Can we improve C IV-based single-epoch black hole mass estimations?

J. E. Mejía-Restrepo,[1,2]★ B. Trakhtenbrot,[3]† P. Lira[1] and H. Netzer[4]

[1]*Departamento de Astronomía, Universidad de Chile, Camino el Observatorio 1515, Santiago 7591245, Chile*

[2] *European Southern Observatory, Casilla 19001, Santiago 19, Chile*

[3]*Department of Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, CH-8093 Zurich, Switzerland*

[4]*School of Physics and Astronomy, Tel Aviv University, Tel Aviv 69978, Israel*

ABSTRACT

In large optical surveys at high redshifts ($z > 2$), the C IV $\lambda 1549$ broad emission line is the most practical alternative to estimate the mass ($M_{\rm BH}$) of active supermassive black holes. However, mass determinations obtained with this line are known to be highly uncertain. In this work, we use the Sloan Digital Sky Survey Data Release 7 and 12 quasar catalogues to statistically test three alternative methods put forward in the literature to improve CIV-based $M_{\rm BH}$ estimations. These methods are constructed from correlations between the ratio of the C IV $\lambda 1549$ linewidth to the low-ionization linewidths (H α, H β, and MgII $\lambda 2798$) and several other properties of rest-frame ultraviolet emission lines. Our analysis suggests that these correction methods are of limited applicability, mostly because all of them depend on correlations that are driven by the linewidth of the CIV profile itself and not by an interconnection between the linewidth of the C IV line and the linewidth of the low ionization lines. Our results show that optical CIV-based mass estimates at high redshift cannot be a proper replacement for estimates based on infrared spectroscopy of low ionization lines like H α, H β, and MgII.

Key words: quasars: general – quasars: emission lines – quasars: super massive black holes – galaxies: nuclei.

1 INTRODUCTION

Accurate determinations of supermasssive black hole (SMBH) masses ($M_{\rm BH}$s) are essential to fully understand SMBH – the physics, demographics, and relations with galaxies. The single-epoch (SE) black hole mass estimation method is commonly used on large samples of unobscured, type-I active galactic nuclei (AGN; McLure & Dunlop 2004; Onken & Kollmeier 2008; Shen et al. 2008; Fine et al. 2010; Rafiee & Hall 2011; Trakhtenbrot & Netzer 2012). This method relies on two basic ingredients: (1) the assumption of virialized gas kinematics in the broad line region (BLR) and (2) the empirical relation from reverberation mapping (RM) experiments between the BLR size ($R_{\rm BLR}$) and the continuum luminosity ($L_\lambda \equiv \lambda L(\lambda)$) at particular wavelength (λ) where $R_{\rm BLR} \propto (L_\lambda)^\alpha$ with $\alpha \sim 0.5-0.7$ (Kaspi et al. 2000, 2005; Bentz et al. 2009; Park et al. 2012; Bentz et al. 2013).

Under these assumptions, the width of the broad emission lines, such as the full width at half-maximum (FWHM), is a good proxy for the virial velocity of the BLR clouds. $M_{\rm BH}$ can thus be expressed as

$$M_{\rm BH} = f G^{-1} R_{\rm BLR} v_{\rm BLR}^2 = K(L_\lambda)^\alpha {\rm FWHM}^2. \qquad (1)$$

Here G is the gravitational constant and f is a geometrical factor that accounts for the unknown structure and inclination to the line of sight of the BLR. In this paper, we assume $f = 1$, which is an appropriate median value for $M_{\rm BH}$ estimates using the FWHM (Woo et al. 2015). However, there is a large uncertainty in this value (of at least a factor of 2; e.g. Onken et al. 2004; Woo et al. 2013; Shankar et al. 2016; Batiste et al. 2017) that can be even larger if f depends on luminosity and/or other line properties (e.g. equivalent widths, line offsets, FWHM; Collin et al. 2006; Shen 2013; Mejía-Restrepo et al. 2017).

The most reliable RM-based $R_{\rm BLR}-L$ relation is the $R_{\rm BLR}({\rm H}\,\beta)-L_{5100}$ relation. This relation is the only one that has been established for a large number of sources and covering a broad luminosity range ($10^{43}\,{\rm erg\,s^{-1}} \lesssim L_{5100} \lesssim 10^{46}\,{\rm erg\,s^{-1}}$). Consequently, SE $M_{\rm BH}$ calibrations for other lines are often re-calibrated to match $M_{\rm BH}$ measurements based on H β and L_{5100}. Such re-calibrations are used to determine $M_{\rm BH}$ values at different redshifts where the H β is not available due to observational limitations. In optical surveys, the H α and H β lines can be used up to $z \lesssim 0.8$ (e.g. Greene & Ho 2005; Netzer & Trakhtenbrot 2007; Xiao et al. 2011; Shen & Liu 2012), the Mg II $\lambda 2798$ (hereafter Mg II) can assist for $M_{\rm BH}$ on sources where $0.6 \lesssim z \lesssim 2.2$ (e.g. McLure & Jarvis 2002; Vestergaard & Osmer 2009; Wang et al. 2009; Trakhtenbrot et al. 2011; Shen & Liu 2012; Trakhtenbrot & Netzer 2012), and the CIV $\lambda 1549$

★ Email: jmejiar@eso.org

† Zwicky postdoctoral fellow

line (hereafter C IV) is used to estimate black hole masses at even higher redshifts ($2.0 \lesssim z \lesssim 5.0$; Vestergaard & Peterson 2006; Park et al. 2013).

M_{BH} calibrations based on low ionization lines (i.e. H α, and Mg II) generally show good agreement with the H β M_{BH} estimator with a typical scatter of $\lesssim 0.2$ dex (Greene & Ho 2005; Xiao et al. 2011; Trakhtenbrot & Netzer 2012). However, the analogous recalibration using the C IV high ionization line is more problematic and shows large scatter (0.4–0.5 dex), possibly driven by several causes. First, the width of C IV is only weakly correlated, if at all, with the width of the low ionization lines and presents large scatter in many AGN samples (e.g. Baskin & Laor 2005; Netzer et al. 2007; Shang et al. 2007; Shen et al. 2008; Fine et al. 2010; Ho et al. 2012; Shen & Liu 2012; Trakhtenbrot & Netzer 2012; Tilton & Shull 2013). Secondly, the C profiles show large IVblueshifts (up to several thousands km s^{-1}; Richards et al. 2002; Baskin & Laor 2005; Shang et al. 2007; Richards et al. 2011) that indicates non-virial motions. Thirdly, Denney (2012) found that the core of the broad C IV line does not reverberate in response to continuum variations. This implies that not only the innermost but also the *outermost* C IV-emitting regions may not be virialized.

Given that C IV is the most widely used line for M_{BH} determination at $z \gtrsim 2$ in optical surveys, it is crucial to design practical methodologies to mitigate the issues related to C IV-based M_{BH} determinations. There have been many efforts to improve single-epoch M_{BH} determinations from C IV (e.g. Vestergaard & Peterson 2006; Assef et al. 2011; Denney et al. 2013; Park et al. 2013; Runnoe et al. 2013; Tilton & Shull 2013; Brotherton et al. 2015; Coatman et al. 2016). The studies of Assef et al. (2011), Denney et al. (2013), Park et al. (2013), and Tilton & Shull (2013) claimed that in spectra of limited signal to noise (S/N) and/or spectral resolution, FWHM(C IV) measurements are underestimating the 'real' linewidths, in objects with strong intrinsic absorption features that cannot be deblended from the emission lines. This would partially explain the fact that about 40 per cent of quasars show C IV profiles narrower than the H β profiles (Trakhtenbrot & Netzer 2012), in contrast to expectations from RM experiments, which finds that the bulk of the C IV-emitting region is interior to that of the low ionization lines (Peterson et al. 2005; Kaspi et al. 2007; Trevese et al. 2014, Lira et al. 2017 submitted). However, Corbin & Boroson (1996) found that even objects with no evidence of absorption features show FWHM(C IV) < FWHM(H β).

Additionally, Assef et al. (2011), Denney et al. (2013), Park et al. (2013), and Tilton & Shull (2013) also found that in high-quality spectrum (S/N $\gtrsim 10$, $R \gtrsim 2000$), the line dispersion (σ_{line}) of the C IV line provides C IV-based M_{BH} estimations in better agreement with H β-based estimations than the FWHM of the C IV line. However, in Mejía-Restrepo et al. (2016), we analysed the impact of using different approaches to trace the continuum emission underneath the emission lines in estimating σ_{line} and FWHM of broad emission lines. Our results indicate that σ_{line} measurements are very sensitive by the continuum determination even in high-quality spectra. On the other hand, we found that FWHM measurements are only weakly affected by this effect ever since the line profile is not affected by strong absorption feature. This result suggests the usage of the FWHM, instead of σ_{line}, for C IV-based M_{BH} estimations after neglecting those objects with strong absorption features.

Recently, Coatman et al. (2017) found a strong correlation between the blueshift of the C IV line centroid and the C IV/H α FWHM ratio for a sample of 66 high luminosity ($10^{46.5}$ erg s^{-1} < L_{Bol} < $10^{47.5}$ erg s^{-1}) and high redshift quasars ($z > 2.1$). They suggested that this correlation can assist to improve C IV-based black hole

masses reducing the scatter between C IV and H α based M_{BH} determinations from 0.40 to 0.24 dex. However, this procedure is not applicable to large optical surveys because of the difficulty to accurately determine the AGN redshift, necessary to compute the C IV blueshift, without information from low ionization lines. Runnoe et al. (2013) and Brotherton et al. (2015) used a sample of 85 low-redshift ($0.03 < z < 1.4$) and low-to-moderate luminosity ($10^{43.4}$ erg s^{-1} < L_{5100} < $10^{46.5}$ erg s^{-1}) AGN with quasi-simultaneous ultraviolet (UV) and optical rest-frame spectra to propose a method to rehabilitate C IV for M_{BH} determination. The method consisted of using a correlation that they found between the Si IV+O IV] $\lambda 1400$/C IV line peak intensity ratio and the H β/C IV FWHM ratio. They claim that by using this correlation, it is possible to predict FWHM(H β) from measurements of the Si IV+O IV] $\lambda 1400$ (hereafter Si IV+O IV) emission to obtain more accurate C IV-based mass measurements. They specifically claim that the scatter between C IV and H β estimations is reduced from 0.43 to 0.33 dex.

In our previous studies, we presented a sample of 39 high-quality, simultaneous (rest-frame) UV-optical spectra of type-1 AGN at z $\sim$ 1.5 obtained with X-Shooter (Capellupo et al. 2015, 2016). Using this sample, in Mejía-Restrepo et al. (2016) we were able to reproduce the correlation found in Runnoe et al. (2013), but with a weaker statistical significance. Mejía-Restrepo et al. (2016) also found a similar but alternative correlation between the C III] $\lambda 1909$/C IV line peak intensity ratio and the H β/C IV FWHM ratio. In general, we found that the ratios of FWHM(C IV) to the FWHM of the H α, H β, and Mg II low ionization lines are correlated with both, the Si IV+O IV/C IV and the C III] $\lambda 1909$/C IV line peak ratios. In spite of these correlations, we found that none of them are able to reduce the scatter between C IV-based M estimations and the $_{BH}$low-ionization line M_{BH} estimations.

It is important to point out that the findings of Coatman et al. (2017), Mejía-Restrepo et al. (2016), and Runnoe et al. (2013) are all obtained from relatively small samples (230, 69, and 39 objects, respectively) that map different regions in the parameter space of the AGN population (see Section 2). Thus, the significance of their findings may not be applicable to the overall population of non-obscure type-I AGN population.

