galaxy spectrum in the 2dF Galaxy Redshift Survey (Baldry et al. 2002). This has an equivalent width for Ca II K of $7.3 \pm 0.2 \text{ \AA}$. The maximum W_λ found for Ca II K in our analysis is $\sim 4\text{--}5 \text{ \AA}$, thus at the faintest luminosities the host galaxy could be contributing a significant fraction of the continuum. If the equivalent width of the line in the galaxy spectrum is $W_{\lambda, \text{gal}}$, then

$$W_\lambda = \frac{W_{\lambda, \text{gal}}}{L_{\text{tot}}/L_{\text{gal}}} = \frac{W_{\lambda, \text{gal}}}{10^{-0.4(M_{\text{tot}} - M_{\text{gal}})}}, \quad (6)$$

where M_{tot} and M_{gal} are the absolute B -band magnitudes from the total (AGN+host) and host components. Assuming a linear relation between M_{gal} and M_{QSO} such that $M_{\text{gal}} = A + BM_{\text{QSO}}$ implies that

$$M_{\text{gal}} = A - 2.5B \log(10^{-0.4M_{\text{tot}}} - 10^{-0.4M_{\text{gal}}}), \quad (7)$$

which can be solved numerically for M_{gal} , and substituted into equation (6). The resulting best fit is

$$M_{\text{gal}} = (-11.10 \pm 0.97) + (0.417 \pm 0.045)M_{\text{QSO}}, \quad (8)$$

which is shown in Fig. 8 (solid line). This fit is very similar to the previous power-law fit (dotted line), only diverging at faint magnitudes. Also plotted is the relation found by Schade et al. (2000) of

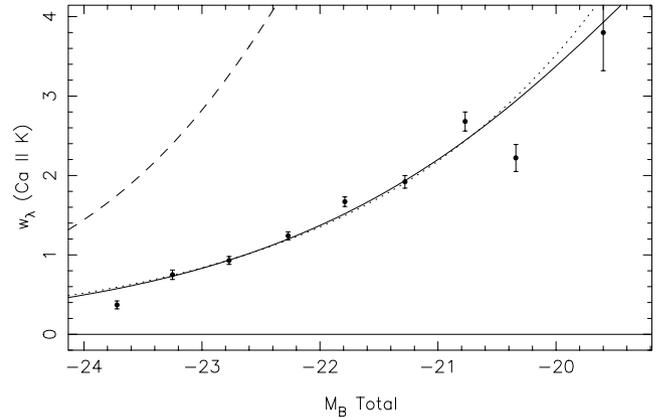


Figure 8. The relationship between $W_\lambda(\text{Ca II K})$ and M_B for composites subdivided by luminosity only. Three models are shown, $M_{\text{gal}} = A + BM_{\text{QSO}}$ (solid line), $M_{\text{gal}} = A + BM_{\text{tot}}$ (dotted line) and $M_{\text{gal}} = A + BM_{\text{QSO}}$ with parameters set from Schade et al. (2000) (dashed line).

$M_{\text{gal}} = -17.25 + 0.21M_{\text{QSO}}$, this is clearly discrepant with our data. Reducing the intrinsic strength of Ca II K to $\sim 4.5 \text{ \AA}$ does not remove this discrepancy. We note that Schade et al. have demonstrated that for low-redshift AGN ($z \sim 0.1$) at about L^* the host galaxies are in almost all respects no different from normal galaxies. The one difference found is a bias towards spheroidal morphologies. This could imply that the stellar populations in AGN host galaxies are older than in average galaxies, but this is by no means certain given that AGN activity could also be accompanied by enhanced star formation.

It is also possible to use the measured Ca II K line together with a mean galaxy spectrum to determine the expected spectral properties of other features not produced by the AGN, in particular, the narrow forbidden oxygen emission lines. In the mean 2dF Galaxy Redshift Survey spectrum the flux emitted in [O III] $\lambda 5007$ is about a factor of 2 lower than the flux absorbed in Ca II K. However, even in our faintest composites where the host galaxies contribute the most, the [O III] line has an W_λ a factor of over 4 greater than Ca II K. Thus, assuming that the host galaxy spectral energy distribution (SED) is similar to that of a normal galaxy implies that the [O III] emission is coming from the AGN, rather than the host galaxy.

