Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **492,** 3580–3601 (2020)
Advance Access publication 2020 January 8

doi:10.1093/mnras/staa018

Exploring the diversity of Type 1 active galactic nuclei identified in SDSS-IV/SPIDERS

Julien Wolf,[1]★ Mara Salvato ,[1] Damien Coffey,[1] Andrea Merloni,[1] Johannes Buchner,[1] Riccardo Arcodia,[1] Dalya Baron,[2] Francisco J. Carrera,[3] Johan Comparat,[1] Donald P. Schneider[4,5] and Kirpal Nandra[1]

[1]*Max-Planck-Institut für extraterrestrische Physik, Gießenbachstraße 1, D-85748 Garching, Germany*
[2]*School of Physics and Astronomy, Tel-Aviv University, Tel Aviv 69978, Israel*
[3]*Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), E-39005 Santander, Spain*
[4]*Department of Astronomy and Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA*
[5]*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*

Accepted 2019 December 16. Received 2019 December 10; in original form 2019 June 12

ABSTRACT

We present a statistical analysis of the optical properties of an X-ray-selected Type 1 active galactic nucleus (AGN) sample, using high signal-to-noise ratio (S/N>20) spectra of the counterparts of the *ROSAT*/2RXS sources in the footprint of the SDSS-IV/SPIDERS (Spectroscopic IDentification of eROSITA Sources) programme. The final sample contains 2100 sources. It significantly extends the redshift and luminosity ranges ($z \sim 0.01$–0.80 and $L_{0.1-2.4\,\mathrm{keV}} \sim 2.0 \times 10^{41}$–$1.0 \times 10^{46}$ erg s^{-1}) used so far in this kind of analysis. By means of a principal component analysis, we derive eigenvector (EV) 1 and 2 in an eleven-dimensional optical and X-ray parameter space, which are consistent with previous results. The validity of the correlations of the Eddington ratio L/L_{Edd} with EV1 and the black hole mass with EV2 is strongly confirmed. These results imply that L/L_{Edd} and black hole mass are related to the diversity of the optical properties of Type 1 AGNs. Investigating the relation of the width and asymmetry of H β and the relative strength of the iron emission $r_{\mathrm{Fe\,II}}$, we show that our analysis supports the presence of a distinct kinematic region: the very broad line region. Furthermore, comparing sources with a red-asymmetric broad H β emission line to sources for which it is blue asymmetric, we find an intriguing difference in the correlation of the Fe II and the continuum emission strengths. We show that this contrasting behaviour is consistent with a flattened, stratified model of the broad-line region, in which the Fe II-emitting region is shielded from the central source.

Key words: accretion, accretion discs – techniques: spectroscopic – galaxies: active – galaxies: nuclei.

1 INTRODUCTION

The exploration of the multiwavelength emission of active galactic nuclei (AGNs), including their spectral features, has motivated the development of unification schemes for this class of persistent and extremely luminous objects (e.g. Antonucci 1993; Urry & Padovani 1995; Netzer 2015; Padovani et al. 2017; see also Elitzur 2012, for some caveats about such unification schemes). Simplifying the complex and historically developed AGN zoo is an appealing idea, since most of the identified species share common spectral properties over large ranges of luminosity and redshift. In this context, two main subclasses have been defined based on their optical emission-line properties: Type 1, which have both broad and narrow optical/UV emission lines, and Type 2, which show only narrow lines. These two classes are unified via the hypothesis that the broad lines in Type 2 sources are obscured (e.g. Netzer 2008; Hickox & Alexander 2018). There is, none the less, a substantial variety in the properties within the Type 1 population. The broad lines encode information about the geometry and the kinematics of the inner regions of the system: the broad-line region (BLR; e.g. Rees, Netzer & Ferland 1989; Peterson 2006; Gaskell 2009). Photoionized by the continuum emission of the central source, the BLR emits transition lines, which are expected to be broadened by the motion of the gas (e.g.

★ E-mail: jwolf@mpe.mpg.de

Baldwin et al. 1995; King 2016). The virial broadening of the emission lines can, under certain assumptions, be used to measure the mass $M_{\rm BH}$ of central supermassive black holes (SMBHs). This mass estimation method is calibrated using estimates of the BLR radius from reverberation mapping (Bahcall, Kozlovsky & Salpeter 1972; Blandford & McKee 1982; Peterson 1993; Gebhardt et al. 2000; Kaspi et al. 2005; Bentz & Katz 2015; Shen et al. 2015). The response delay of the BLR emission to variations in the ionizing continuum emission from the central source (i.e. the accretion disc) is used to determine the distance between the BLR and the nucleus, assuming a Keplerian velocity field. Broad, low-ionization lines, such as H β (rest-frame wavelength at 4862 Å) and the Mg II λ2800 Å doublet,[1] are generally thought to yield more reliable virial broadening estimators than, for instance, the high-ionization transition C IV λ1549 Å (e.g. Trakhtenbrot & Netzer 2012; Mejia-Restrepo et al. 2016; Mejía-Restrepo et al. 2018). This is in part due to the presence of prominent blueshifted components in the C IV λ1549 Å line profiles of strong accretors, which may not be related to the orbital motion of the gas and would affect the inferred velocities (e.g. Richards et al. 2002; Zamanov et al. 2002; Marziani et al. 2017; Sulentic et al. 2017). However, the H β lines possess profiles that are difficult to reconcile with simple Keplerian motion around the central black hole, and could significantly affect the $M_{\rm BH}$ estimates (Negrete et al. 2018). Accurate $M_{\rm BH}$ measurements are important for the development of galaxy evolution models. It is therefore critical to understand the mechanisms responsible for the observed broad line shapes.

The debate on the actual BLR kinematics, physical processes, and geometry has flourished through the past decades; the variety of the often asymmetric (sometimes double-peaked) broad lines and their variability has led to an abundant diversity of competing models. It was not until recently that the kinematics of this region could be resolved for the first time, favouring a flattened BLR as an orbiting extension of the accretion disc (Gravity Collaboration 2018). A flattened BLR geometry was initially supported by the relation between the width of H β and the radio core-to-lobe ratio. (Wills & Browne 1986; Jarvis & McLure 2006; Brotherton, Singh & Runnoe 2015) and by the double shoulders and peaks observed in certain broad emission lines (Oke 1987; Chen & Halpern 1989; Chen, Halpern & Filippenko 1989; Eracleous & Halpern 1994, 2003; Storchi-Bergmann et al. 2016).The double-peaked profile variability has been shown not to correlate with continuum changes. The peak shifts are therefore not a reverberation effect (Wanders & Peterson 1996). The strong variability of the double-peaked emission profiles has also been discussed in the light of different disc models by Lewis, Eracleous & Storchi-Bergmann (2010), which can account for some of the observed line profiles.