The purpose of this work is to test the validity of these empirical alternatives to improve C IV-based M_{BH} estimations on large AGN samples with survey-grade spectroscopic data. To accomplish this goal, in this paper we use data from the Sloan Digital Sky Survey (SDSS York et al. 2000), specifically from the SDSS-II data release 7 and the SDSS-III data release 12 quasar spectroscopic catalogues (DR7Q and DR12Q, respectively, Schneider et al. 2010; Pâris et al. 2017). All the alternatives that we are testing here stand on correlations that relate FWHM (C IV) /FWHM (H β) with the C IV line itself or with the properties of emission lines or continuum windows that are close to the C IV line. Due to the lack of simultaneous coverage of C IV and H β lines in the optical SDSS survey, we will carry our analysis in terms of FWHM (C IV) /FWHM (Mg II) $\equiv$ FWHM [C IV/Mg II] instead of FWHM (C IV) /FWHM (H β). This is justified as it is well known that FWHM(Mg II) is tightly correlated with FWHM(H β) and that M_{BH} estimations from these two emission lines are known to agree within 0.2 dex of accuracy (e.g. Wang et al. 2009; Shen & Liu 2012; Trakhtenbrot & Netzer 2012; Mejía-Restrepo et al. 2016). Because of the limitation in S/N of the SDSS data and the difficulties in measuring the σ_{line} of C IV, in this paper we do not thoroughly explore the usage of this quantity for C IV measurements; we, however, briefly analyse the feasibility of its usage in a high-quality spectra subsample of the SDSS DR12Q catalogue.

This paper is structured as follows. In Section 2, we present the samples and introduce the most relevant parameters that we measured for our analysis. In Section 3, we present and discuss our main results and in Section 4 we highlight our most important findings. Throughout this paper, we assume a flat Λ cold dark mattercosmology with the following values for the cosmological parameters: $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$, and $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

2 SAMPLES, DATA, AND ANALYSIS

In this section, we describe in detail two large samples, namely, the SDSS DR7Q and the DR12Q samples, as well as three small samples taken from Mejía-Restrepo et al. (2016), Runnoe et al. (2013), and Coatman et al. (2017). We also describe the spectral fitting procedure and the emission line and continuum properties that were obtained for the analysis presented here.

2.1 Large samples

To accomplish our goal, we need to guarantee the simultaneous coverage of the Si IV+O IV], C IV, C III], and Mg II emission lines. According to the spectroscopic coverage of the DR7Q (3800–9200 Å) and the DR12Q (3600–10 400 Å) samples, we selected objects with $1.8 < z < 2.0$ and $1.70 < z < 2.3$, respectively. These redshift constraints translate into a total of 4817 objects for the DR7Q catalogue and 69092 objects for the DR12Q catalogue. Although the objects from the SDSS DR7 sample are also included in the DR12Q sample, to construct the DR12Q catalogue, the objects from the DR7Q catalogue were re-observed. The time interval between observations is at least 4 years.

For the DR7Q sample, we used the redshift estimations from Hewett & Wild (2010), which provides important improvements to the SDSS redshift estimations with a reduction of a factor of 6 of the systematic uncertainties with respect to SDSS redshift estimations. For the DR12Q sample, Hewett & Wild (2010) redshift calculations are not available. We therefore adopt the visually inspected SDSS redshift estimations described in Pâris et al. (2017). Consequently, line shift estimations in the SDSS DR12Q sample are less reliable than in the DR7Q sample.

It is also important to note that because the survey was designed to map the large-scale structure of the Universe at high redshift, the DR12Q catalogue is primarily biased towards $z > 2$ sources (Schlegel, White & Eisenstein 2009). In particular, in our subsample of the DR12Q catalogue, 75 per cent of the objects are at $z > 2.0$.

2.2 Small samples

We complement our analysis with three additional smaller samples with considerably higher quality spectroscopic data. These samples correspond to the original samples used to propose the different methodologies to improve C IV-based M_{BH} estimations that we described in the introduction.

The first sample is described in Capellupo et al. (2016, hereafter the X-Shooter sample) consisting of 39 radio-quiet (RQ) quasars observed by the X-Shooter spectrograph that guaranteed simultaneous observations of the rest-frame UV and optical range. The sample comprises objects with $1.45 < z < 1.69$ and $10^{44.8}\mathrm{erg\,s^{-1}} < L_{1450} < 10^{46.8}\mathrm{erg\,s^{-1}}$.

The second of these samples is described in Runnoe et al. (2013, hereafter R13 sample) consisting of 69 objects including 37 radio-loud (RL) and 32 RQ quasars with nearly simultaneous observations of the (rest-frame) X-ray, UV, and optical range. This sample is a subset of the Tang et al. (2012) sample and comprises objects with $0.03 < z < 1.4$ and $10^{43.6}\mathrm{erg\,s^{-1}} < L_{1450} < 10^{46.7}\mathrm{erg\,s^{-1}}$.

Finally, the sample described in Coatman et al. (2017, hereafter the C17 sample) consists of a compilation of 230 RQ quasars with $10^{45.7}\mathrm{erg\,s^{-1}} < L_{1350} < 10^{47.7}\mathrm{erg\,s^{-1}}$. This sample comprises sources from Shen & Liu (2012), Coatman et al. (2016), and Shen et al. (2016). All the sources have non-simultaneous optical observations (from SDSS) and ground-based near-infrared observations, which at the redshift range of the sample ($1.5 < z < 4.0$) correspond to the rest-frame UV and optical range, respectively. This sample has not reported Mg II emission line measurements, but includes H α emission line measurements that can be used as a proxy for Mg II line measurements (see e.g. Shen & Liu 2012; Mejía-Restrepo et al. 2016).

2.3 Line and continuum measurements

For each object in the SDSS DR7Q and DR12Q samples, we fitted the line profiles of the Si IV+O IV], C IV, C III], and Mg II emission lines as described in Appendix A. From the best-fitting model of the emission lines of each object, we measured the line FWHM, the velocity dispersion (σ_{line}; following Peterson et al. 2004), the rest-frame equivalent width ($EW(line)$), the integrated line luminosity ($L(line)$), and the luminosity at the peak of the fitted profile ($L_{peak}(line) \equiv 4\pi D_L^2 F_{\lambda,peak}(line)$). As line blueshift indicators, we measured two different quantities: (1) the shift of the emission line peak (Δv_{peak}) and (2) the line centroid shift defined as shift in the flux-weighted central wavelength (Δv_{line}, following Peterson et al. 2004). We also computed the monochromatic luminosities at different wavelengths ($L_\lambda \equiv \lambda L(\lambda)$). We particularly measured L_{1350}, L_{1450}, L_{2000}, and L_{3000} that correspond to continuum bands adjacent to the Si IV+O IV], C IV, C III], and the Mg II emission lines, respectively. Finally, from the large DR7Q and DR12Q samples, we excluded broad absorption line quasars (BALQSOs) and objects with unreliable fits following the strategy described in AppendixA. We ended up with 3267 objects from the DR7Q catalogue (out of 4817) and 35674 from the DR12Q catalogue (out of 69062 objects).

In the case of the X-Shooter, R13, and C17 samples, we also extracted the measurements of the aforementioned quantities whenever available from the published data in Mejía-Restrepo et al. (2016), Runnoe et al. (2013), and Coatman et al. (2017), respectively. Although the fitting approaches in each of these papers are not identical, they follow similar procedures and then provide comparable measurements.

Mejía-Restrepo et al. (2016) showed that σ_{line} and $L(line)$ are very sensitive to the continuum placement because of their strong dependence on the line wings. Analogously, Δv_{line} (C IV), one of the most widely used blueshift indicators, is also affected by the continuum placement. This fact motivates us to use the alternative blueshift estimators Δv_{peak} (see definition above). Similarly, $EW(line)$ is also sensitive to continuum placement. Therefore, in addition to $EW(line)$, we also use $L_{peak}(line)/L_\lambda$ because of its weaker dependency on continuum placement.

We thus have a set of quantities that are weakly sensitive to continuum placement given by FWHM, $L_{peak}(line)$, Δv_{peak}, and $L_{peak}(line)/L_\lambda$, as well as a set of quantities that are strongly affected by the placement of the continuum emission given by σ_{line}, $L(line)$, Δv_{line}, and $EW(line)$. We emphasize, however, that the latter quantities are also important because they carry information about the broadest components of the emission lines and therefore of the innermost region of the BLR.

From all the quantities considered here, the most relevant parameters for our analysis are the following:

(i) L_{1450}
(ii) L_{3000}
(iii) FWHM(C IV)
(iv) FWHM(Mg II)
(v) FWHM $[$C IV/Mg II$]$ $\equiv$ FWHM (C IV) /FWHM (Mg II)
(vi) L_{peak} $[$C IV/SiO IV$]$ $\equiv$ L_{peak} (C IV) $/L_{\text{peak}}$ (Si IV + O IV])
(vii) L_{peak} $[$C IV/C III]$]$ $\equiv$ L_{peak} (C IV) $/L_{\text{peak}}$ (C III])
(viii) L_{peak} $[$C IV/1450Å$]$ $\equiv$ L_{peak} (C IV) $/L$ $(1450Å)$
(ix) Δv_{line} (C IV), blueshift of the C IV line centroid.
(x) Δv_{peak} (C IV), blueshift of the C IV line peak.

In Figs 1–5, we present relevant information associated with these quantities. First, in Fig. 1, we show the bi-bdimensional distribution of $\log$ FWHM (C IV) versus $\log L_{1450}$ (left column) and $\log$ FWHM (Mg II) versus $\log L_{3000}$ (right column) for the DR7Q (top panels) and the DR12Q (bottom panels) samples. We continue with Fig. 2 where we show the bi-dimensional distribution of $\log$ FWHM (C IV) versus $\log$ FWHM (Mg II) for the DR7Q (left) and the DR12Q (right) samples. We also show in Figs 3 and 4 the bi-dimensional distributions of $\log$ FWHM (C IV) and $\log$ FWHM $[$C IV/Mg II$]$ versus $\log L_{\text{peak}}$ [C IV/SiO IV] (left), $\log L_{\text{peak}}$ [C IV/C III]] (centre), and $\log L_{\text{peak}}$ $[$C IV/1450Å$]$ (right) for the DR7Q and the DR12Q samples, respectively. Finally, in Fig. 5, we show the bi-dimensional distributions of $\log$ FWHM $[$C IV/Mg II$]$ versus Δv_{peak} (C IV) and Δv_{line} (C IV) for the DR7Q (columns 1 and 2 from left to right) and the DR12Q (columns 3 and 4) samples, respectively. We also show the cumulative distribution function (CDF) of all the quantities and superimpose the relevant information from the X-Shooter, R13, and C17 samples whenever available. In all these figures, we show the Pearson correlation coefficient (r_{P}) for the DR7Q and DR12Q bi-dimensional distribution of the associated quantities. The associated probability for upholding the null hypothesis (P_{P}) is also shown whenever $P_{\text{P}} > 1 \times 10^{-8}$; otherwise, it is approximated to $P_{\text{P}} = 0$.

As can be seen in all these figures, all samples used for this paper are subject to different limitations and, potentially, different selection effects and biases. On the one hand, large samples have the advantage of better sampling the overall quasar population. However, they not only have limited data quality, but are also incomplete at low luminosities (because of flux limits) and high luminosities (because of the upper redshift cuts; Labita et al. 2009). On the other hand, our small samples have very good data quality, but cannot statistically represent the overall quasar population. For further details, in AppendixB, we discuss the particular advantages and limitations related to the large and small samples used for this work.

3 RESULTS AND DISCUSSION

In this section, we explore in detail the different methods suggested by the aforementioned authors to improve C IV-based M_{BH} estimations. First, we will analyse our results from the largest SDSS DR7Q and DR12Q samples to subsequently contrast them with those obtained from the X-Shooter, R13, and C17 samples and discuss the possible problems in the analysis done with such small samples.

3.1 SDSS DR7Q and DR12Q samples

In Table A3, we present the correlation matrix associated with the most relevant measurements relating the C IV and Mg II lines, and the continuum emission from the accretion disc in both SDSS samples.

One important result shown in this table, as well as in Fig. 2, is the very weak (or absent) correlation between FWHM(C IV) and FWHM(Mg II) ($0 < r_{\text{P}} < 0.13$, $P_{\text{P}} < 2 \times 10^{-13}$) that inhibits the possibility to derive reliable C IV-based M_{BH} estimations by only comparing the C IV with the Mg II linewidths. One alternative to overcome this issue is by means of a correlation between the ratio of FWHM(C IV) to the FWHM of the low ionization lines, and other emission lines and/or continuum property. This would provide a simple procedure to predict the width of low ionization lines in terms of FWHM(C IV) and other observed properties as already proposed by Runnoe et al. (2013), Mejía-Restrepo et al. (2016), and Coatman et al. (2017).