The situation regarding the [O II] $\lambda 3727$ line is different. In the mean galaxy spectrum the line has a relative flux that is similar to that in the Ca II K line, and has an equivalent width of $\sim 10.7 \pm 0.3 \text{ \AA}$. Comparing the correlations of Ca II K and [O II] we find that close to the faint end of the distribution, at $M_B = -20$, the [O II] W_λ is 70–100 per cent of what would be predicted from the host galaxy. As the [O II] correlation with W_λ is flatter than that of Ca II K the simple assumption of a constant host galaxy SED would predict that the increased fraction of the [O II] flux is emanating from the AGN at higher luminosities (~ 65 per cent at $M_B = -24$). However, the difference is not as large as in the case of the [O III] line and it is possible to speculate that the [O II] line is formed in powerful star-forming regions and the star formation rate in the host galaxy is increasing with AGN luminosity. In this case, all the [O II] emission could be caused by the host galaxy. The [O II] line would then provide a very useful diagnostic tool for the study of star formation in high-redshift QSOs. A further test of this would be to investigate the velocity distribution of the different narrow lines. If, for example, the [O II] line has a significantly lower velocity dispersion, more consistent with typical galaxies, than [O III], this would be good

evidence for the [O II] being due mostly to the host galaxy. Here we proceed by assuming that a major fraction of the [O II] line emission originates in the NLR of the AGN. The former possibility of the starburst origin will be investigated further by Corbett et al. (in preparation).

5.2 The narrow-line region

The lines that are thought to be emitted within the NLR of the AGN are [O II], [O III], [Ne III] and [Ne V] (see, however, the comment concerning [O II] in the previous section). Of those, only [O II] and [Ne V] show clear and significant correlations with M_B . The partial correlation analysis for both lines is consistent with the hypothesis that the correlation is solely caused by variations in M_B , with no correlation with redshift. The slope of the [Ne V] correlation for the magnitude only divided composites is 0.095 ± 0.016 , consistent with the correlation found in the M_B - z composites, which has a slope of 0.101 ± 0.012 . This line has the highest ionization potential (here and below we consider the lower ionization potential; the energy required to ionize the lower ionization ion) of any of the narrow lines we investigate (126.21 eV). The [O II] line has the lowest ionization potential and the other two have intermediate values: 63.45 eV for [Ne III] and 54.93 eV for [O III]. The [Ne III] line shows only a marginal detection of a correlation (97 per cent significant) and has a significantly flatter slope, 0.031 ± 0.017 . The [O III] line shows no evidence of a correlation and all the measured equivalent widths in the M_B composites are within a range of $\Delta \log W_\lambda = 0.15$. Plotting the measured slope as a function of ionization energy for these narrow lines (open triangles in Fig. 9) we see no obvious trend. However, if much of the [O II] emission is caused by star formation, as suggested in the earlier section, there may be a correlation in a sense that a steeper slope corresponds to higher ionization energy. Given the small number of points and the additional assumption about [O II], we cannot use this trend to infer the NLR physics.

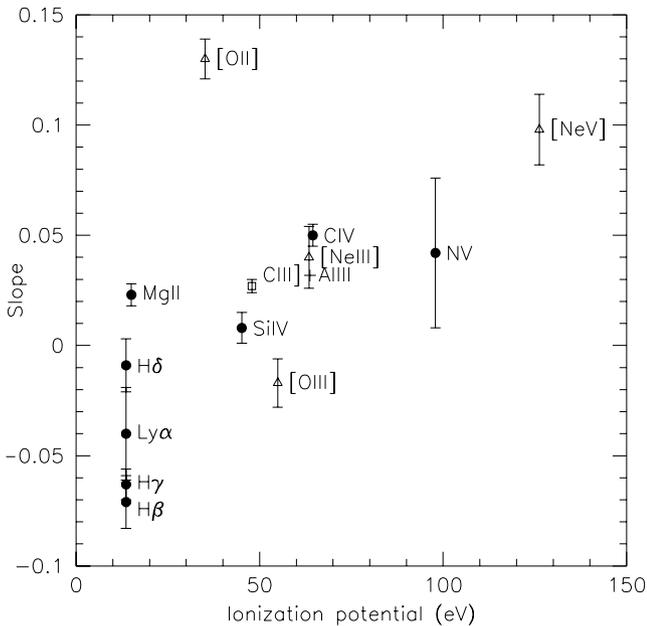


Figure 9. The measured slopes of the $\log W_\lambda$ versus M_B correlation for different lines as a function of lower ionization potential for broad permitted lines (filled circles), narrow forbidden lines (open triangles) and semiforbidden C III] (open square).