A potentially useful approach to constrain and explore the inner structure of an AGN is to acquire an understanding of their spectral diversity. For instance, in their seminal paper, Boroson & Green (1992) performed a principal component analysis (PCA) of a set of optical, radio, and X-ray features of 87 Type 1 AGNs. Through the orientation of the first principal component (eigenvector 1, EV1) in their parameter space, they established an anticorrelation of the equivalent width of the narrow [O III] lines at 4959 and 5007 Å with the relative strength of the iron emission $r_{\rm Fe\,II} = F({\rm Fe\,II}\,\lambda4570)/F({\rm H}\,\beta)$, and the full width at half-maximum (FWHM) of the H β line as central markers of diversity in their sample. A large number of subsequent studies have investigated the EV1 correlation planes (e.g. Sulentic, Marziani & Dultzin-Hacyan 2000a; Sulentic et al. 2000b; Marziani et al. 2001; Shang et al. 2003; Grupe 2004; Kuraszkiewicz et al. 2009; Mao, Wang & Wei 2009; Tang et al. 2012; Shen & Ho 2014). In particular, Sulentic et al. (2000a, 2000b) and Marziani et al. (2001) established the foundations of the four-dimensional EV1 (4DE1) formalism, including $\rm FWHM_{H\,\beta}$ and $r_{\rm Fe\,II}$ as two of the main correlates of EV1 in Boroson & Green (1992). These two quantities are, respectively, related to $M_{\rm BH}$ (since $\rm FWHM_{H\,\beta}$ is used as virial broadening estimator) and the Eddington ratio $L/L_{\rm Edd}$ (e.g. Grupe et al. 1999; Marziani et al. 2001; Netzer & Trakhtenbrot 2007; Sun & Shen 2015; Du et al. 2016; Panda, Marziani & Czerny 2019c). The 4DE1 parameter space was extended with the soft X-ray photon index $\Gamma_{\rm soft}$ and the centroid shifts from rest-frame wavelength of C IV λ1549 at 50 per cent fractional intensity (Sulentic et al. 2007b). The careful exploration of this parameter space led to a proposed two-population paradigm in the low-redshift universe, with an empirical separation at $\rm FWHM_{H\,\beta} \approx 4000\,km\,s^{-1}$ (Sulentic et al. 2000a, b; see section 5 of Marziani et al. 2018, and references therein for a full review of the evidence supporting the population A/B classification). The optical plane of 4DE1 (i.e. $\rm FWHM_{H\,\beta}$ versus $r_{\rm Fe\,II}$) was established as an analogue to the stellar Hertzsprung–Russel diagram for the Type 1 AGN population. The domain occupied by Type 1 AGNs in this plane has been presented as a quasar main sequence (Marziani et al. 2001; Sulentic, Marziani & Zamfir 2011). The viewing angle and the accretion power have been investigated as potential drivers of this main sequence (Shen & Ho 2014; Panda et al. 2018). Zamfir et al. (2010, henceforth Z10) studied the characteristics of the H β emission of 477 optically selected AGNs ($z < 0.7$) in the optical 4DE1 plane. They reported that the line shapes distinctively changed along the sequence, the asymmetry and shifts scaling with both of the optical dimensions, corroborating early results by Sulentic et al. (2002).

In this work, we explore the EV1 correlation space exploiting high signal-to-noise ratio (S/N) optical spectra for the largest sample of X-ray-selected Type 1 AGNs used to date for this type of analysis. The X-ray selection allows us to construct a less biased sample of AGNs, since the host X-ray-related processes are believed to be particularly weak compared to high-energetic AGN emission. At the same time, X-rays are able to penetrate the host galaxies and in general large column densities, thus making the sample more complete than with other AGN selection techniques (e.g. Singh 2013; Padovani et al. 2017). The Type 1 AGNs in our sample were detected by *ROSAT* (Voges et al. 1999; Boller et al. 2016), while their multiwavelength counterparts were determined in Salvato et al. (2018) and spectroscopically followed up in the SDSS-IV/SPIDERS programme (Dwelly et al. 2017; Comparat et al. 2019). These sources were presented in Coffey et al. (2019, henceforth C19), where they were selected according to the width of their broad lines. The final sample contains 2100 sources, which is a factor of 4–18 larger than samples used in previous studies of this type. It significantly extends the probed redshift and luminosity ranges ($z \sim 0.01$–0.80 and $L_{0.1\text{–}2.4\,\rm keV} \sim 2.0 \times 10^{41}$–$1.0 \times 10^{46}\,\rm erg\,s^{-1}$). We perform a PCA on the optical and X-ray features of these sources, which allows us to determine the markers of spectral diversity in our sample. We specifically investigate the role played by $L/L_{\rm Edd}$ and $M_{\rm BH}$ for the total variance in our data. The asymmetry of the broad H β contributes strongly to the orientation of EV1 through the chosen spectral parameter space. We thus follow in the footsteps of Z10 and connect our

[1] In an effort to simplify the notation, we will henceforth refer to this line as Mg II.

global statistical results to the shape of the broad H β emission line, which is discussed in the context of BLR stratification and low-ionization outflows.

The paper is organized as follows: In Section 2, we briefly present the SPIDERS AGN value-added catalogue of C19 and how we created our subsample. We also outline how the additional properties, which were not provided in C19, were determined. In Section 3, we describe the distribution of the emission-line asymmetry index. The core results of the direct correlation analysis and PCA on our sample are presented in Section 4. We demonstrate how the derived physical parameters $L/L_{\rm Edd}$ and $M_{\rm BH}$ scale with the obtained principal components in Section 5. The distribution of our sample in the EV1 plane and the possible presence of a distinct kinematic BLR region are investigated in Section 6. We divided the sample in two distinct subsets according to the sign of their H β asymmetry indices. From the comparison of parameter correlations in the two subsamples, we detect an interesting contrast in the scaling relation between source luminosity and the equivalent width of the iron emission. This is discussed in Section 7. Conclusions are presented in Section 8. Throughout this work, we adopt a standard cosmology: $\Omega_{\rm M} = 0.3$, $\Omega_{\Lambda} = 0.7$, and $H_0 = 70\,{\rm km\,s^{-1}\,Mpc^{-1}}$.

2 DATA

The sample of Type 1 AGNs analysed in this work is extracted from the SDSS-IV/SPIDERS AGN catalogue, presented in C19. The original catalogue compiles spectral information through SDSS DR14[2] (Abolfathi et al. 2018), for 7344 Type 1 AGNs detected in the Second *ROSAT* All-Sky Survey (*ROSAT*/2RXS; Boller et al. 2016) and 1157 Type 1 AGNs in the *XMM–Newton* Slew 1 catalogue (XMMSL1; Saxton et al. 2008), for which the multiwavelength counterparts were determined in Salvato et al. (2018). Optical spectra of the counterparts were obtained from two different spectrographs, SDSS and BOSS. These instruments cover different optical wavelength ranges: 3800–9200 Å for SDSS (counting both channels, with a spectral resolution ranging from 1850 to 2200) and 3600–10 400 Å for BOSS (spectral resolution of 1560–2270 in the blue channel, and 1850–2650 in the red channel). A technical summary of the SDSS-IV survey is provided by Blanton et al. (2017). The instrument is presented in Gunn et al. (2006). Finally, a detailed description of the SDSS and BOSS spectrographs is given by Smee et al. (2013).

For the sources in the C19 catalogue, the authors fit the spectroscopy following a procedure detailed in C19. Here we report only the key features. In C19, the spectral regions centred around the H β and Mg II lines (4420–5500 and 2450–3050 Å, respectively) were fitted with a multicomponent continuum model and a series of Gaussian functions. This work will specifically focus on the broad H β emission line and the narrow [O III] λ4959 Å and [O III] λ5007 Å forbidden transition lines. The C19's fitting algorithm used up to four Gaussian functions to fit the H β line. The best-fitting model (and thus the required number of Gaussian components) was automatically selected on the basis of the Bayesian Information Criterion (BIC). Of the four Gaussians, one accounts for the narrow core, while the remaining potential three components of H β are defined as broad if they fulfil the criterion: ${\rm FWHM_{H\,\beta}} > 800\,{\rm km\,s^{-1}}$. This relatively complex model allows one to trace not only typical broad bimodal profiles above ${\rm FWHM_{H\,\beta}} \sim 1000\,{\rm km\,s^{-1}}$, but also distinct broader components (see section 8.1 in C19, and Section 6 of this work). For instance, Marziani et al. (2010) propose a broad-line decomposition into broad, very broad, and blue components (called 'BC', 'BLUE', and 'VBC', respectively; for a review, see section 6.2 of Marziani et al. 2018), which the model defined in C19 may individually trace. Up to two Gaussians were used for each of the [O III] lines: one fitting the narrow core and one tracing shifted wings. The continuum model was constructed from a power law, a host galaxy component, and an Fe II template. The iron emission and the galaxy contribution are obtained, respectively, from a normalized narrow-line Seyfert 1 I Zw 1 template (Boroson & Green 1992) and an early-type galaxy template.[3] Morphological studies (e.g. Grogin et al. 2005) have shown that AGNs are mainly found in bulge-dominated galaxies. Bulges are predominantly home to old star populations. C19 thus justify the use of an early-type galaxy template for the spectral host contribution by the fact that SDSS fibres sample only the central regions of targeted galaxies. The early-type approximation for the host contribution was also used in previous work (e.g. Calderone et al. 2017). The template has H β absorption features. Comparing a subsample with strong host contribution to a subsample with low host contribution at z<0.2, we found no significant impact of these features on H β line shape diagnostics. The fitting parameters are listed in the catalogue of C19 along with monochromatic continuum and X-ray luminosities and derived parameters such as bolometric luminosities, and estimates for $M_{\rm BH}$ and $L/L_{\rm Edd}$.