3.1.1 Line peak ratios

We first explore the statistical significance of the anticorrelations that link FWHM $[$C IV/Mg II$]$ with L_{peak} [C IV/SiO IV], L_{peak} [C IV/C III]], and L_{peak} $[$C IV/1450Å$]$ and that are used to improve C IV-based M_{BH} estimations. The reason to include L_{peak} $[$C IV/1450 Å$]$ in this analysis, which has not been considered in the literature, is its independence from other emission lines.[1] Hereafter, we will refer to these three quantities as the line peak ratio quantities.

Figs 3 and 4 show that FWHM $[$C IV/Mg II$]$ and FWHM(C IV) are anticorrelated with L_{peak} $[$C IV/1450 Å$]$, L_{peak} [C IV/SiO IV], and L_{peak} [C IV/C III]] in both SDSS quasar samples.[2] Additionally, the corresponding values of r_{P} suggest that in most cases, the anticorrelations of the three line peak ratio quantities with FWHM(C IV) are tighter than those with FWHM $[$C IV/Mg II$]$ with the exception of L_{peak} [C IV/C III]] in the DR7 sample where the anticorrelations are of comparable strength (Fig3, middle column). In addition to this, the data presented in Table 1 show that the scatter of the correlations associated with FWHM(C IV) is smaller than in those associated with FWHM $[$C IV/Mg II$]$ in both SDSS quasar samples.

One possibility to explain this behaviour is that the correlations related to FWHM $[$C IV/Mg II$]$ are driven by the more fundamental FWHM(C IV) correlations. This interpretation is supported by the tight correlation between FWHM(C IV) and FWHM $[$C IV/Mg II$]$ that we find in both SDSS samples ($r_{\text{p}} = 0.71$ in both cases). Thus, FWHM $[$C IV/Mg II$]$ is just increasing the scatter of the original correlations with FWHM(C IV).

To test this idea, we first divided our DR7Q and DR12Q samples into two subgroups: objects with FWHM (Mg II) < 3000 km s^{-1} (narrow group) and objects with FWHM (Mg II) > 4500 km s^{-1} (broad group). Then, we binned each group by the line peak ratio quantities with a bin size of 0.2 dex. For each bin, we computed the median value of FWHM $[$C IV/Mg II$]$ and FWHM(C IV) and the corresponding 16 and 84 percentiles to quantify the dispersion in each bin. The light turquoise and red solid lines in Figs 3 and 4 represent the median values and corresponding 1σ dispersion in the narrow and broad groups, respectively. We can see that light turquoise and red lines are very close to each other

---

[1]We also considered the possibility of using the EW(C IV) for our analysis. However, it shows weaker correlations with FWHM $[$C IV/Mg II$]$ and FWHM(C IV) than L_{peak} [C IV/SiO IV], L_{peak} [C IV/C III]], and L_{peak} $[$C IV/1450Å$]$.

[2]Notice that Figs 3 and 4 map the same dynamical range for $\log$ FWHM (C IV) and $\log$ FWHM $[$C IV/Mg II$]$ (1.6 dex in both cases) as well as for $\log L_{\text{peak}}$ [C IV/SiO IV], $\log L_{\text{peak}}$ [C IV/C III]], and $\log L_{\text{peak}}$ $[$C IV/1450Å$]$ (a total of 2.5 dex in all of them).
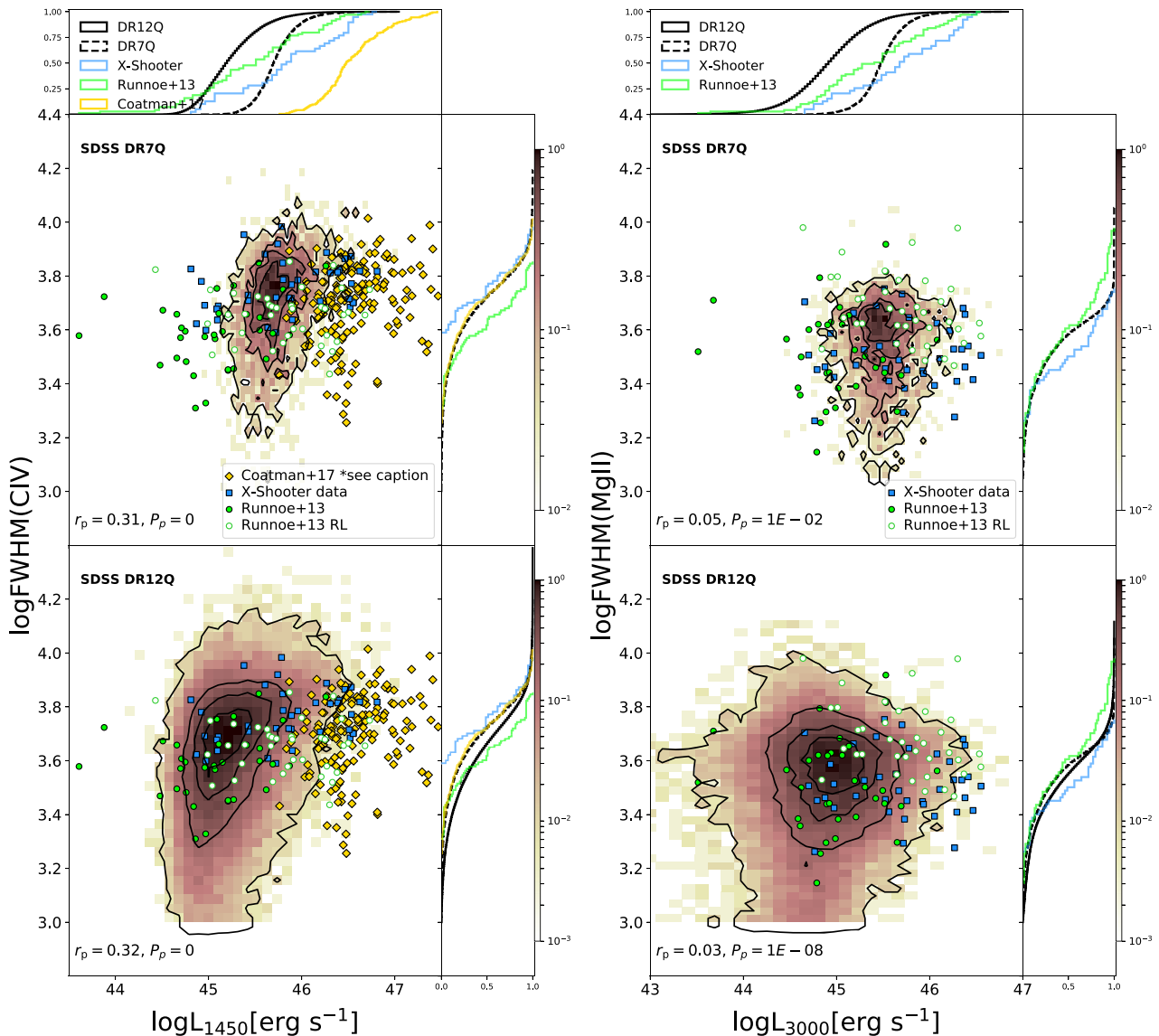
Figure 1. Left column: L_{1450}-FWHM(C IV) bi-dimensional distributions in the SDSS DR7Q (top left) and DR12Q (bottom left) samples. The intensity of the colour represents the relative density of points as shown in the colour bars on the right. The black thin lines represent the 25 per cent, 50 per cent, 75 per cent, and 99 per cent contours centred at the maximum probability point. In the top and right side diagrams, we show the projected CDFs of L_{1450} and FWHM(C IV), respectively. We superimpose in each panel analogue data from the small X-Shooter, R13, and C17 samples as summarized in the legend. R13 is differentiated by RL and RQ objects in the scatter plots. For the C17 sample, we show L_{1350} as a proxy for L_{1450} (L_{1450} is not listed in Coatman et al. 2016). We also show the Pearson correlation coefficient (r_P) and associated probability for upholding the null hypothesis (P_P) for the DR7Q and DR12Q bi-dimensional distribution. Right column: L_{3000}-FWHM(Mg II) bi-dimensional distributions. Description is identical to the left column. C17 data are not available for these quantities.

and that their dispersion bars overlap in the diagrams associated with FWHM(C IV) (bottom panels). In those panels, the median red lines are just *slightly above* the light turquoise lines (roughly 0.07 dex) in all diagrams. However, in those diagrams associated with FWHM $[$C IV/Mg II$]$ (middle panels), we can see a clearer separation between light turquoise and red lines. Particularly, red lines (FWHM (Mg II) > 4500 km s^{-1}) are now roughly 0.3 dex *under* the light turquoise lines (FWHM (Mg II) < 3000 km s^{-1}) in all diagrams. This indicates that FWHM(Mg II) is driving the dispersion in the correlation between FWHM $[$C IV/Mg II$]$ and the line peak ratio quantities.

To obtain further support for the previous finding, we looked at the residuals of the line peak ratios when expressed as a function of FWHM $[$C IV/Mg II$]$. In the case that FWHM(Mg II)

is driving the dispersion in the FWHM $[$C IV/Mg II$]$ correlations, we would find significant anticorrelation between these residuals and FWHM(Mg II). To address this, we fit the line peak ratios in terms of FWHM(C IV) and FWHM $[$C IV/Mg II$]$ using bisector linear regressions. We find that for the peak ratios as functions of FWHM $[$C IV/Mg II$]$, all the line-peak-ratio residuals are significantly anticorrelated with FWHM(Mg II) ($|r_p| > 0.41$ in both samples) as expected. Moreover, for the line peak ratios versus FWHM(C IV), we do not find any significant correlations between any of the residuals with FWHM(Mg II) ($|r_p| < 0.23$ in both samples).