Previous analyses have tentatively detected correlations between narrow-line strength and luminosity (Green et al. 2001). However, these were generally in data sets with a large fraction of non-detections, as they used single objects instead of composite spectra. The suggested correlation of line intensity with luminosity in [O III] demonstrated in Fig. 2 does not appear to be borne out in the correlation analysis involving equivalent width measurements. The apparent variation could be caused by a real variation in the velocity width of the [O III] lines while the total flux remains approximately constant.

A simple interpretation of the [Ne V] and [O II] correlation involves the ‘disappearing NLR’ model. This idea is based on the fact that the NLR size scales with the source luminosity to some power

$$R_{\text{NLR}} = R_0(L/L_0)^\beta, \quad (9)$$

which is reasonable given the similar correlation found for the broad-line region (BLR) (Kaspi et al. 2000) and the actual observed NLR size in a number of sources. Luminosity scaling suggests $0.5 < \beta < 0.7$, where the lowest value is obtained from the similarity of the narrow-line spectrum in low- and high-luminosity AGN (e.g. Netzer 1990) and $\beta = 0.7$ is the value obtained by Kaspi et al. (2000) for the BLR. For nearby low-luminosity Seyfert 1s, $R_0 \simeq 500$ pc. Thus, the 5-mag range in M_B observed in our sample translates to $R_{\text{NLR}} = 5$ –13 kpc for the highest-luminosity AGN, i.e. the scale of the entire galaxy, and indeed, recent *Hubble Space Telescope* imaging of the NLR in several radio-loud quasars shows equation (9) to hold for R_{NLR} up to ~ 10 kpc with $\beta = 0.5$ (Bennert et al. 2002). It is therefore possible that the NLR gas, if it were there in the first place, has long left the galaxy and most high-luminosity AGN contain weak or non-existent NLRs. Moreover, if $\beta = 0.7$ as suggested here, and if the NLR density is independent of size, we can perhaps explain the decreasing equivalent width of [Ne V] as being a result of a decreasing ionization parameter with luminosity.

The [Ne III] W_λ correlation with luminosity is marginal although with the right trend (Fig. 7). However, the above simplified model does not explain the different behaviour of the [O III] line that shows no obvious correlation. We note, however, that the [Ne III], [Ne V] and [O II] lines are measured over a larger magnitude range, compared with [O III], because the [O III] lines are lost off the red end of the spectrum at much lower redshifts (and hence lower luminosities in a flux-limited sample) than the other emission lines. Thus, the reality of the model will have to be tested when more [O III] line measurements of higher-luminosity AGN are available. Needless to say, if the model suggested here is confirmed by future observation, it will have serious implications for searches for luminous *type 2* QSOs, which simply may not have luminous narrow-emission-line regions.

5.3 The broad-line region

Finally, we consider emission from the BLR. Most of the strong emission lines we analyse emanate from this region and we will discuss each of the lines in turn.

The Ly α line is the only one to show a clear correlation with redshift (top left-hand plot in Fig. 5). We see an increase in Ly α strength with redshift. However, this is likely to be caused by a combination of increased Ly α forest absorption and our normalization of the continuum. As we have no knowledge of the continuum shape of our spectra, we cannot extrapolate a power-law slope from the red to the blue side of Ly α . Thus, if absorption increases, our continuum point on the blue side of the line will be increasingly

depressed, resulting in an overestimate of the line flux. The N v line, close to Ly α will also be affected by this problem, and is generally dominated by the errors in subtracting the Ly α components. Previous work (Francis & Koratkar 1995) has found correlations between the strength of Ly α and redshift, with higher-redshift QSOs having weaker Ly α . Francis & Koratkar attribute this largely to an increase in Ly α forest absorption in their flux-calibrated spectra. In our composite spectra, which are continuum divided, this same Ly α forest absorption causes an underestimate of the continuum level. This in turn produces the positive correlation of W_λ with redshift.

The Si iv + O iv] lines had to be fitted as one feature, as they could not be de-blended. Apart from this they are relatively clear of contamination. This blend shows no evidence whatsoever for any correlation with luminosity, with equivalent widths that are constant to ~ 15 per cent over a factor of ~ 40 in luminosity. This is in disagreement with Green et al. (2001) who find a significant correlation with $W_\lambda \propto L^{-0.30 \pm 0.08}$. We find a shallower slope that is consistent with zero. We note that Green et al. measure their correlation over a range that is only a factor of ~ 10 in luminosity.