In C19, the single-epoch $M_{\rm BH}$ estimation method was used for the derivation of the black hole masses (McLure & Jarvis 2002; Vestergaard 2002; Vestergaard & Peterson 2006; Assef et al. 2011; Shen & Liu 2012 and Shen 2013, for a review). The masses stem from the Assef et al. (2011) calibration, which builds on the FWHM of H β and the BLR radius–luminosity relation from Bentz et al. (2009). For this paper, we have reconstructed the best-fitting models from the parameters listed in C19. Two examples of H β fit reconstructions are shown in Appendix A of the present work. The details of the broad H β line decomposition are also presented.

2.1 Sample construction and selection pipeline

We are primarily interested in Type 1 AGN spectral properties, which can be associated with the physics of the BLR and the narrow-line region (NLR). We thus assembled a parameter subset inspired by the classical quasar PCA papers by Boroson & Green (1992) and Grupe (2004). They included emission properties of the H β, Fe II, and [O III] lines at 5007[4] and He II λ1640 Å, as well as X-ray and optical monochromatic luminosities to trace the emission of the hot electron corona and the accretion disc. We confined the analysis to sources detected in 2RXS only (7344 sources), as they were six times more numerous than the XMMSL1 sources in the SPIDERS AGN sample. This also simplifies the analysis, given the fact that the X-ray fluxes from the two surveys are measured in different bands and at different times (see Saxton et al. 2008; Boller et al. 2016). More specifically, we retained the 2RXS sources, for which the X-ray flux in the 0.1–2.4 keV band is provided in C19 (7154 sources). The missing 190 sources have no X-ray flux mea-

[2]SDSS DR14 VACs.

[3]SDSS spectral cross-correlation templates, Template 24.

[4]We choose to work with the [O III] transition line at 5007 Å and not 4959 Å, since this line will be less affected by interline fit contamination with H β. We argue that any blue asymmetries and shifts in [O III] λ5007 Å should be detectable in [O III] λ4959 Å. Throughout this paper, [O III] λ5007 Å will be referred to as [O III].

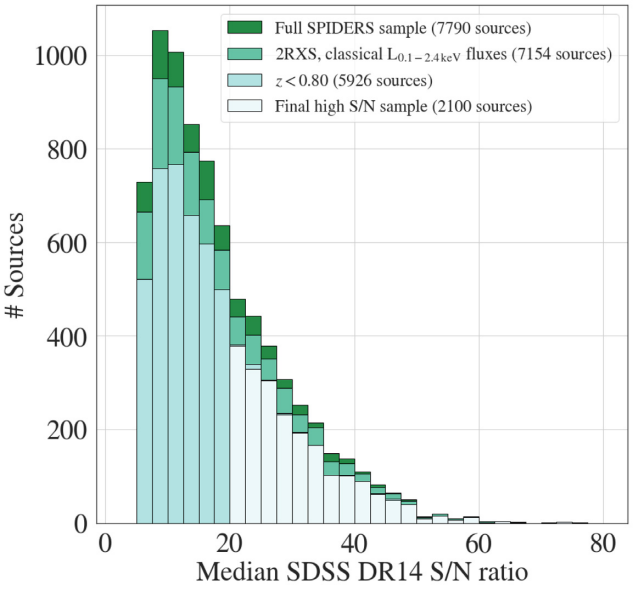

Figure 1. Distribution of the median S/N ratios per resolution element of the 2100 SDSS DR14 sources in our final sample (white). Also presented are the subsamples of the full SPIDERS sample in C19 after two consecutive quality cuts: the presence of a 0.1–2.4 keV flux and the redshift restriction $z < 0.8$.

Table 1. Summary of the sample construction pipeline.

Sample cuts (Initial SPIDERS AGN sample: 7790 sources)	Remaining sources
2RXS flux	7154
$z < 0.8$	5926
S/N $\geq$ 20	2124
Robust asymmetry index/black hole masses	2100

surement listed. They correspond to sources with low photon count rates.

Since we use the monochromatic luminosity at 5100 Å to trace the continuum emission, we limit the sample to $z = 0.8$ so that the rest-frame 5100 Å are covered by the spectral window. The redshift cut ensures the presence of fits of the H β centred region. These constraints yield 5926 sources. We further limit our analysis to sources with good SDSS S/N per resolution element in order to improve the significance of our results. Our analysis relies on the accurate measurement of line shapes, and low-quality spectra might strongly affect the fits. For this reason, we consider only sources with median S/N $\geq$ 20. This critical restriction drops our total source count to 2124 objects. The SDSS S/N distribution of the original sample and its consecutive cuts are shown in Fig. 1. Finally, 24 sources were excluded due to a lack of $M_{\rm BH}$ estimates, extreme uncertainties on the fit parameters of H β or absent [O III]. A visual inspection of these sources' spectra revealed that these issues might be linked to strong continuum emission. Table 1 summarises the consecutive cuts of the sample construction.

For the 2100 sources, the following subset of spectral properties was compiled (unless specified otherwise, all wavelengths, energies, equivalent widths are defined in the rest frame):

(i) $\mathbf{FWHM_{H\,\beta}}$: full width at half-maximum of H β (broad component) [km s^{-1}];
(ii) $\boldsymbol{F}$**([O III])**$/\boldsymbol{F}$**(H** $\boldsymbol{\beta}$**)**: flux ratio of [O III] λ5007 Å and H β;
(iii) $\boldsymbol{r_{Fe\,II}} = \boldsymbol{F}$**(Fe II)**$/\boldsymbol{F}$**(H** $\boldsymbol{\beta}$**)**: flux ratio of H β and the total Fe II (Fe II emission in the 4434–4684 Å range);
(iv) $\boldsymbol{L_{5100\,Å}}$: optical continuum monochromatic luminosity at 5100 Å from the power-law component (without host and Fe II contributions) [erg s^{-1} Å^{-1}];
(v) $\boldsymbol{L_X}$: soft observed X-ray luminosity in the 0.1–2.4 keV band [erg s^{-1}];
(vi) $\boldsymbol{W}$**(Fe II)**: equivalent width of Fe II (4570 Å blend)[5] [Å];
(vii) $\boldsymbol{W}$**([O III])**: equivalent width of [O III] λ5007 Å (full profile)[5] [Å];
(viii) $\boldsymbol{W}$**(H** $\boldsymbol{\beta}$**)**: equivalent width of the broad component of H β[5] [Å];
(ix) $\boldsymbol{W}$**(He II)**: equivalent width of He II at 4686 Å[5] [Å];
(x) $\mathbf{\Delta\lambda_{[O\,III]}}$: asymmetry index for [O III] λ5007 Å;[6]
(xi) $\mathbf{\Delta\lambda_{H\,\beta}}$: asymmetry index for the broad component of H β.[6]

It is worth to highlight the following points:

(i) Sources for which the parameters relative to the Fe II emission were not listed in C19 typically correspond to AGNs with very weak Fe II emission. We manually set their $r_{\rm Fe\,II}$ and iron equivalent widths to zero. We thus implicitly assume that their iron emission is too weak to be disentangled from the AGN continuum emission.