An additional test consists of estimating the statistical significance of the difference between the correlation coefficients associated with FWHM(C IV) and those associated with
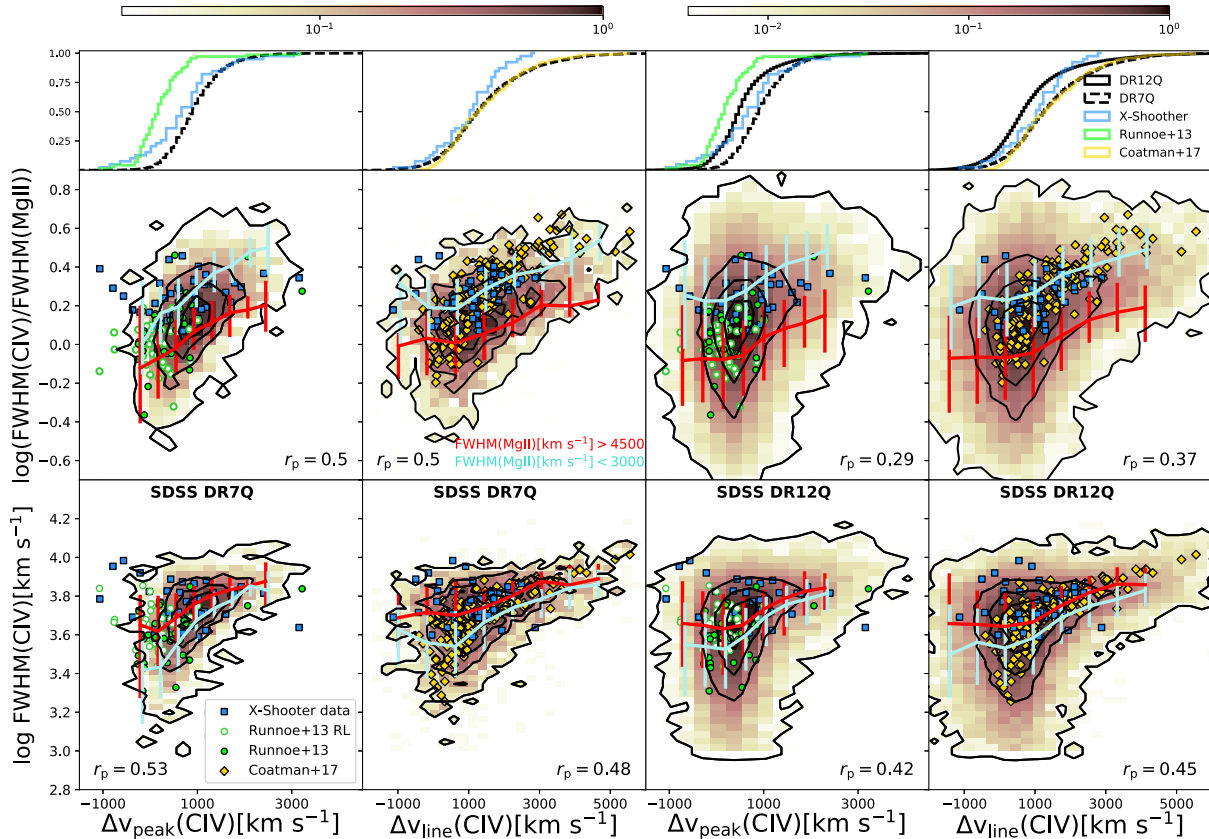
Figure 5. Bi-dimensional distributions of FWHM (C IV) /FWHM (Mg II) (top panels) and FWHM (C IV) (bottom panels) versus the C IV blueshift proxies Δv_{peak}(left) and Δv_{line} (right) in the SDSS DR7Q (two left columns) and DR12Q (two right columns) samples. The intensity of the colour represents the relative density of points as shown in the colour bar on the top. The black thin lines represent the 25 per cent, 50 per cent, 75 per cent, and 99 per cent contours centred at the maximum probability point. The small samples are superimposed as indicated in the legends. The projected CDFs of Δv_{peak} and Δv_{line} are also shown in the top diagrams. In C17 sample, there are not available measurements for the FWHM(Mg II). We then used 0.75FWHM(H α) as a proxy for FWHM(Mg II) where 0.75 represents the median value of FWHM (Mg II) /FWHM (H α) in the X-Shooter sample from Mejía-Restrepo et al. (2016). Coloured trend lines represent the median values of FWHM [C IV/Mg II] (middle panels) and FWHM(C IV) (bottom panels) as a function of the different blueshift estimators for objects with FWHM (Mg II) < 3000 km s^{-1} (light turquoise) and FWHM (Mg II) > 4500 km s^{-1} (red). The error bars represent the 1σ dispersion of the points around these trends. Measurements of the blueshift are less reliable in the DR12Q sample than in the DR7Q sample because of the redshift determination (see Section 2). We also show the correlation coefficient r_P. In all these cases, $P_P < 10^{-90}$ and are not shown in the panels.

FWHM [C IV/Mg II] in both SDSS quasar samples. The William's test, using Fisher-z transformations, provides a procedure to test the relative significance of the difference between two Pearson correlation coefficients obtained from the same sample and sharing one common variable (Dunn & Clark 1969). By applying this method to the correlations of FWHM(C IV) and FWHM [C IV/Mg II] with the common variable L_{peak} [C IV/SiO IV], in both SDSS quasar samples, we find an associated probability of $P_{William} < 10^{-5}$ for upholding the null hypothesis that both correlation coefficients are equal. This result discards the equivalence of both the FWHM(C IV)-L_{peak} [C IV/SiO IV] and the FWHM [C IV/Mg II]-L_{peak} [C IV/SiO IV] correlation coefficients. We find similar behaviours for the case of the FWHM(C IV)-L_{peak} [C IV/1450 Å] and FWHM [C IV/Mg II]-L_{peak} [C IV/1450 Å] correlations where we find $P_{William} < 10^{-14}$ in both samples. Finally, for the case of the FWHM(C IV)-L_{peak} [C IV/C III] and FWHM [C IV/Mg II]-L_{peak} [C IV/C III] correlations, we find $P_{William} = 10^{-9}$ in the DR12Q sample, while for the DR7 samples the correlation coefficients are identical (see Fig. 3).

From all the evidence that we have collected, we can conclude that the prescriptions proposed by Runnoe et al. (2013) and Mejía-Restrepo et al. (2016) are of limited applicability for cor-

recting C IV-based estimates of M_{BH} because the correlations between the line peak ratios and FWHM(C IV) are statistically stronger and very likely driving the weaker correlations associated with FWHM [C IV/Mg II].

3.1.2 C IV blueshifts

We continue to test whether or not the Δv_{line} (C IV) proposed by Coatman et al. (2017) can be used to improve C IV-based measurements. In addition to Δv_{line} (C IV), we will also include Δv_{peak} (C IV) in our analysis. The reason for this choice is the better stability of Δv_{peak} (C IV) to continuum placement as we discussed in Section 2.

In Fig. 5, we show the bi-dimensional distribution of FWHM [C IV/Mg II] and FWHM(C IV) versus the C IV blueshift indicators Δv_{peak} (C IV) and Δv_{line} (C IV) in both SDSS samples. In each panel, we map the same dynamic range for FWHM(C IV) and FWHM [C IV/Mg II]. We also present the Pearson correlation coefficients for each diagram.

Fig. 5 demonstrates that FWHM [C IV/Mg II] and FWHM(C IV) are both correlated with Δv_{peak} (C IV) and Δv_{line} (C IV) in both SDSS samples. It is also noticeable that in most cases, the cor-
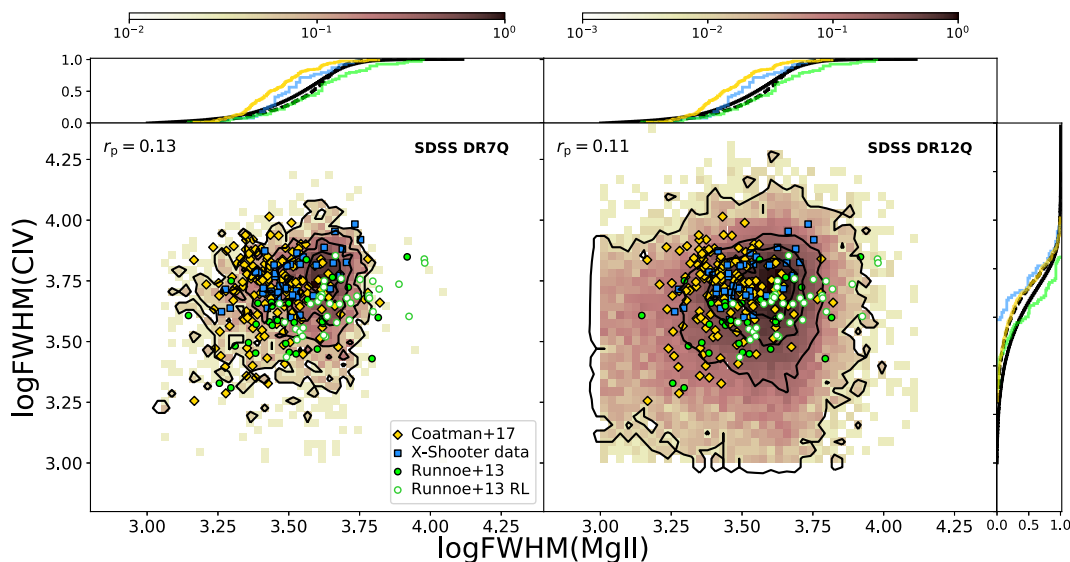
Figure 2. FWHM(Mg II)-FWHM(C IV) bi-dimensional distributions in the SDSS DR7Q (left) and DR12Q (right) samples. The intensity of the colour represents the relative density of points as shown in the colour bars on the right. The black thin lines represent the 25 per cent, 50 per cent, 75 per cent, and 99 per cent contours centred at the maximum probability point. In the top and right side diagrams, we show the projected CDFs of FWHM(Mg II) and FWHM(C IV), respectively. We superimpose in each panel analogue data from the small X-Shooter, R13, and C17 samples as in Fig. 1. R13 is differentiated by RL and RQ objects in the scatter plots. We also show the Pearson correlation coefficient (r_P) for the DR7Q and DR12Q bi-dimensional distribution.
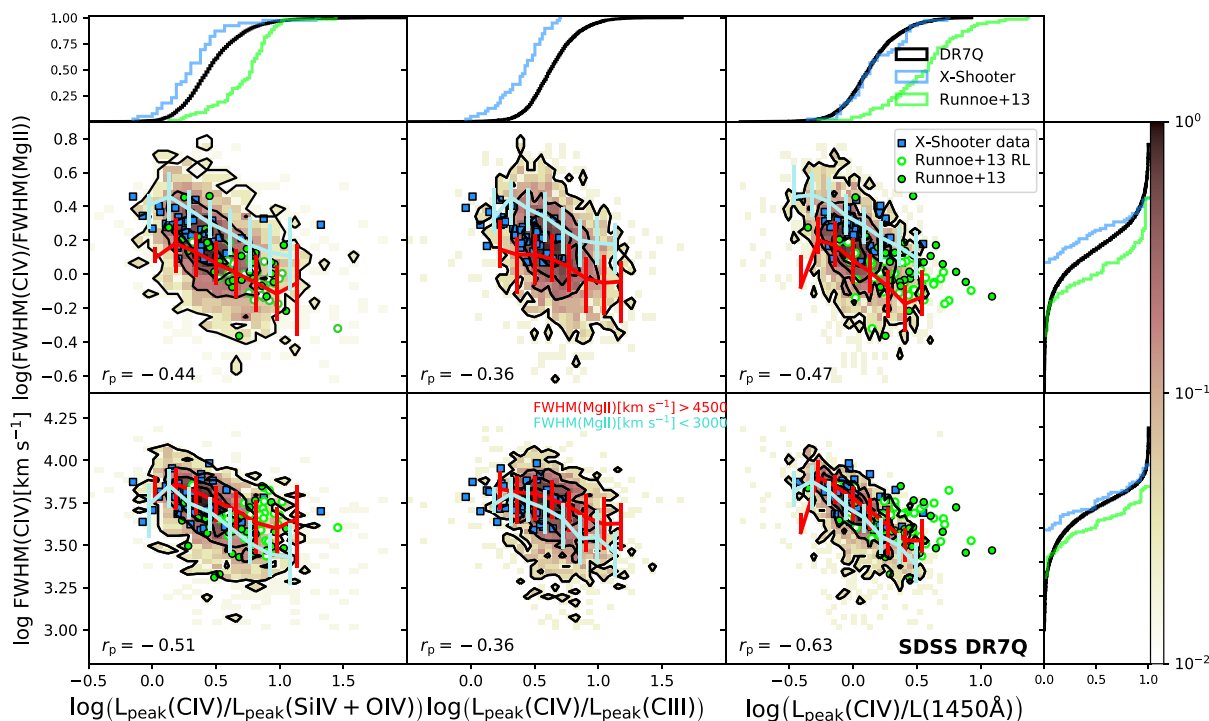