The C iv line shows a highly significant correlation of equivalent width with luminosity, $W_\lambda \propto L^{-0.128 \pm 0.015}$. The C iii] blend (C iii], Al iii and Si iii]) correlation is flatter, with $W_\lambda \propto L^{-0.070 \pm 0.008}$. However, this blend also shows apparent significant correlations with redshift which could be, in some part, a result of the combination of lines taken. For Mg ii, we find $W_\lambda \propto L^{-0.058 \pm 0.013}$ with significant departures away from the correlation at low luminosity, $M_B > -22$. Checks of the individual line fits show no obvious systematic problems, and it appears that these deviations from the simple power law are real. If they were also present in the C iv or C iii] lines and occurred at the same luminosity we would not see the effect as the faintest QSOs that have these lines visible are too bright ($M_B = -23$ to -24). An alternative cause is a change in the SED, e.g. of the host galaxy, but this effect is not seen in the narrow [Ne v] emission line, which spans a similar range in luminosity. However, we also note that the Mg ii emission line lies on top of the broad Fe ii emission feature that we have treated as a continuum. This Fe ii emission may contribute a significant proportion (~ 25 per cent) of this continuum emission and hence any variation in the strength of the Fe ii emission as a function of luminosity will affect the Mg ii equivalent width measurements.

Perhaps the biggest surprise of this analysis is the Balmer line correlation, which is markedly different from that of the broad UV emission lines. Although the broad H δ line shows no significant correlation, the equivalent widths of the H γ and H β lines show a *positive* correlation with luminosity. There is a hint that the strength of the correlation increases for lower-order Balmer lines with $W_\lambda \propto L^{0.160 \pm 0.020}$ and $W_\lambda \propto L^{0.185 \pm 0.030}$ for H γ and H β , respectively. This *inverse Baldwin effect* has not been seen previously. Note that for the H γ line we included the narrow component and the [O iii] $\lambda 4363$ line in the equivalent width, while for H β we subtract the narrow component. However, the positive correlations with luminosity remain if the narrow H γ component is subtracted from the emission line or the narrow H β component is included.

We plot the measured slopes of equivalent width dependence on luminosity against the ionization potential (filled circles in Fig. 9) for all the major emission lines in this study. Excluding the [O ii] line, which could be contaminated by emission from the host galaxy, we see a correlation between the slope of the Baldwin relation and the ionization potential.

A possible explanation for this effect is that the SED of the ionizing continuum may steepen towards lower energies with increasing luminosity, resulting in more photons being available to ion-

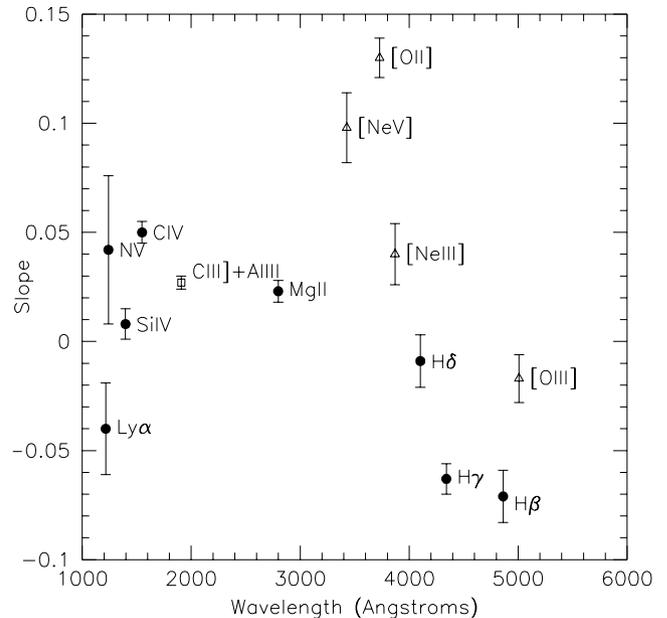


Figure 10. The measured slopes of the log W_λ versus M_B correlation for different lines as a function of rest wavelength for broad permitted lines (filled circles), narrow forbidden lines (open triangles) and semiforbidden C iii] (open square).

ize hydrogen but relatively fewer with energies greater than 64 eV available to ionize C iv. Alternatively, the ionization parameter may change as a function of luminosity, as discussed below.

On the other hand, the correlations we measure between equivalent width and luminosity may be caused by changes in the continuum flux under the lines rather than the line flux itself. We plot the slope of the equivalent width dependence on luminosity versus the rest wavelength of the features in question (Fig. 10). In this case, there may be a trend for the slope of the Baldwin relation to increase with decreasing wavelength. However, the broad lines show a much flatter correlation with wavelength than the narrow lines (even excluding [O ii]); arguing against an effect caused by a simple continuum variation.