(ii) All the spectral properties of our defined subset are available in C19 except for the equivalent widths and asymmetry indices, which were derived as presented in Section 2.2 and Appendix A.

(iii) The $L_{5100\,Å}$ provided in C19 includes the host contribution. However, in the present analysis we aim at tracing the accretion disc emission with $L_{5100\,Å}$. We thus derived the monochromatic luminosity from the reconstructed power-law model at rest frame, removing the Fe II and host contribution. The uncertainties are obtained from the errors on the normalization of the slope. The considerable degeneracy between the host, iron, and power-law components at lower redshift ($z < 0.2$) due to the stronger host contribution makes it challenging to correctly perform an AGN-host decomposition. It is the principal limitation of deriving monochromatic luminosities directly from the fitted power-law model. The relative contributions of the Fe II complex and the host galaxy to the monochromatic flux at 5100 Å are illustrated in Fig. 2. While the Fe II contribution at 5100 Å to the total monochromatic flux is marginal over the complete redshift range, the host galaxy strongly contributes to the monochromatic flux at lower redshifts.

(iv) The X-ray fluxes were obtained by C19, following Dwelly et al. (2017), from an absorbed power law ($\Gamma = 2.4$). In addition to these classical fluxes, C19 provide Eddington bias-corrected 2RXS fluxes using a Bayesian prior (Kraft, Burrows & Nousek 1991; Laird et al. 2009; Georgakakis & Nandra 2011). The Bayesian flux estimates reach unrealistic low values for sources with low count rates. C19 allow up to a factor of 10 difference with the classical fluxes. Sources for which the Bayesian fluxes do not meet this requirement only have classical flux measurements. We chose to use the uncorrected X-ray flux measurements in this work, since using Bayesian fluxes would reduce the size of our final sample by a factor of 2. We note that a test run of the statistical pipeline presented in this work on a smaller sample with Bayesian soft X-ray luminosities yielded similar results to those presented here for the entire sample using the nominal fluxes.

Fig. 3 presents the soft X-ray luminosity–redshift distribution of

[5]cf. Appendix A.
[6]cf. Section 2.2.

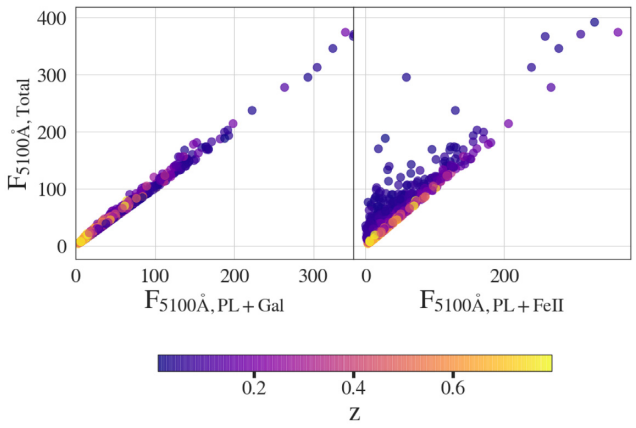

Figure 2. The relation to the monochromatic fluxes at 5100 Å, as measured from the power-law model with Fe II and host contributions ($F_{5100\,Å,Total}$), the power-law model with the Fe II emission ($F_{5100\,Å,PL+Fe\,II}$), and the power-law model with the host emission ($F_{5100\,Å,PL+Gal}$). The fluxes are in units of 10^{-17} erg cm^{-2} s^{-1} Å^{-1}.

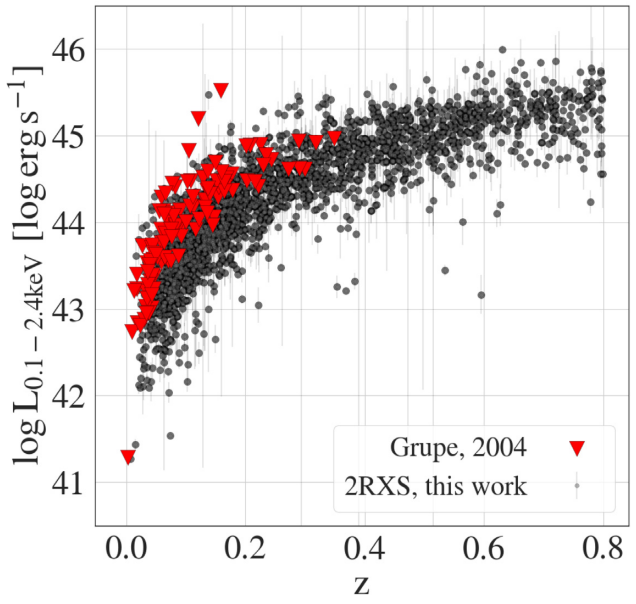

Figure 3. The observed soft X-ray luminosity–redshift distribution for our 2100 2RXS sources. The red triangles are the 110 AGNs used in Grupe (2004), which all lie at the brighter end of our sample for equivalent redshifts. Our sample significantly extends both the redshift and the luminosity ranges.

our sources, and compares our sample to that of Grupe (2004). Our sample spans a range of soft X-ray luminosities from 1.9×10^{41} to 9.9×10^{45} erg s^{-1}, with redshifts up to 0.8. Compared to Grupe (2004), our sample extends to lower luminosities at a given redshift and also probes a higher redshift regime. The larger sample will in particular allow the placement of more stringent constraints on the relation of EV1 and EV2 to physical parameters (see Fig. 10).

2.2 Measuring asymmetry in emission lines: motivation and method

Emission lines in AGNs show asymmetries, which encode precious information on the geometry and kinematics of the emitting region. In the case of broad emission lines, asymmetry is defined as the relative displacement of the centroid wavelength of the peak and base components of the profile. The peak wavelength of the profile is in practice measured at a high fraction of the broad component intensity (without the narrow core) and is related to the classical broad component (Brotherton et al. 1994; Popovic et al. 2002; Adhikari et al. 2016). The base of the emission-line profile is, as its name indicates, measured at lower fractional intensities and is expected to trace a distinct, red- or blueshifted, sometimes broader component. The shifted VBC is believed to originate from the presence of a distinct emitting region in the BLR: the very broad line region (VBLR). The VBLR is expected to be composed of highly ionized gas situated even closer to the central black hole. It is further discussed in Section 6 of this work. The concept of BLR stratification arose from reverberation studies. Peterson & Ferland (1986) reported variability in the profiles of He II 4686 Å and H β in NGC 5548, most notably the appearance of broader line components while the continuum luminosity increased. Measuring the asymmetry of broad emission lines and including the indices in a statistical analysis therefore offer an approach to investigate how the relative kinematics of assumed layers in the BLR impacts the diversity of observed Type 1 AGN. One can also use the asymmetry index to trace the relative displacement of narrow emission lines and their shifted wings. This is particularly useful to trace the presence of outflows in the NLR (e.g. Heckman et al. 1981; Zamanov et al. 2002; Zakamska et al. 2016; Rakshit & Woo 2018; Wang, Xu & Wei 2018). In this case, the displacement of peak and base components is measured for the full profile (i.e. narrow core and shifted wings).

The asymmetry index was not included in C19 and was determined here starting from the reconstruction of the lines, using the Gaussian models listed in the catalogue. In this work, we derived the following asymmetry indices:

(i) $\Delta\lambda_{H\beta}$ for the broad component of H β. From the C19 catalogue, we retain the Gaussians with FWHM $> 800\,\mathrm{km\,s^{-1}}$ as broad components. From the four Gaussian components used to fit the full H β profile, one was fixed by C19 at FWHM $< 800\,\mathrm{km\,s^{-1}}$ to force the fit of the narrow emission line. Up to three Gaussians were thus used to fit the broad profile.