Figure 3. Bi-dimensional distributions of FWHM [C IV/Mg II] (top row) and FWHM(C IV) (bottom row) versus L_{peak} [C IV/SiO IV] (left column), L_{peak} [C IV/C III] (centre column), and L_{peak} [C IV/1450 Å] (right column) in the **SDSS DR7Q** sample. The intensity of the colour represents the relative density of points as shown in the colour bar on the right. The black thin lines represent the 25 per cent, 50 per cent, 75 per cent, and 99 per cent contours centred at the maximum probability point. The projected CDFs of each of the quantities are also shown in the right and top side diagrams. We superimpose to each panel analogue data of the X-Shooter and R13 small samples as indicated in the legends. Coloured trend lines represent the median values of FWHM [C IV/Mg II] (middle panels) and FWHM(C IV) (bottom panels) as a function of the different line peak ratios for objects with FWHM (Mg II) < 3000 km s^{-1} (light turquoise) and FWHM (Mg II) > 4500 km s^{-1} (red). The error bars represent the 1σ dispersion of the points around these trends. Note that the dynamic range that is shown for FWHM(C IV) and FWHM(Mg II) coincides (1.6dex). The same situation occurs with the dynamic range that is shown for L_{peak} [C IV/SiO IV], L_{peak} [C IV/C III], and L_{peak} [C IV/1450 Å] (2.5dex). We also show the correlation coefficient r_P. In all these cases, $P_P <\, <10^{-8}$ and are not shown in the panels.
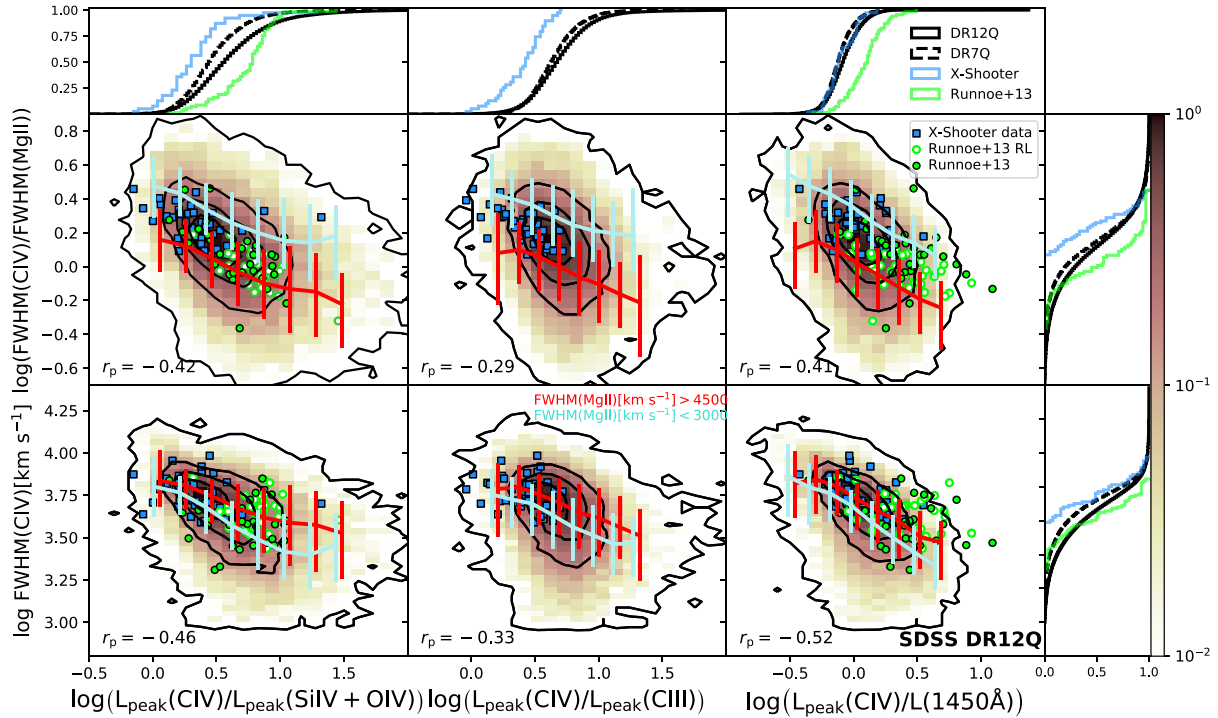
Figure 4. As in Fig. 3, but for the **SDSS DR12Q** sample.

Table 1. Scatter found in correlations between the listed quantities in the DR7Q and DR12Q samples.

| | Scatter | | | |
| | FWHM $[C\,IV/Mg\,II]$ | | FWHM(C IV) | |
	DR7Q	DR12Q	DR7Q	DR12Q
$L_{peak}\,[C\,IV/Si\,IV]$	0.21	0.27	0.17	0.21
$L_{peak}\,[C\,IV/C\,III]]$	0.22	0.29	0.19	0.23
$L_{peak}\,[C\,IV/1450\,\text{Å}]$	0.20	0.25	0.14	0.19
Δv_{peak} (C IV)	0.19	0.28	0.16	0.22
Δv_{line} (C IV)	0.21	0.29	0.19	0.24

relations between both blueshift estimators and FWHM(C IV) are tighter than with FWHM $[C\,IV/Mg\,II]$. The only exception is with Δv_{line} (C IV) in the DR7Q sample where both correlations show similar significance. We can also notice in Table 1 that the scatter of the FWHM(C IV) correlations is smaller than the scatter in the corresponding FWHM $[C\,IV/Mg\,II]$ correlations in both SDSS samples. These results would indicate that the correlations associated with FWHM(C IV) are driving the correlations associated with FWHM $[C\,IV/Mg\,II]$, similarly to what we found in Section 3.1.1.

We repeated the same three tests described in Section 3.1.1 to further check the reliability of these results. First, when dividing the samples into two subsets according to their FWHM(Mg II) and binning by the blueshift indicators in bins of 700 km s^{-1}, we find that the separation of the median trends of the narrow group from the median trends of the broad group (light turquoise and red lines in Fig. 5) is increased from roughly −0.08 dex in the FWHM(C IV) diagrams to roughly 0.25 dex in the FWHM $[C\,IV/Mg\,II]$ diagrams. Secondly, the residuals of the blueshift indicators expressed as a function of FWHM $[C\,IV/Mg\,II]$ show significant anticorrelations with FWHM(Mg II) ($|r_s| > 0.36$ in both samples). In contrast, when the blueshift indicators are expressed as a function of FWHM(C IV),

we find no correlations with FWHM(Mg II) ($|r_s| < 0.10$ in both samples). Finally, the relative significance test shows that FWHM(C IV) correlations are indeed stronger than the FWHM $[C\,IV/Mg\,II]$ correlations ($P_{William} < 0.006$) for Δv_{peak} (C IV) and Δv_{line} (C IV) in the DR12Q sample and for Δv_{peak} (C IV) in the DR7Q sample. In the case of Δv_{line} (C IV) for the DR7 sample, we find that its Pearson correlation coefficients with FWHM(C IV) and FWHM $[C\,IV/Mg\,II]$ (0.48 and 0.50, respectively) are not statistically different from each other ($P_{William} = 0.1$, smaller than 2σ significance). All these results support the idea that the correlations between the C IV blueshifts and FWHM(C IV) are the main drivers for the correlations with FWHM $[C\,IV/Mg\,II]$. This, in turn, suggests that the prescription introduced by Coatman et al. (2017) may have limited applicability to improve C IV mass measurements.

3.2 Small samples

We continue our analysis exploring the small samples described in Section 2. Below we analyse the behaviour of the line peak ratios and the blueshift relations with FWHM(C IV) and FWHM $[C\,IV/Mg\,II]$ in those samples and discuss the similarities and differences with respect to our findings in the large SDSS samples.

3.2.1 Line peak ratios

Given that line peak information is only available for the X-Shooter and R13 samples, we limit the analysis of the line peak ratios to these two samples. In addition to the SDSS data, in Figs 3 and 4, we also show the data points and distribution functions associated with the X-Shooter (light blue squares) and R13 samples (lime open dots and lime filled dots for RL and RQ objects, respectively).

In Table 2, we show the correlation coefficients of the line peak quantities versus (1) FWHM(C IV) and (2) FWHM $[C\,IV/Mg\,II]$ in the following configurations:

Table 2. Pearson correlation coefficients of (1) log FWHM (C IV) and (2) log FWHM [C IV/Mg II] versus the line peak quantities in (*a*) the X-Shooter sample, (*b*) the R13 sample, (*c*) the combination of the X-Shooter and R13 samples, and (*d*) the combination of the X-Shooter and R13 samples excluding the RL objects. L_{peak} [C IV/C III]] measurements are not available for the R13 sample.

| | log FWHM (C IV) | | | | | | | | log FWHM [C IV/Mg II] | | | | | | | |
| | X-Shooter[a] | | R13[b] | | Both[c] | | Both RQ[d] | | X-Shooter[a] | | R13[b] | | Both[c] | | Both RQ[d] | |
	r_P	P_P	r_P	P_P	r_P	P_P	r_P	P_P	r_P	P_P	r_P	P_P	r_P	P_P	r_P	P_P
log L_{peak} [C IV/SiO IV]	-0.21	2×10^{-1}	-0.04	8×10^{-1}	-0.45	1×10^{-8}	-0.54	1×10^{-6}	-0.32	4×10^{-2}	-0.60	3×10^{-8}	-0.65	1×10^{-11}	-0.54	1×10^{-6}
log L_{peak} [C IV/1450 Å]	-0.48	2×10^{-3}	-0.30	1×10^{-2}	-0.53	3×10^{-9}	-0.60	3×10^{-8}	-0.40	1×10^{-2}	-0.36	2×10^{-3}	-0.59	2×10^{-11}	-0.59	8×10^{-8}
log L_{peak} [C IV/C III]]	-0.23	2×10^{-1}	–	–	–	–	–	–	-0.60	4×10^{-5}	–	–	–	–	–	–

(i) The individual X-Shooter and R13 samples (subsamples *a* and *b*).

(ii) The combination of the X-Shooter and the R13 samples including RL objects (subsample *c*).

(iii) The combination of the X-Shooter and the R13 samples excluding RL objects (subsample *d*).

We remark that both the X-Shooter and the R13 samples are not complete. Indeed, as discussed in Appendix B, the two samples are mapping totally different regions in the parameter spaces determined by (1) the line peak quantities versus FWHM [C IV/Mg II] and by (2) the line peak quantities versus FWHM(C IV) (see also Figs 3 and 4). Consequently, after the exclusion of the RL objects from the R13 sample, the combination of the X-Shooter and the R13 samples (i.e. subsample *d*) maps the parameter space of RQ type-1 AGN over wider ranges in L_{1450}, FWHM(C IV), FWHM [C IV/Mg II], and the line peak quantities.

The correlation test presented in Table 2 suggests that for the individual samples (*a* and *b*) in most cases, the correlations of FWHM [C IV/Mg II] with the line peak quantities are tighter than those associated with FWHM(C IV), in contrast to what we find for the large SDSS samples. The same behaviour is found for the combination of both samples including the RL objects (subsample *c*). However, when the RL objects are excluded from the analysis (subsample *d*), the FWHM [C IV/Mg II] and FWHM(C IV) correlation coefficients are statically indistinguishable. Thus, the results from subsample *d*, which better maps our parameter space, are in agreement with our results from the large SDSS samples. This may indicate that the large fraction of RL objects in the R13 sample (37/69) are probably artificially strengthening the correlations associated with FWHM [C IV/Mg II]. This is probably caused by relatively broad Mg II profiles shown by RL objects (log FWHM (Mg II) [km s^{-1}] $\gtrsim$ 3.5, see right-hand panel of Fig. 1).

3.2.2 C IV blueshifts

Before reporting the results of our comparative analysis for the C17 sample, we note that this sample leans towards high-luminosity sources as can be seen in Fig. 1. Nonetheless, its Δv_{line} (C IV) and FWHM(C IV) distribution are in very good agreement with the SDSS DR7Q sample (see AppendixB and Figs 1 and 5 for details).

Using the results reported in Coatman et al. (2017), we find that in their sample, Δv_{line} (C IV) is very tightly correlated with FWHM(C IV) ($r_p = 0.82$). However, we also find that the Δv_{line} (C IV)-FWHM [C IV/Hα] Pearson correlation coefficient is essentially equal ($r_p = 0.83$). It is also remarkable that fig. 9 in Coatman et al. (2017) shows that the scatter in the FWHM [C IV/Hα] versus Δv_{line} (C IV) correlation is clearly dominated by FWHM(Hα). These results support our hypothesis that the FWHM(CIV)-Δv_{line} (C IV) correlation is the driver of the FWHM [C IV/Hα]-Δv_{line} (C IV) correlation.

3.3 Resampling tests

Here we present different tests designed to check the validity of the findings presented above. They consist of re-sampling our SDSS DR7Q and DR12Q samples in four different ways:

(i) Flat distribution in log L_{1450}.
(ii) Flat distribution in log FWHM.
(iii) Flat distribution in log FWHM [C IV/Mg II].
(iv) Flat distribution in both log L_{1450} and log FWHM simultaneously.