Our results imply that either (i) $L(\text{H}\beta)/L(\text{C iv})$ increases with luminosity, (ii) $L(4860\text{\AA})/L(1550\text{\AA})$ decreases with luminosity or a combination of the two. We have no way of answering this question directly since our data are not flux calibrated and we cannot measure the above luminosity ratios. However, we can refer to earlier findings discussing the line and continuum luminosities in smaller, less complete samples.

Earlier studies of AGN SEDs (e.g. Vanden Berk et al. 2001 and references therein) show that the slope of the power-law continuum changes dramatically near to the H γ and H β lines. Bluewards of about 4500 \AA the continuum slope has $\alpha \sim -0.5$, while redwards of this point the slope is more like $\alpha \sim -1.6$. This has been interpreted as being a result of the different continuum processes contributing to the SED at different wavelengths (Laor 1990). At short wavelengths most of the emission is caused by accretion discs (e.g. Laor & Netzer 1989 and references therein) while at longer wavelengths dust emission, combined with non-thermal emission in radio-loud sources, is more important. Bearing this in mind we can speculate that the relative contribution is luminosity dependent in a sense that the accretion disc contribution is more important at long wavelengths in higher-luminosity sources. This is exactly the trend observed by Laor (1990) in his study of the continuum emission in

high-luminosity AGN. Other ideas are related to the accretion disc inclination (Netzer 1985; Netzer, Laor & Gondhalekar 1992; Wilkes et al. 1999), changes in the ionizing luminosity as a function of luminosity (Espey & Adreadis 1999; Green 1998; Korista, Baldwin & Ferland 1998; Wandel 1999) (considered above) or emission by optically thin gas (Shields, Ferland & Peterson 1995). Most of these models predict a similar trend for all lines albeit with a different slope. No existing model, except for that involving changes in the relative luminosity of the two components, can explain the observed difference between C IV and H β .

Another possibility is to test the dependence of optical and UV line ratios, such as $L(\text{H}\beta)/L(\text{Ly}\alpha)$, on luminosity and continuum shape as was done for example by Netzer et al. (1995) for a small (20) sample of radio-loud AGN. That study shows a clear correlation of $L(\text{H}\beta)/L(\text{Ly}\alpha)$ with $L(4861 \text{ \AA})/L(1216 \text{ \AA})$. However, the small size of the sample, and the small luminosity range, prevents any clear conclusion.

Finally, we note that the H β luminosity range and the C IV luminosity range are very different, with almost no overlap, because of the $z-M_B$ correlation in our sample. It is therefore possible that the *real* Baldwin relationship for *all lines* is more complicated than previously assumed, showing both rising and falling branches. This may be related to the unusual shape of the M_B-W_λ curve of the Mg II line seen in Fig. 7. This idea can only be studied by obtaining good H β measurements in high-luminosity AGN.

6 CONCLUSIONS

We have used composite QSO spectra to make the most accurate determination of line and continuum correlations to date. We see the Baldwin effect in a number of lines. In general, the equivalent width correlations are primarily with luminosity and not redshift. The broad UV lines generally show strong anticorrelations with luminosity, although somewhat flatter than previous determinations. The Balmer line equivalent widths, in contrast, show an inverse Baldwin effect, and are positively correlated with luminosity. We postulate that this difference could be caused by a different combination of disc and non-disc components in AGN of different luminosity.

Some, but not all, narrow forbidden lines also show anticorrelations with luminosity. A possible explanation is that the NLR becomes more extended, and fainter by comparison, at high luminosity. This has important implications concerning the possible detection of type 2 QSOs at high redshifts. By comparing the strength of the Ca II K absorption line with the [O II] emission line via a mean galaxy spectrum we find that at low luminosities most, if not all, of the [O II] flux could come from the host galaxy and not the AGN. This raises the possibility that a large fraction of the observed [O II] in high-luminosity AGN is caused by enhanced nuclear star formation. Using the Ca II K line and assuming a constant SED for the host galaxy, we are able to derive the correlation between the host galaxy and the AGN luminosity, which is $L_{\text{gal}} \propto L_{\text{QSO}}^{0.417 \pm 0.045}$.

A number of areas have still to be investigated. First is the detailed shapes and positions of the lines, which will be discussed in a forthcoming paper (Corbett et al., in preparation). Secondly, after looking at the mean properties as a function of redshift and luminosity, we should also investigate the variance about this mean by fitting individual spectra.