(ii) $\Delta\lambda_{[O\,III]}$ for the full profile of [O III] λ5007 Å. Here, both narrow and broad components are used for the measure. In this work, we use the description of the asymmetry $\Delta\lambda$ measurement for any emission line first introduced by Heckman et al. (1981), and further detailed in Winkler & Chauke (2014). The method requires the identification of the wavelengths at which the line reaches 15 and 80 per cent of its maximum intensity, which required the reconstruction of the multi-Gaussian line fits as described in Appendix A. We determined the maximum of the full line profile (or alternatively the broad components for H β) and measure 15 and 80 per cent of this maximum. A schematic view of this measurement is presented in Fig. 4. The asymmetry index ($\Delta\lambda$) is defined by

$$\Delta\lambda = \frac{\alpha_b - \alpha_r}{\alpha_b + \alpha_r} = \frac{(\lambda_{c,high} - \lambda_{b,low}) - (\lambda_{r,low} - \lambda_{c,high})}{(\lambda_{c,high} - \lambda_{b,low}) + (\lambda_{r,low} - \lambda_{c,high})} \quad (1)$$

$$= \frac{\lambda_{b,high} + \lambda_{r,high} - \lambda_{b,low} - \lambda_{r,low}}{\lambda_{r,low} - \lambda_{b,low}}. \quad (2)$$

Here, $\lambda_{b,\,high}$ and $\lambda_{r,\,high}$ are measured at 80 per cent fractional intensity and $\lambda_{b,\,low}$ and $\lambda_{r,\,low}$ at 15 per cent fractional intensity. $\lambda_{c,\,high}$ is the wavelength of the centre of the line at 80 per cent fractional intensity, i.e. the mid-point between $\lambda_{b,\,high}$ and $\lambda_{r,\,high}$. The parameters α_b and α_r are the distances from $\lambda_{c,\,high}$ to the

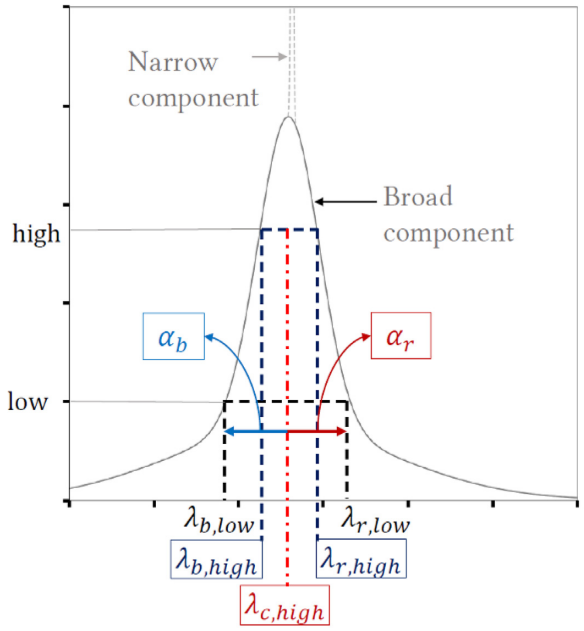

Figure 4. Diagram for the determination of the asymmetry parameter using the Heckman et al. (1981) definition, adapted from Winkler & Chauke (2014). For H β, the broad component is used without the narrow line. For [O III], the asymmetry of the full profile (narrow and wing) is measured.

blue and red edges of the line at 15 per cent fractional intensity. A negative (positive) $\Delta\lambda$ corresponds to a blueward (redward) asymmetry, and its value ranges from -1 to 1. This measure is not sensitive to the continuum emission (cf. correlation matrix shown in Section 4.1). We propagated the uncertainties of the FWHM of the fits into the measure of asymmetry and obtain typical errors of $\sigma_{\Delta\lambda} \sim 0.05$. An important caveat is that using the uncertainties of the FWHM from the full profile does not account for the lower, distinct and possibly more contaminated line components. Examples of fit reconstructions of red- and blue-asymmetric H β from our sample are displayed in Appendix B, Fig. B1, alongside their Gaussian decomposition.

3 IMPACT OF FIT CONTAMINATION ON THE Hβ AND [O III] ASYMMETRY INDEX DISTRIBUTION

The distributions for H β and [O III] asymmetry indices are shown in Fig. 5. The indices clearly follow very different distributions for the two emission lines. The following statistics were derived for the distributions of asymmetries with base centroids measured at 25 per cent fractional intensity (i.e. the blue histograms in Fig. 5). The asymmetry indices of the [O III] are predominantly blueward (skewness of distribution: $\gamma = -0.54$), which is consistent with the presence of a blueshifted wing detected in a two-Gaussian model. The mean value of $\Delta\lambda_{[O\,III]}$ is ~ -0.10. 1546 sources with [O III] asymmetry below zero and 547 above zero are counted, confirming the excess of the blue-asymmetric [O III] lines. The distribution of the H β asymmetry appears positively asymmetric (skewness of distribution: $\gamma = 0.36$). The mean value of $\Delta\lambda_{H\beta}$ is $+0.03$. An excess of red-asymmetric H β emitters is clearly measured (1244 sources out of 2100 with $\Delta\lambda_{H\beta} > 0$). The fractional intensity at which the lower velocity shift is measured for the asymmetry index has important repercussions on the overall distribution of asymmetries as Fig. 5 demonstrates. The distributions of $\Delta\lambda_{H\beta}$ and $\Delta\lambda_{[O\,III]}$ are displayed for different fractional intensities at which $\lambda_{b,\,low}$ and $\lambda_{r,\,low}$ are set: 5, 10, 15, 20, and 25 per cent. While we always detect an excess in redward asymmetry for the H β profiles, the lower the percentage at which $\lambda_{b,\,low}$ and $\lambda_{r,\,low}$ are measured, the more the asymmetry index distribution is bimodal, with a clear separation between the symmetric and red-asymmetric H β modes. It is unclear if this redward excess in H β asymmetries is influenced by [O III] and continuum contamination in the fit. The bimodality of the $\Delta\lambda_{H\beta}$ distribution measured at lower base intensities, however, clearly indicates a preferred displacement between peak and base components, which could arise from the near-systematic contamination by [O III] or its blue wing.

A similar effect is observed for the distribution of the [O III] asymmetries: it becomes double peaked, with a distinct second mode appearing blueward of the symmetric sources for lower base intensities. For the measurement of the asymmetry index, we set the base of the lines at 15 per cent fractional intensity, in order to avoid the detection of overlaying line models. This choice is motivated by two aspects of the distributions seen in the two panels of Fig. 5: 15 per cent fractional intensity, the peak of the symmetric sources is more sensitive to low intensity, and shifted profile components are detected by the measurement than at 25 and 20 per cent. Furthermore, for this measurement configuration (peak at 80 per cent and base at 15 per cent fractional intensity), the bimodality, which we identify as a signature of model degeneracy, is not yet detected.

To further investigate the risk of model degeneracy, the dependence of the [O III] asymmetry index on the shifts of the H β base component c_{15} (the centroid shift at 15 per cent fractional intensity with respect to the rest-frame wavelength) is presented in Fig. 6. We clearly observe an excess of sources in the quadrant of the blue-asymmetric [O III] lines and redshifted H β base components. The coloured markers correspond to high signal-to-noise ratio sources, which also appear to preferentially populate the lower right quadrant in the figure. 1196 sources out of 2100 (57 per cent) have $\Delta\lambda_{[O\,III]} < 0$ and $c_{15} > 0$. In the high S/N regime, i.e. S/N > 35, 278 out of 498 sources (56 per cent) have $\Delta\lambda_{[O\,III]} < 0$ and $c_{15} > 0$. A binomial test showed that the clustering in the lower right quadrant is indeed significant. For both c_{15} and $\Delta\lambda_{[O\,III]}$, the null hypothesis that the sources distribute equally on both sides of zero was rejected with p-values $p < 10^{-6}$.[7] This result underlines the possibility of model degeneracy in the H β region. For example, in their study of extremely reddened quasars, Perrotta et al. (2019) recently presented rare sources for which H β and [O III] appear to blend (see their fig. 1). Disentangling the presence of a shifted VBC from fit contamination would offer a cleaner window on the kinematics of the BLR; however, the fitting procedure in C19 did not allow for such a decomposition.