The motivation for these tests is to check whether our findings are biased by the concentrated distribution in L_{1450}, FWHM(C IV), and FWHM [C IV/Mg II] and/or the known correlation between L_{1450} and FWHM(C IV) that we observe in Fig. 1 (see also Appendix B). To this end, we first divided our SDSS samples in bins of 0.5 dex in L_{1450} starting at log L_{1450} = 45.0 erg s^{-1} in the DR7Q sample and at log L_{1450} = 44.5 erg s^{-1} in the DR12Q sample. For FWHM(C IV) and FWHM [C IV/Mg II], we divided our sample in bins of 0.4 dex. To guarantee an equal number of objects in each bin, we selected 23 objects from the DR7Q sample and 100 objects from the DR12Q sample, in each realization of the re-sampling simulation. We finally subdivided our SDSS samples in bi-dimensional bins of L_{1450} and FWHM(C IV) of 0.5 and 0.4 dex, respectively. For each bin, we selected 13 and 30 objects in the DR7Q and DR12Q samples. In this case, the total number of sources in each re-sampling test is comparable with the size of the small X-Shooter, R13, and C17 samples. To account for statistical variance because of the limited sampling, we repeated these procedures 100 times. We find the FWHM(CIV)-associated correlations show larger or equivalent statistical significance than the FWHM [C IV/Mg II] correlations as suggested by the William's method (see Table 3 for details), which is consistent with what we found for the large SDSS samples.

3.4 Signal-to-noise analysis

We also considered the possibility that our results may be affected by the limited quality of the SDSS spectroscopic data. To this end, we selected a high-quality subsample of the SDSS DR12Q catalogue consisting of objects with S/N > 10 at 1700 Å and at 3000 Å with a binning of 0.75 Å/pixel, following Denney et al. (2013). We found a total of 2230 objects meeting these criteria.

As can be seen in Table A2, our analysis on this high-quality subsample is consistent with our central findings on the entire SDSS DR12Q sample. In particular, we find that the correlations connecting the line peak quantities and the CIV blueshifts with FWHM(C IV) are stronger than those with FWHM [C IV/Mg II]. The only exception is with Δv_{line} (C IV), where both correlations are of comparable strength. Similarly, in Table A2 we also show that when we further limit our analysis to the 483 objects with S/N > 20, the correlations related to FWHM(CIV) are always stronger than those related to FWHM [C IV/Mg II].

Table 3. Absolute values of the Pearson correlation coefficients ($|r_p|$) for the quantities in the first column versus (1) log FWHM (C IV) and (2) log FWHM [C IV/Mg II]. The central values correspond to the medians obtained from the 100 randomly generated subsamples selected to have flat distributions in L_{1450} (*a*), FWHM(C IV) (*b*), FWHM [C IV/Mg II] (*c*), and L_{1450}-FWHM(C IV) (*d*) from the SDSS DR7Q and DR12Q samples. Errors correspond to the central 68 per cent of the $|r_p|$ distribution.

Random sampling ⇒	$L_{1450}{}^a$				FWHM(C IV)^b				FWHM [C IV/Mg II]^c				L_{1450}-FWHM(C IV)^d			
	DR12Q		DR7Q		DR12Q		DR7Q		DR12Q		DR7Q		DR12Q		DR7Q	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
L_{peak} [C IV/SiO IV]	$0.63^{+0.03}_{-0.03}$	$0.55^{+0.03}_{-0.04}$	$0.59^{+0.10}_{-0.11}$	$0.52^{+0.09}_{-0.09}$	$0.58^{+0.05}_{-0.04}$	$0.55^{+0.05}_{-0.05}$	$0.61^{+0.08}_{-0.06}$	$0.54^{+0.08}_{-0.08}$	$0.58^{-0.06}_{-0.04}$	$0.57^{+0.04}_{-0.05}$	$0.62^{+0.08}_{-0.10}$	$0.57^{+0.08}_{-0.08}$	$0.49^{+0.02}_{-0.02}$	$0.47^{+0.02}_{-0.02}$	$0.61^{+0.04}_{-0.04}$	$0.56^{+0.05}_{-0.04}$
L_{peak} [C IV/C III]]	$0.37^{+0.04}_{-0.04}$	$0.37^{+0.04}_{-0.04}$	$0.37^{+0.11}_{-0.11}$	$0.37^{+0.11}_{-0.12}$	$0.42^{+0.05}_{-0.06}$	$0.41^{+0.04}_{-0.05}$	$0.44^{+0.11}_{-0.11}$	$0.42^{+0.11}_{-0.12}$	$0.42^{+0.06}_{-0.04}$	$0.37^{+0.05}_{-0.07}$	$0.49^{+0.05}_{-0.13}$	$0.49^{+0.09}_{-0.10}$	$0.36^{+0.03}_{-0.02}$	$0.36^{+0.02}_{-0.03}$	$0.46^{+0.05}_{-0.06}$	$0.42^{+0.05}_{-0.05}$
L_{peak} [C IV/1450 Å]	$0.70^{+0.03}_{-0.02}$	$0.56^{+0.03}_{-0.04}$	$0.72^{+0.06}_{-0.08}$	$0.52^{+0.09}_{-0.11}$	$0.62^{+0.06}_{-0.04}$	$0.52^{+0.05}_{-0.04}$	$0.67^{+0.06}_{-0.08}$	$0.53^{+0.05}_{-0.06}$	$0.59^{+0.05}_{-0.05}$	$0.49^{+0.05}_{-0.05}$	$0.73^{+0.07}_{-0.09}$	$0.61^{+0.09}_{-0.09}$	$0.57^{+0.01}_{-0.02}$	$0.49^{+0.02}_{-0.02}$	$0.69^{+0.03}_{-0.04}$	$0.54^{+0.05}_{-0.04}$
Δv_{peak} (C IV)	$0.48^{+0.04}_{-0.05}$	$0.42^{+0.02}_{-0.03}$	$0.60^{+0.09}_{-0.11}$	$0.56^{+0.06}_{-0.09}$	$0.44^{+0.05}_{-0.06}$	$0.37^{+0.05}_{-0.05}$	$0.62^{+0.08}_{-0.09}$	$0.58^{+0.06}_{-0.10}$	$0.42^{+0.06}_{-0.05}$	$0.34^{+0.06}_{-0.06}$	$0.67^{+0.06}_{-0.07}$	$0.64^{+0.06}_{-0.06}$	$0.48^{+0.03}_{-0.02}$	$0.41^{+0.03}_{-0.02}$	$0.60^{+0.04}_{-0.04}$	$0.57^{+0.04}_{-0.04}$
Δv_{line} (C IV)	$0.62^{+0.03}_{-0.03}$	$0.54^{+0.04}_{-0.04}$	$0.54^{+0.08}_{-0.09}$	$0.56^{+0.06}_{-0.09}$	$0.42^{+0.05}_{-0.06}$	$0.41^{+0.06}_{-0.06}$	$0.55^{+0.09}_{-0.09}$	$0.55^{+0.09}_{-0.08}$	$0.42^{+0.05}_{-0.04}$	$0.40^{+0.05}_{-0.05}$	$0.63^{+0.07}_{-0.10}$	$0.65^{+0.06}_{-0.08}$	$0.46^{+0.03}_{-0.02}$	$0.45^{+0.04}_{-0.03}$	$0.52^{+0.05}_{-0.05}$	$0.56^{+0.05}_{-0.05}$

The usage of these high-S/N subsamples also allows us to test the results of several previous studies that suggested that, in high-quality spectra, the σ_{line} of C IV provides more accurate M_{BH} estimates than the FWHM of C IV. Because of the lack of H β measurements in our sample, we used the FWHM and the σ_{line} of Mg II as proxies for the H β measurements. In contrast to the results of Denney et al. (2013), we found weak or even no correlations between the σ_{line} of C IV and, the σ_{line} and FWHM of MgII in both of the high-quality subsamples (i.e. those with S/N > 10 and/or >20; correlation coefficients in the range $0.04 < r_p < 0.19$). This indicates that the σ_{line} of C IV cannot be reliably used to provide accurate estimates of M_{BH} even in high-quality data, mainly because of its instability to the continuum placement.

3.5 Principal component analysis

We conducted a principal component analysis on the correlation matrices presented in Table A3 to find different groups of interconnected variables and obtain the amount of variance driven by each group. In Table 4, we show the correlation coefficients between the first three proper vectors and the quantities that define them. We can observe that the first proper vector (PV1) is responsible for 38 per cent and 32 per cent of the variance in the DR7Q and DR12Q samples, respectively. In both cases, FWHM(C IV) and L_{peak} [C IV/1450 Å] show the strongest correlations with PV1, indicating that the FWHM(C IV)-L_{peak} [C IV/1450 Å] anticorrelation drives PV1 and consequently a large percentage of the variance in the SDSS samples. FWHM [C IV/Mg II], L_{1450}, L_{peak} [C IV/SiO IV], L_{peak} [C IV/C III]], Δv_{peak} (C IV), Δv_{line} (C IV), and EW(C IV) and L_{peak} (C IV) $/L$ (C IV) also show important correlations with PV1. However, these correlations are basically caused by the strong relations of these quantities with FWHM(C IV) and L_{peak} [C IV/1450 Å].

The second proper vector (PV2) is responsible for 16 per cent and 15 per cent of the variance in the DR7Q and DR12Q samples, respectively, and is strongly correlated with FWHM(Mg II) in both samples. It also shows a strong correlation with FWHM (Mg II) $/\sigma$ (Mg II) and a strong anticorrelation with FWHM [C IV/Mg II], which is basically inherited from their correlations with FWHM(Mg II). Given that by definition PV2 is linearly independent of the other proper vectors, this result reveals that FWHM(Mg II) is basically independent of any C IV-related quantity and may indicate that C IV and Mg II profiles show completely independent behaviours.

Finally, the third proper vector (PV3) drives 13 per cent and 14 per cent of the variance of the SDSS samples, respectively, and is strongly correlated with EW(C IV) and L_{peak} (C IV) $/L$ (C IV), which are both strongly correlated with each other because of their dependence on L (C IV). It also shows an important correlation with FWHM [C IV/Mg II], which is not correlated with any of these quan-

tities. Thus, PV3 does not provide a link between C IV and Mg II properties.

4 SUMMARY AND CONCLUSIONS

C IV-based M_{BH} estimations are known to be problematic. In the past few years, Runnoe et al. (2013), Mejía-Restrepo et al. (2016), and Coatman et al. (2017) provided alternative methods attempting to improve C IV-based masses. All these methods were based on correlations between different observables associated with the C IV emission and the ratio of FWHM(C IV) and the FWHM of low ionization lines (i.e. H α, H β, and MgII). Despite the good quality of the data used in these works, all these methods were derived using small samples with limited coverage of the parameter space of the observables involved in each method.

Using SDSS DR7Q and DR12Q samples (which are more representative of the quasar population), we showed that all these methods are of limited applicability to improve C IV-based M_{BH} estimations. In fact, we find that the aforementioned methods depend on correlations that are actually driven by the FWHM of the C IV profile itself and IVnot by an interconnection between FWHM(C IV) and the FWHMs of the low ionization lines. Additionally, our analysis suggests that from all the correlations that we considered with FWHM(C), those that involve L_{peak} [C IV/1450 Å] are the tightest ones. We also find that other quantities considered in this work (L_{peak} [C IV/SiO IV], L_{peak} [C IV/C III]], Δv_{peak} (C IV), and Δv_{line} (C IV)) are all tightly correlated with L_{peak} [C IV/1450Å] (see Table A3).

Further support for these conclusions comes from principal component analysis, which reveals that the PV1 is mostly driven by the anticorrelation between FWHM(C IV) and L_{peak} [C IV/1450 Å]. This occurs in such a way that the relations between these quantities and FWHM [C IV/Mg II], L_{1450}, EW(C IV), L_{peak} [C IV/SiO IV], L_{peak} [C IV/C III]], Δv_{peak} (C IV), and Δv_{line} (C IV) are basically driven by the FWHM(C IV)-L_{peak} [C IV/1450 Å] anticorrelation. Notably, the PV2 is mostly driven by FWHM(MgII) and shows no correlation with any C IV-related quantity. This suggests that the properties of the MgII and C IV profiles are independent of each other. In other words, there is no possibility to relate the nonvirialized C IV emission with the virialized Mg II emission.