ACKNOWLEDGMENTS

The 2dF QSO Redshift Survey was based on observations made with the Anglo-Australian Telescope and the UK Schmidt Telescope.

We warmly thank all the present and former staff of the Anglo-Australian Observatory for their work in building and operating the 2dF and 6dF facilities. KR was supported by an AAO summer vacation studentship during the course of this work. HN thanks the director and staff of the AAO for their hospitality and support during a 2-month sabbatical visit in early 2002.

REFERENCES

- Bailey J., Glazebrook K., Offer A., Taylor K., 2002, MNRAS, submitted
 Baldry I.K. et al., 2002, ApJ, 569, 582
 Baldwin J.A., 1977, ApJ, 214, 679
 Baldwin J.A., Wampler E.J., Gaskell C.M., 1989, ApJ, 388, 630
 Bennert N., Falcke H., Schulz H., Wilson A., Wills B.J., 2002, ApJ, 574, L105
 Boyle B.J., 1990, MNRAS, 243, 231
 Boyle B.J., Shanks T., Peterson B.A., 1988, MNRAS, 235, 935
 Boyle B.J., Shanks T., Croom S.M., Smith R.J., Miller L., Loaring N., Heymans C., 2000, MNRAS, 317, 1014
 Brotherton M.S., Wills B.J., Francis P.J., Steidel C.C., 1994a, ApJ, 430, 495
 Brotherton M.S., Wills B.J., Steidel C.C., Sargent W.L.W., 1994b, ApJ, 423, 131
 Cristiani S., Vio R., 1990, A&A, 227, 385
 Croom S.M., Smith R.J., Boyle B.J., Shanks T., Loaring N.S., Miller L., Lewis I.J., 2001, MNRAS, 322, L29
 Espey B., Adreadis S., 1999, in Ferland G., Baldwin J., eds, Proc. ASP Conf. Ser. Vol. 162, Quasars and Cosmology. Astron. Soc. Pac., San Francisco, p. 351
 Francis P.J., Koratkar A., 1995, MNRAS, 274, 504
 Francis P.J., Hewett P.C., Foltz C.B., Chaffee F.H., Weymann R.J., Morris S.L., 1991, ApJ, 373, 465
 Green P.J., 1998, ApJ, 498, 170
 Green P.J., Forster K., Kuraszkiwicz J., 2001, ApJ, 556, 727
 Hewett P.C., Foltz C.B., Chaffee F.H., 1995, AJ, 109, 1498
 Kaspi S., Smith P.S., Netzer H., Maoz D., Jannuzi B., Giveon U., 2000, ApJ, 533, 631
 Korista K., Baldwin J., Ferland G., 1998, ApJ, 507, 24
 Laor A., 1990, MNRAS, 246, 369
 Laor A., Netzer H., 1989, MNRAS, 238, 897
 Macklin J.T., 1982, MNRAS, 199, 1119
 Netzer H., 1985, MNRAS, 216, 63
 Netzer H., 1990, in Blandford R.D., Netzer H., Woltjer L., Courvoisier T.J.-L., Mayor M., eds, Active Galactic Nuclei (SAA-FEE 1990). Springer-Verlag, Berlin
 Netzer H., Laor A., Gondhalekar P.M., 1992, MNRAS, 254, 15
 Netzer H., Brotherton M.S., Wills B.J., Han M., Wills D., Baldwin J.A., Ferland G.J., Browne I.W.A., 1995, ApJ, 448, 27
 Schade D.J., Boyle B.J., Letawsky M., 2000, MNRAS, 315, 498
 Schlegel D.J., Finkbeiner D.P., Davis M., 1998, ApJ, 500, 525
 Schneider D.P. et al., 2002, AJ, 123, 567
 Shields J.C., Ferland G.J., Peterson B.M., 1995, ApJ, 441, 507
 Smith R.J., Croom S.M., Boyle B.J., Shanks T., Loaring N.S., Miller L., 2002, MNRAS, submitted
 Vanden Berk D.E. et al., 2001, AJ, 122, 549
 Wandel A., 1999, ApJ, 527, 649
 Wilkes B.J., Kuraszkiwicz J., Green P.J., Mathur S., McDowell J.C., 1999, ApJ, 513, 76
 Wills B.J., Brotherton M.S., Fang D., Steidel C.C., Sargent W.L.W., 1993, ApJ, 415, 563
 Zamorani G., Marano B., Mignoli M., Zitelli V., Boyle B.J., 1992, MNRAS, 256, 238