4 STATISTICAL ANALYSIS

This section presents the various statistical tests performed on the selected optical properties of our sample of 2100 SPIDERS AGNs, listed in Section 2.

[7]Throughout this work, ρ denotes correlation coefficients, while p stands for p-value.

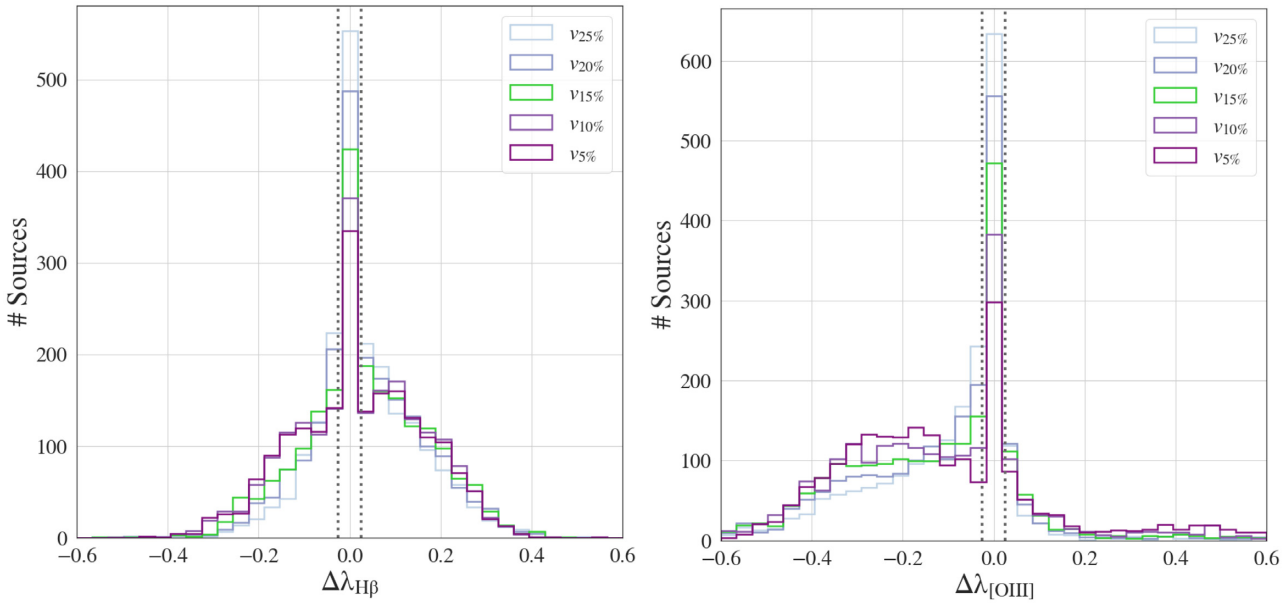

Figure 5. *Left*: The H β asymmetry distribution in the final source subsample for different levels of base fractional intensity: 5, 10, 15, 20, and 25 per cent. A clear inflection redward of the symmetric H β profiles arises for bases measured at lower intensities. Typical 1σ errors are shown centred at zero by two dotted vertical bars. The $\Delta\lambda_{\mathrm{H}\,\beta}$ values derived for a fractional base intensity of 15 per cent are used in this paper (green histogram). *Right*: The [O III] asymmetry distribution for different levels of base fractional intensity: 5, 10, 15, 20, and 25 per cent. A clear inflection can be detected here, bluewards of the symmetric [O III] profiles for bases measured at lower intensities. Typical 1σ errors are shown centred at zero by two dotted vertical bars.The green histogram was obtained from $\Delta\lambda_{[\mathrm{O\,III}]}$ measurements with base fractional intensity at 15 per cent as used in this work.

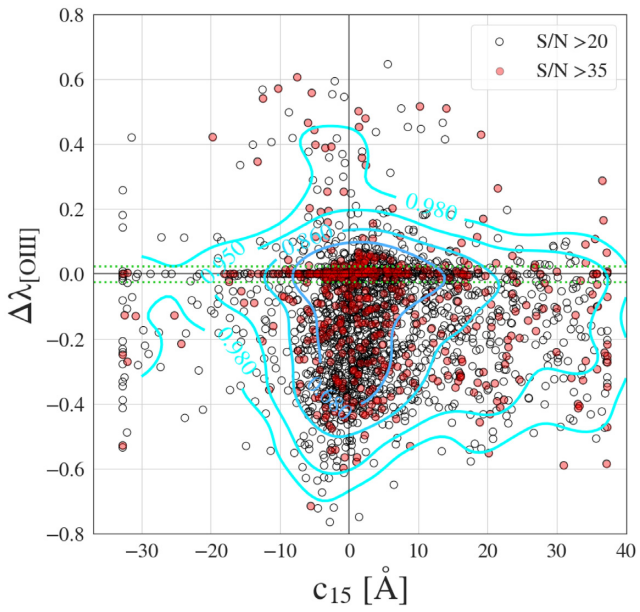

Figure 6. The asymmetry index of [O III] as a function of the H β centroid shift at 15 per cent fractional intensity. The excess of sources in the lower right quadrant suggests an interline contamination for redshifted H β components and blueshifted [O III] components. The filled circles mark the sources with S/N above 35. The green dotted horizontal lines indicate typical 1σ uncertainties in the asymmetry parameter. The density contours containing 68, 86, 95, and 98 per cent of all the sources in our sample were obtained from a Gaussian kernel density estimate.

4.1 Direct correlation

Our initial step is to generate a Spearman rank correlation matrix for our sample (Fig. 7). Statistically insignificant correlations, defined by imposing a threshold of $p_{\mathrm{max}} = 0.05/55$ on the p-values, were masked. The threshold p_{max} applies the Bonferroni correction for multiple statistical hypothesis testing (e.g. Haynes 2013). Positive (negative) coefficients indicate a positive (negative) monotonic relation between two parameters. We have included a dendrogram (tree diagram) that clusters our data hierarchically using correlation as distance metric (for a review, see Baron 2019). The clustered correlation matrix reveals significant ($p < 10^{-6}$) structures in our parameter subset, some of which are also previously reported in Boroson & Green (1992), Grupe (2004), and Shen & Ho (2014). In particular,

(i) A clear anticorrelation ($\rho = -0.53$) arises between $\mathrm{FWHM}_{\mathrm{H}\,\beta}$ and $r_{\mathrm{Fe\,II}}$. This behaviour is the well-known anticorrelation between the first two dimensions of the 4DE1 (e.g. Sulentic et al. 2000a, 2000b; Marziani et al. 2003a, b; Sulentic, Dultzin-Hacyan & Marziani 2007a; for a review, see Marziani et al. 2018 and Sulentic & Marziani 2015). In addition, $W(\mathrm{Fe\,II})$ and $\mathrm{FWHM}_{\mathrm{H}\,\beta}$ are anticorrelated ($\rho = -0.38$), as expected.

(ii) Similarly, an even stronger anticorrelation ($\rho = -0.61$) is measured between $W([\mathrm{O\,III}])$ and $r_{\mathrm{Fe\,II}}$, in agreement with the main EV1 anticorrelation.