A possible explanation for this could be associated with the fact that the more luminous a quasar is, the lower its EW(C IV). This is the well-known C IV Baldwin effect (Baldwin 1977; Baskin & Laor 2004, 2005; Richards et al. 2011; Ge et al. 2016). Both the quasar luminosity and EW(C IV) are known to be related with the C IV blueshift, the C IV asymmetry, and the relative strength of the X-ray emission (e.g. Richards et al. 2011; Shen et al. 2016). Indeed, if we take L_{peak} [C IV/1450 Å] as a proxy for the EW of C IV and consider the anticorrelation between FWHM(C IV) and L_{1450}, we can

Table 4. Pearson correlation coefficients between the listed quantities and the three first proper vectors (PV1, PV2, and PV3) obtained from principal component analysis for the DR7Q and DR12Q correlation matrices shown in Table A3.

| | DR7 | | | DR12 | | |
| | PV1 | PV2 | PV3 | PV1 | PV2 | PV3 |
Cumulative variance	38 per cent	54 per cent	67 per cent	32 per cent	47 per cent	61 per cent
$\log$ FWHM (C IV)	-0.81	0.27	-0.23	-0.75	0.21	0.25
$\log$ FWHM (Mg II)	0.07	0.86	0.33	0.10	0.87	-0.30
$\log$ FWHM $[$C IV/Mg II$]$	-0.72	-0.39	-0.43	-0.68	-0.45	0.41
$\log L_{1450}$	-0.48	0.25	-0.17	-0.62	0.22	-0.01
Δv_{peak} (C IV)	0.73	0.09	0.09	0.64	-0.04	0.01
Δv_{line} (C IV)	0.74	0.15	0.15	0.71	0.01	-0.04
$\log L_{peak} [$C IV/1450 Å$]$	0.86	-0.04	-0.37	0.79	0.02	0.33
$\log L_{peak} [$C IV/SiO IV$]$	0.77	0.05	-0.18	0.77	0.00	0.13
$\log L_{peak} [$C IV/C III]$]$	0.60	0.25	-0.45	0.61	0.04	0.30
$\log$ EW(C IV)	0.47	0.27	-0.74	0.27	0.38	0.84
$\log \left(L_{peak} (\text{C IV}) / L (\text{C IV}) \right)$	0.74	-0.40	0.37	0.45	-0.43	-0.66
$\log (\text{FWHM (Mg II)} / \sigma (\text{Mg II}))$	-0.13	0.76	0.31	-0.05	0.76	-0.30
$\log(L_{3000}/L_{1450})$	0.14	-0.41	0.40	-0.13	-0.03	-0.19

conclude that the very tight anticorrelation between FWHM(C IV) and $L_{peak} \left[\text{C IV/1450 Å} \right]$ can be seen as inherited from the Baldwin effect.

Our analysis implies that the well-characterized M_{BH} estimations from the low ionization lines cannot be accurately predicted from the emission line properties of the C IV line. Consequently, systematic infrared spectroscopic observations of large samples of quasars are required to guarantee the coverage of low ionization lines and the proper determination of the SMBH masses at $z \gtrsim 2$. Achieving accurate C IV mass estimations requires, apart from a robust determination of the R_{BLR} (C IV)–L_{1450} relation, an extensive analysis of the C IV emission line itself to further understand the virialized component of the C IV line.

REFERENCES

Assef R. J. et al., 2011, ApJ, 742, 93
Baldwin J. A., 1977, ApJ, 214, 679
Baskin A., Laor A., 2004, MNRAS, 350, L31
Baskin A., Laor A., 2005, MNRAS, 356, 1029
Batiste M., Bentz M. C., Raimundo S. I., Vestergaard M., Onken C. A., 2017, ApJ, 838, L10
Bentz M. C., Peterson B. M., Netzer H., Pogge R. W., Vestergaard M., 2009, ApJ, 697, 160
Bentz M. C. et al., 2013, ApJ, 767, 149
Brotherton M. S., Runnoe J. C., Shang Z., DiPompeo M. A., 2015, MNRAS, 451, 1290
Capellupo D. M., Netzer H., Lira P., Trakhtenbrot B., Mejía-Restrepo J., 2015, MNRAS, 446, 3427
Capellupo D. M., Netzer H., Lira P., Trakhtenbrot B., Mejía-Restrepo J., 2016, MNRAS, 460, 212

Coatman L., Hewett P. C., Banerji M., Richards G. T., 2016, MNRAS, 461, 647
Coatman L., Hewett P. C., Banerji M., Richards G. T., Hennawi J. F., Prochaska J. X., 2017, MNRAS, 465, 2120
Collin S., Kawaguchi T., Peterson B. M., Vestergaard M., 2006, A&A, 456, 75
Corbin M. R., Boroson T. A., 1996, ApJS, 107, 69
Denney K. D., 2012, ApJ, 759, 44
Denney K. D., Pogge R. W., Assef R. J., Kochanek C. S., Peterson B. M., Vestergaard M., 2013, ApJ, 775, 60
Dunn O. J., Clark V., 1969, J. Am. Stat. Assoc., 64, 366
Fine S., Croom S. M., Bland-Hawthorn J., Pimbblet K. A., Ross N. P., Schneider D. P., Shanks T., 2010, MNRAS, 409, 591
Ganguly R., Brotherton M. S., Cales S., Scoggins B., Shang Z., Vestergaard M., 2007, ApJ, 665, 990
Ge X., Bian W.-H., Jiang X.-L., Liu W.-S., Wang X.-F., 2016, MNRAS, 462, 966
Greene J. E., Ho L. C., 2005, ApJ, 630, 122
Hewett P. C., Wild V., 2010, MNRAS, 405, 2302
Ho L. C., Goldoni P., Dong X.-B., Greene J. E., Ponti G., 2012, ApJ, 754, 11
Kaspi S., Smith P. S., Netzer H., Maoz D., Jannuzi B. T., Giveon U., 2000, ApJ, 533, 631
Kaspi S., Maoz D., Netzer H., Peterson B. M., Vestergaard M., Jannuzi B. T., 2005, ApJ, 629, 61
Kaspi S., Brandt W. N., Maoz D., Netzer H., Schneider D. P., Shemmer O., 2007, ApJ, 659, 997
Labita M., Decarli R., Treves A., Falomo R., 2009, MNRAS, 399, 2099
McLure R. J., Dunlop J. S., 2004, MNRAS, 352, 1390
McLure R. J., Jarvis M. J., 2002, MNRAS, 337, 109
Mejía-Restrepo J. E., Trakhtenbrot B., Lira P., Netzer H., Capellupo D. M., 2016, MNRAS, 460, 187
Mejía-Restrepo J. E., Lira P., Netzer H., Trakhtenbrot B., Capellupo D. M., 2017, Nat. Astron., 2, 63
Netzer H., Trakhtenbrot B., 2007, ApJ, 654, 754
Netzer H., Lira P., Trakhtenbrot B., Shemmer O., Cury I., 2007, ApJ, 671, 1256
Onken C. A., Kollmeier J. A., 2008, ApJ, 689, L13
Onken C. A., Ferrarese L., Merritt D., Peterson B. M., Pogge R. W., Vestergaard M., Wandel A., 2004, ApJ, 615, 645
Pâris I. et al., 2017, A&A, 597, A79
Park D. et al., 2012, ApJ, 747, 30
Park D., Woo J.-H., Denney K. D., Shin J., 2013, ApJ, 770, 87
Peterson B. M. et al., 2004, ApJ, 613, 682
Peterson B. M. et al., 2005, ApJ, 632, 799
Rafiee A., Hall P. B., 2011, ApJS, 194, 42

Richards G. T., Vanden Berk D. E., Reichart T. A., Hall P. B., Schneider D. P., SubbaRao M., Thakar A. R., York D. G., 2002, AJ, 124, 1

Richards G. T. et al., 2011, AJ, 141, 167

Runnoe J. C., Brotherton M. S., Shang Z., DiPompeo M. A., 2013, MNRAS, 434, 848

Schlegel D., White M., Eisenstein D., 2009, in astro2010: The Astronomy and Astrophysics Decadal Survey, National Academies Press, Washingtong, DC, p. 314

Schneider D. P. et al., 2010, AJ, 139, 2360

Shang Z., Wills B. J., Wills D., Brotherton M. S., 2007, AJ, 134, 294

Shankar F. et al., 2016, MNRAS, 460, 3119

Shen Y., 2013, Bull. Astron. Soc. India, 41, 61

Shen Y., Liu X., 2012, ApJ, 753, 125

Shen Y., Greene J. E., Strauss M. A., Richards G. T., Schneider D. P., 2008, ApJ, 680, 169

Shen Y. et al., 2016, ApJ, 831, 7

Tang B., Shang Z., Gu Q., Brotherton M. S., Runnoe J. C., 2012, ApJS, 201, 38

Tilton E. M., Shull J. M., 2013, ApJ, 774, 67

Trakhtenbrot B., Netzer H., 2012, MNRAS, 427, 3081

Trakhtenbrot B., Netzer H., Lira P., Shemmer O., 2011, ApJ, 730, 7

Trevese D., Perna M., Vagnetti F., Saturni F. G., Dadina M., 2014, ApJ, 795, 164

Vestergaard M., Osmer P. S., 2009, ApJ, 699, 800

Vestergaard M., Peterson B. M., 2006, ApJ, 641, 689

Wang J.-G. et al., 2009, ApJ, 707, 1334

Woo J.-H., Schulze A., Park D., Kang W.-R., Kim S. C., Riechers D. A., 2013, ApJ, 772, 49

Woo J.-H., Yoon Y., Park S., Park D., Kim S. C., 2015, ApJ, 801, 38

Xiao T., Barth A. J., Greene J. E., Ho L. C., Bentz M. C., Ludwig R. R., Jiang Y., 2011, ApJ, 739, 28

York D. G. et al., 2000, AJ, 120, 1579

APPENDIX A: FITTING PROCEDURE, MEASUREMENTS, AND BALQSO EXCLUSION

Our broad emission line modelling follows the procedure presented in Mejía-Restrepo et al. (2016) and Trakhtenbrot & Netzer (2012). Very briefly, the most prominent lines (Si IV+O IV], C IV, C III], and Mg II) are modelled using two Gaussian components, while other weak emission lines are modelled with a single Gaussian (including He II1640, N IV1718, Si III]1892). The central wavelength of each Gaussian component is restricted to move within $1000 \, \text{km s}^{-1}$ around the laboratory central wavelength. The C IV and He II1640 are allowed to be blue shifted up to $5000 \, \text{km s}^{-1}$.

We automatized the procedure by introducing some additional steps to the line and continuum fitting. We first proceed to fit and subtract the continuum emission within a pair of continuum windows around each line. These continuum windows are set at the wavelengths that we list in Table A1. We subsequently fit the emission line following the 'local' approach described in Mejía-Restrepo et al. (2016). After this, we obtain the residuals of the fitting and remove the pixels with the 3 per cent most negative fluxes within the continuum windows. The purpose of this step is to exclude from the fitting strong absorption features. We repeat the entire procedure three times to guarantee convergence. We also exclude from our sample objects with final reduced χ^2 larger than 3.

To avoid C IV BALQSOs, we excluded from the sample objects with more than 7 per cent of the pixels with negative C IV residuals. To test the performance of this automatic selection method, we compare the objects that are flagged as BALQSOs using our method with the sample of 562 manually classified BALQSOs from the SDSS-DR2 quasar data base that is described in Ganguly et al. (2007). With our criterion, we flagged as BALQSO a total of 573/5088 ob-

Table A1. Spectral pseudo-continuum windows used for our line fitting procedure under the *local* continuum approach. [1]For each object, we manually adjusted the continuum bands, using the listed wavelength ranges as a reference.

Line complex	——— Continuum windows[1] ———	
Si IV+O IV]	1340–1360Å	1420–1460Å
C IV	1430–1460Å	1680–1720Å
C III]	1680–1720Å	1960–2020Å
Mg II	2650–2670Å	3020–3040Å

jects from the SDSS-DR2 quasar catalogue at $1.7 < z < 2.0$. From these objects, we found that 560/562 objects are also classified as BALQSO in the manually selected Ganguly et al. (2007) sample. These results translate into a successful identification rate of 99.6 per cent and a false-positive identification rate of 2.3 per cent. After excluding the BALQSO candidates and objects with unreliable fits, we end up with 3267 from the DR7Q catalogue (out of the originally selected 4817 objects) and with 35 674 objects from the DR12Q catalogue (out of the originally selected 69 092 objects).