(iii) The X-ray and continuum luminosities, L_{X} and $L_{5100\,\text{Å}}$, are related to the emission-line properties in a similar manner. There is a strong correlation ($\rho = 0.54$ for $L_{5100\,\text{Å}}$ and $\rho = 0.55$ for L_{X}) with $W(\mathrm{H}\,\beta)$. The correlation between $L_{5100\,\text{Å}}$ and $W(\mathrm{H}\,\beta)$ is consistent with photoionization models in which increasing emission of the central engine results in a more luminous BLR. L_{X}

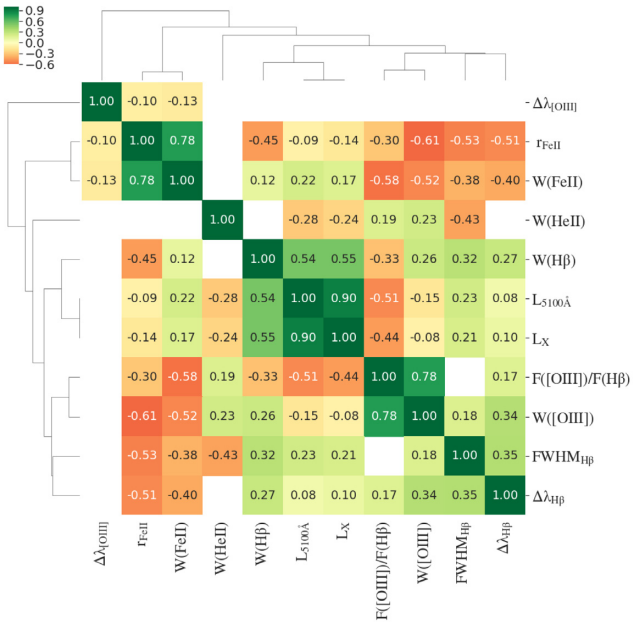

Figure 7. Spearman rank correlation matrix for our parameters. A positive (negative) coefficient indicates a positive (negative) monotonic correlation. Statistically insignificant correlations (with p-values >0.05/55) have been masked. The matrix was hierarchically clustered using correlation as a distance measure. A dendrogram displays the average linkage of the parameters. Previously reported correlations are confirmed for a larger range of redshift and luminosities.

and $L_{5100\,\text{Å}}$ decrease strongly ($\rho = -0.44$ and -0.51, respectively) with increasing flux ratio $F([\text{O III}])/F(\text{H}\,\beta)$.

(iv) The asymmetry index $\Delta\lambda_{\text{H}\,\beta}$ is positively correlated to $W([\text{O III}])$ ($\rho = 0.34$) and the $\text{FWHM}_{\text{H}\,\beta}$ ($\rho = 0.35$). We also report its relatively strong anticorrelation with $r_{\text{Fe II}}$ ($\rho = -0.51$) and $W(\text{Fe II})$ ($\rho = -0.40$). Despite using a different convention for the asymmetry index, Boroson & Green (1992) describe a similar correlation behaviour of $\Delta\lambda_{\text{H}\,\beta}$.

The key result here is not the rediscovery of correlations mapped out by, e.g. Boroson & Green (1992), Grupe (2004), and Shen & Ho (2014), but the fact that these relations hold for type 1 AGN in general, up to a redshift limit of at least $z = 0.80$ and for luminosities up to $L_{\text{X}} \sim 10^{46}\,\text{erg s}^{-1}$. In addition, a new insight provided by the correlation matrix is the correlation behaviour of $\Delta\lambda_{[\text{O III}]}$ with parameters related to the BLR emission. Namely, the asymmetry index of [O III] appears to be marginally related to the rest of the parameters, suggesting the absence of kinematic linkage between the inner region of the AGN and the NLR. This result is in contrast to Zamanov et al. (2002), who reported evidence for correlation between the shifts of the high ionization lines C IV λ1549 Å and the shifts of the [O III] lines, inferring a possible linkage between the NLR and BLR.

4.2 Partial correlation

In order to study the correlations with the $\Delta\lambda_{\text{H}\,\beta}$ parameter more closely, we perform a partial correlation analysis of $\text{FWHM}_{\text{H}\,\beta}$, $W([\text{O III}])$, $r_{\text{Fe II}}$, and $\Delta\lambda_{\text{H}\,\beta}$ (e.g. Baba, Shibata & Sibuya 2004). For this exercise, we first generate the Pearson correlation coefficients for this subset of parameters, and then measure the strength of each of these correlations while controlling for the other confounding

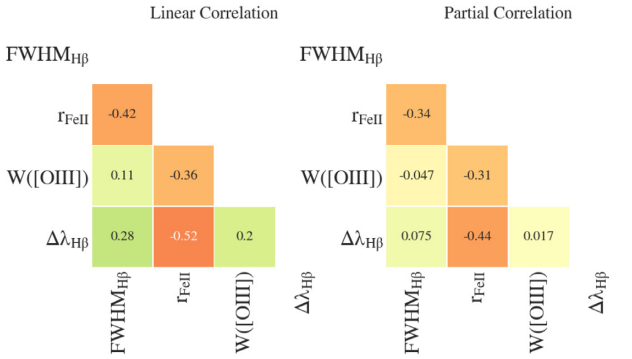

Figure 8. Left-hand panel: Pearson Correlation coefficients for the parameter subset: $\text{FWHM}_{\text{H}\,\beta}$, $W([\text{O III}])$, and $r_{\text{Fe II}}$, $\Delta\lambda_{\text{H}\,\beta}$. Right-hand panel: partial correlation coefficients (correlation coefficients, when the effects of confounding variables are removed).

variables. The results are shown in Fig. 8. When we control for $\text{FWHM}_{\text{H}\,\beta}$ and $r_{\text{Fe II}}$, the correlation between $W([\text{O III}])$ and $\Delta\lambda_{\text{H}\,\beta}$ significantly decreases (we measure a drop $\rho_P = 0.20$, $p < 10^{-7}$ to $\rho_{P,\,\text{partial}} = 0.017$, $p = 0.45$). The p-value of the latter partial correlation is above the corrected significance threshold: $p = 0.05/6$. Similarly, when we control for $W([\text{O III}])$ and $r_{\text{Fe II}}$, the strength of the $\text{FWHM}_{\text{H}\,\beta}$ versus $\Delta\lambda_{\text{H}\,\beta}$ correlation decreases ($\rho_P = 0.28$, $p < 10^{-6}$ to $\rho_{P,\,\text{partial}} = 0.075$, $p \sim 10^{-4}$). The anticorrelation between $\Delta\lambda_{\text{H}\,\beta}$ and $r_{\text{Fe II}}$ is, however, less affected when we control for $\text{FWHM}_{\text{H}\,\beta}$ and $W([\text{O III}])$ ($\rho_{P,\,\text{partial}} = -0.44$, $p < 10^{-6}$ instead of $\rho_P = -0.52$, $p < 10^{-6}$). These results demonstrate that $\Delta\lambda_{\text{H}\,\beta}$ parameter is principally related to the relative strength of the iron emission and is marginally linearly correlated to $\text{FWHM}_{\text{H}\,\beta}$, indicating that the broader profiles tend to be more red-skewed, as also suggested by Marziani et al. (2013a).

4.3 PCA

The second step consisted of running a PCA on the standardized data set to determine which parameters are contributing most to the total variance of our sources. It is particularly important to scale the parameters to unit variance since PCA is sensitive to the variance of the parameter distributions. PCA is an orthogonal linear transformation that is often used as a dimensionality-reduction algorithm (for a review, see Jolliffe & Cadima 2016). It yields the eigenvectors in parameter space, which point in the direction of total maximal variance in the data set. For a $N \times M$ data set, the first component is found by minimizing the distance between points in the M-dimensional parameter space and their orthogonal projections on to an M-dimensional vector, which simultaneously increases the variance of the projected points. The second component is chosen to be orthogonal to the first and is determined in the same manner given this condition. The associated eigenvalues measure the amount of explained variance in each principal component (eigenvector). The variables (our parameters) are linked to the principal components by linear coefficients. These coefficients yield the amount of a variable's variance explained by the component.[8] While PCA is a very common tool in astronomy it comes with several disadvantages, such as the linearity of the dimensionality reduction, which might not capture the full complexity of our

[8]In practice, we made use of a SCIKIT-LEARN implementation of PCA (Pedregosa et al. 2011).