APPENDIX B: SAMPLE COMPARISON

Here we describe in detail the parameter space of the most relevant physical quantities derived for the samples and summarize the most relevant issues associated with each of the small and large samples used in this paper.

B1 Comments on large samples

(i) In both SDSS samples, we can observe that FWHM(C IV) and L_{1450} correlate with each other ($r_p \equiv r_{Pearson} \sim 0.35$). This indicates that, on average, more luminous quasars typically show broader C IV line profiles. However, the FWHM(Mg II) is completely independent of the quasar luminosity ($r_p < 0.05$ in both SDSS samples).

(ii) The DR12Q sample probes considerably fainter sources, compared with the DR7Q sample (by ~ 0.5 dex) and extends to $L_{1450} \gtrsim 10^{44.5} \text{erg s}^{-1}$. This difference between both samples allows us to directly test the impact of luminosity-limited samples as well as data quality in our analysis.

(iii) In the case of $\log \text{FWHM}(\text{C IV})$ and $\log \text{FWHM}(\text{Mg II})$, both SDSS samples span similar ranges, from ~ 3 to ~ 4.2 in $\log \text{FWHM}(\text{C IV})$ and from ~ 3 to ~ 3.8 in $\log \text{FWHM}(\text{Mg II})$. However, the DR12Q sample has a larger fraction of object with low $\log \text{FWHM}(\text{C IV})$ and $\log \text{FWHM}(\text{Mg II})$. Explicitly, 5 per cent (10 per cent) of the objects in the DR12Q sample have $\text{FWHM}(\text{C IV}) \lesssim 3.3$ ($\text{FWHM}(\text{Mg II}) \lesssim 3.3$) versus 2 per cent (6 per cent) in the DR7Q sample. Additionally, the DR12Q sample has a larger fraction of objects with large $\log \text{FWHM}(\text{Mg II})$. The sharp cut at $\log \text{FWHM}(\text{C IV}) = \log \text{FWHM}(\text{Mg II}) = 3$ is imposed by the fitting criterion.

(iv) The $\log \text{FWHM}[\text{C IV/Mg II}]$ distribution in both samples spans over similar ranges, from ~ -0.6 to ~ 0.8. However, the DR12Q sample shows a larger fraction (28 per cent) of objects with $\log \text{FWHM}[\text{C IV/Mg II}] < 0$ than the DR7Q sample (20 per cent).

(v) We can observe that the DR12Q sample shows a larger fraction (roughly 10 per cent) of objects with $\log L_{peak}[\text{C IV/SiO IV}]$ and $\log L_{peak}[\text{C IV/C III}] \gtrsim 1$ than the DR7Q sample (roughly ~ 5 per cent). We need, however, to be cautious about the reliability of such measurements because those objects show $L_{peak}(\text{Si IV} + \text{O IV}])$ and $L_{peak}(\text{C III}])$ weaker than one tenth of

Table A2. Pearson correlation coefficients for the quantities in the first column versus (1) log FWHM (C IV) and (2) log FWHM [C IV/Mg II]. In all cases, it yields $P_P < 1 \times 10^{-13}$.

	All objects		S/N > 10		S/N > 20	
	1	2	1	2	1	2
$\log L_{\text{peak}}$ [C IV/1450 Å]	−0.53	−0.41	−0.60	−0.48	−0.62	−0.47
$\log L_{\text{peak}}$ [C IV/SiO IV]	−0.46	−0.42	−0.52	−0.48	−0.59	−0.49
$\log L_{\text{peak}}$ [C IV/C III]]	−0.33	−0.29	−0.34	−0.33	−0.41	−0.35
Δv_{peak} (C IV)	0.37	0.29	0.55	0.46	0.62	0.45
Δv_{line} (C IV)	0.39	0.37	0.54	0.56	0.57	0.55

Table A3. Pearson correlation coefficients between the listed quantities for the DR7Q and DR12Q samples. We note that in both samples, whenever $|r_P| > 0.1$, it yields $P_P < 1 \times 10^{-10}$.

Property	1	2	3	4	5	6	7	8	9	10	11	12	13
						DR7Q correlations							
1 $\log$ FWHM (C IV)	1	0.17	0.71	0.34	0.51	0.44	−0.63	−0.51	−0.36	−0.11	−0.83	0.22	−0.19
2 $\log$ FWHM (Mg II)	0.17	1	−0.58	0.10	−0.11	−0.19	−0.07	0.03	0.09	0.07	−0.20	0.60	−0.10
3 $\log$ FWHM [C IV/Mg II]	0.71	−0.58	1	0.21	0.50	0.50	−0.47	−0.45	−0.36	−0.15	−0.55	−0.24	−0.09
4 $\log L_{1450}$	0.34	0.10	0.21	1	0.28	0.37	−0.36	−0.36	−0.07	−0.09	−0.44	0.08	−0.16
5 Δv_{peak} (C IV)	0.51	−0.11	0.50	0.28	1	0.68	−0.49	−0.50	−0.39	−0.25	−0.45	0.05	−0.07
6 Δv_{line} (C IV)	0.44	−0.19	0.50	0.37	0.68	1	−0.51	−0.53	−0.34	−0.25	−0.48	−0.03	−0.09
7 $\log L_{\text{peak}}$ [C IV/1450 Å]	−0.63	−0.07	−0.47	−0.36	−0.49	−0.51	1	0.65	0.60	0.77	0.58	−0.20	0.07
8 $\log L_{\text{peak}}$ [C IV/SiO IV]	−0.51	0.03	−0.45	−0.36	−0.50	−0.53	0.65	1	0.47	0.45	0.43	−0.08	0.03
9 $\log L_{\text{peak}}$ [C IV/C III]]	−0.36	0.09	−0.36	−0.07	−0.39	−0.34	0.60	0.47	1	−0.55	0.23	−0.03	−0.15
10 $\log$ EW (C IV)	−0.11	0.07	−0.15	−0.09	−0.25	−0.25	0.77	0.45	−0.55	1	−0.08	−0.06	−0.13
11 $\log \left(L_{\text{peak}} \text{(C IV)}/L \text{(C IV)}\right)$	−0.83	−0.20	−0.55	−0.44	−0.45	−0.48	0.58	0.43	0.23	−0.08	1	−0.23	0.28
12 $\log$ (FWHM (Mg II)/σ (Mg II))	0.22	0.60	−0.24	0.08	0.05	−0.03	−0.20	−0.08	−0.03	−0.06	−0.23	1	−0.19
13 $\log(L_{3000}/L_{1450})$	−0.19	−0.10	−0.09	−0.16	0.07	−0.09	0.07	0.03	−0.15	−0.13	0.28	−0.19	1
						DR12Q correlations							
1 $\log$ FWHM(C IV)	1	0.14	0.71	0.39	0.37	0.39	−0.53	−0.46	−0.33	−0.01	−0.50	0.16	−0.01
2 $\log$ FWHM(Mg II)	0.14	1	−0.60	0.08	0.01	−0.08	−0.01	0.07	0.04	0.08	−0.11	0.58	0.02
3 $\log$ FWHM [C IV/Mg II]	0.71	−0.60	1	0.25	0.29	0.37	−0.41	−0.42	−0.29	−0.07	−0.32	−0.29	−0.03
4 $\log L_{1450}$	0.39	0.08	0.25	1	0.31	0.43	−0.37	−0.53	−0.27	−0.06	−0.29	0.11	0.14
5 Δv_{peak} (C IV)	0.37	0.01	0.29	0.31	1	0.66	−0.39	−0.37	−0.29	−0.13	−0.22	0.01	0.08
6 Δv_{line} (C IV)	0.39	−0.08	0.37	0.43	0.66	1	−0.45	−0.46	−0.29	−0.14	−0.27	−0.02	0.03
7 $\log L_{\text{peak}}$ [C IV/1450 Å]	−0.53	−0.01	−0.41	−0.37	−0.39	−0.45	1	0.59	0.53	0.58	0.28	−0.11	−0.04
8 $\log L_{\text{peak}}$ [C IV/SiO IV]	−0.46	0.07	−0.42	−0.53	−0.37	−0.46	0.59	1	0.47	0.27	0.26	−0.08	−0.09
9 $\log L_{\text{peak}}$ [C IV/C III]]	−0.33	0.04	−0.29	−0.27	−0.29	−0.29	0.53	0.47	1	−0.30	0.16	−0.07	−0.28
10 $\log$ EW(C IV)	−0.01	0.08	−0.07	−0.06	−0.13	−0.14	0.58	0.27	−0.30	1	−0.61	0.00	−0.04
11 $\log \left(L_{\text{peak}} \text{(C IV)}/L \text{(C IV)}\right)$	−0.50	−0.11	−0.32	−0.29	−0.22	−0.27	0.28	0.26	0.16	−0.61	1	−0.10	0.01
12 $\log$ (FWHM (Mg II)/σ (Mg II))	0.16	0.58	−0.29	0.11	0.01	−0.02	−0.11	−0.08	−0.07	0.00	−0.10	1	−0.07
13 $\log(L_{3000}/L_{1450})$	−0.01	0.02	−0.03	0.14	0.08	0.03	−0.04	−0.09	−0.28	−0.04	0.01	−0.07	1

L_{peak} (C IV). Thus, the signal to noise of the C III] and Si IV+O IV] line profiles is probably very low in many of those objects. We can also appreciate that the DR12Q and DR7Q samples show a similar distribution in $\log L_{\text{peak}}$ [C IV/1450 Å].

(vi) As we already explained, DR7Q blueshift estimations are more accurate than in DR12Q. Additionally, we can observe in Fig. 5 that the Δv_{peak} (C IV) and the Δv_{line} (C IV) distributions show a larger fraction of objects with small blueshifts in the DR12Q sample than in the DR7Q sample (6 per cent in DR7Q versus 23 per cent DR12Q and 12 per cent in DR7Q versus 28 per cent in DR12Q for Δv_{peak} (C IV) and Δv_{line} (C IV) < 200 km s^{-1}, respectively). This behaviour is probably caused by objects in the DR12Q sample whose cosmological redshift has been estimated using the C IV profile. This effect artificially biases the C IV blueshifts towards values close to 0. Because of these problems with the DR12Q redshift determinations, we will mainly focus our blueshift analysis on the DR7Q sample.

B2 Comments on small samples

(i) The R13 sample is mostly described by uniform distributions in L_{1450}, FWHM(C IV), and FWHM(Mg II) that are fairly spread around the SDSS data. However, RL AGN are mostly high-luminosity objects and show large FWHM(Mg II) values.

We can also observe that the R13 FWHM [C IV/Mg II] distribution is clearly shifted towards larger values than the peak of the SDSS distributions. In terms of Δv_{peak} (C IV), the sample is shifted towards low values (75 per cent with < 1000 km s^{-1}). Finally, in terms of $\log L_{\text{peak}}$ (C IV) $/L_{\text{peak}}$ (Si IV + O IV)) and $\log L_{\text{peak}}$ (C IV) $/L_{\text{peak}}$ (C III]), the sample is shifted towards large values (75 per cent with $\gtrsim$ 0.5).

(ii) The X-Shooter sample also shows mostly flat distributions in $\log L_{1450}$, $\log$ FWHM (C IV), and $\log$ FWHM (Mg II) that are also fairly spread around the SDSS data. Its $\log$ FWHM [C IV/Mg II] is clearly distributed towards values lower than the peak of the SDSS distribution. Line peak ratios ($\sim$75 per cent with $\log L_{\text{peak}}$ (C IV) $/L_{\text{peak}}$ (Si IV + O IV)) $\lesssim$ 0.5 and $\log L_{\text{peak}}$ (C IV) $/L_{\text{peak}}$ (C III]) $\lesssim$ 0.5) and blueshifts (75 per cent with < 1000 km s^{-1}) are both distributed towards low values with respect to the SDSS distribution peaks.

(iii) The C17 sample is clearly dominated by objects with very large L_{1450} ($L_{1450} \gtrsim 10^{46}$ erg s^{-1}) compared to the other samples. However, its FWHM(C IV) and Δv_{line} (C IV) distributions very closely follow the SDSS DR7Q distributions.

This paper has been typeset from a TEX/LATEX file prepared by the author.