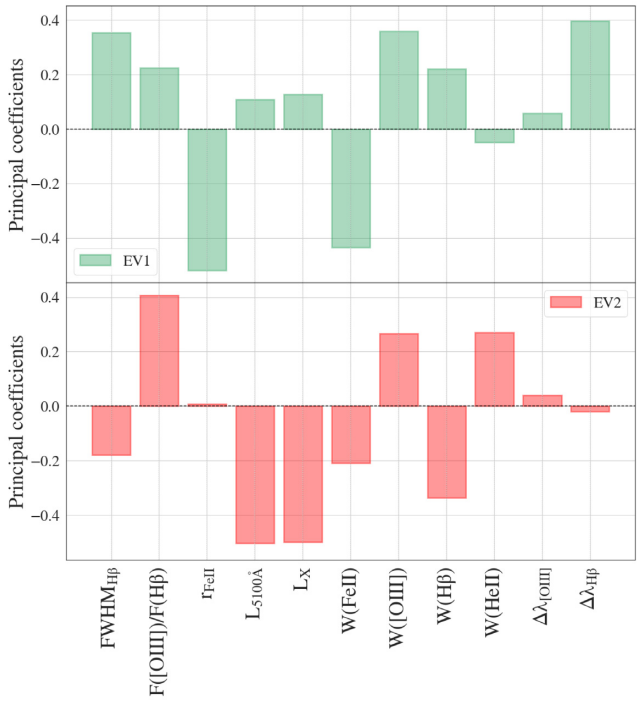

Figure 9. The component coefficients (factor loadings) defining the first and second principal components: EV1 (green bars) and EV2 (red bars). EV1 appears to be heavily dominated by the anticorrelation of $\mathrm{FWHM_H}$ and the strength of the [O III] emission with $r_{\mathrm{Fe\,II}}$. A relatively large linear coefficient links the H β asymmetry to EV1. EV2 is heavily dominated in equal measure by the X-ray and optical luminosities.

data. One must in general be careful with the choice of the particular dimensionality reduction one uses and its interpretation (e.g. see Baron 2019). The first two principal components are presented in Fig. 9. The bar diagrams show the values of the correlation coefficients, which link the parameters (our variables) to the given component. The first and second principal components, respectively, explain 25.2 and 20.3 per cent of the total variance in the selected parameter subset. The first principal component (EV1) is anticorrelated to the strength of the iron emission and correlated to $\mathrm{FWHM_{H\,\beta}}$ and the (relative) strength of the [O III] emission, i.e. the diversity in the selected optical features of our SPIDERS AGN subsample is dominated by the anticorrelation of W([O III]) and W(Fe II). The Balmer profile asymmetries $\Delta\lambda_{\mathrm{H}\,\beta}$ are noticeably correlated to EV1, which corroborates the findings of Z10, who stress that the asymmetry parameter could in essence be used as a surrogate 4DE1 parameter. Z10 note, however, that $\Delta\lambda_{\mathrm{H}\,\beta}$ is also substantially orienting eigenvector 2 (EV2). This behaviour is not observed in our PCA results. Our EV1 is consistent with early results in Grupe (2004). The second principal component, EV2, is strongly dominated by the X-ray and optical continuum luminosities. The flux ratio F([O III])/F(H β) is correlated to EV2. W(H β) once again significantly contributes to the orientation of the principal component. EV2 is also consistent with Grupe (2004).

5 RELATING EV1 AND EV2 TO PHYSICAL PARAMETERS

We now want to extend the analysis on derived AGN properties such as M_{BH} and L/L_{Edd}. In Boroson & Green (1992), EV1 was related to L/L_{Edd}, tracing the accretion power of the observed SMBH. Their EV2 was largely dominated by luminosity and related to the Baldwin effect (Baldwin 1977; Baldwin et al. 1978; Dietrich et al. 2002). L/L_{Edd} has indeed long been considered a prime candidate to explain the observed diversity in optical AGN features (e.g. Sulentic et al. 2000b; Yuan & Wills 2003; Grupe 2004; Kuraszkiewicz et al. 2009; Grupe 2011; Shen & Ho 2014; Bon, Bon & Marziani 2018).

The FWHM of broad emission lines serves as a proxy for the line-of-sight velocity of the BLR gas, which in combination with the distance to the black hole yields an estimate for its M_{BH} by

$$M_{\mathrm{BH}} = \frac{r_{\mathrm{BLR}}\, f_{\mathrm{BLR}} \mathbf{FWHM}_{\mathrm{BLR}}{}^2}{G}, \tag{3}$$

where f_{BLR} denotes the geometric form factor of the BLR. The distance r_{BLR} from the SMBH to the line-emitting region is obtained from the delay in response of the line fluxes to continuum changes. The $\mathrm{FWHM_{BLR}}$ of a line from the emitting region corresponds to the radial-velocity component of the BLR gas. The Keplerian velocity v_{Kepl} of the gas is obtained by correcting for projection effects encoded in the geometric form factor: $v_{\mathrm{Kepl}}{}^2 = f_{\mathrm{BLR}}\mathrm{FWHM_{BLR}}{}^2$. One can construct an expression of f_{BLR}, which takes into account the source orientation θ and the shape of the velocity field described by the motion of the gas in the BLR (e.g. equation 12 in Negrete et al. 2018), by making assumptions about the structure of the disc. M_{BH} and L/L_{Edd} for our sample are derived from H β using the calibration developed in Assef et al. (2011), who based their estimates on the BLR radius–luminosity relation of Bentz et al. (2009).

$$\log\left(\frac{M_{\mathrm{BH}}}{\mathrm{M}\odot}\right) = A + B\log\left(\frac{\lambda L_\lambda}{10^{44}\,\mathrm{erg\,s^{-1}}}\right) + C\log\left(\frac{\mathrm{FWHM_{H\,\beta}}}{1000\,\mathrm{km\,s^{-1}}}\right). \tag{4}$$

L_λ corresponds to the monochromatic luminosity at 5100 Å and the parameters A, B, and C are derived from reverberation mapping studies: $A = 0.895$, $B = 0.52$, and $C = 2$. From these black hole masses, C19 provided estimates of L/L_{Edd} following

$$L_{\mathrm{Bol}}/L_{\mathrm{Edd}} = \frac{L_{\mathrm{Bol}}}{4\pi c \mathrm{G} M_{\mathrm{BH}} m_{\mathrm{p}}/\sigma_{\mathrm{T}}}, \tag{5}$$

where G is the gravitational constant, c is the speed of light, m_{p} is the proton mass, and σ_{T} is the Thomson scattering cross-section. The bolometric luminosities L_{Bol} used here are obtained from the bolometric corrections presented in Richards et al. (2006): $L_{\mathrm{Bol}} = 9.26 L_{5100\,\text{Å}}$. This factor is in principle dependent on $L_{5100\,\text{Å}}$, but for the purpose of this work, this assumption is not crucial.

Fig. 10 displays how L/L_{Edd} and M_{BH} scale with the projection of the sources on to our newly determined EV1 and EV2. We can strongly corroborate that EV1 correlates with L/L_{Edd}. An even stronger anticorrelation is that of EV2 with black hole mass. However, the log M_{BH}–EV2 relation breaks down at lower redshifts ($z \sim 0.2$). In order to determine to which extend this correlation is a side-product of the $L_{5100\,\text{Å}}$ and $\mathrm{FWHM_{H\,\beta}}$ principal coefficients in EV2, we once again perform a partial correlation test. We start by measuring the Spearman rank correlation of EV2 with black hole mass, while controlling for $L_{5100\,\text{Å}}$ and $\mathrm{FWHM_{H\,\beta}}$. In Fig. 11, the Spearman correlation coefficients measured for log M_{BH}, EV2, $L_{5100\,\text{Å}}$, and $\mathrm{FWHM_{H\,\beta}}$ are displayed in the left-hand panel. In the right-hand panel of Fig. 11, the partial correlation coefficient results are presented. As already discussed in Section 4, this exercise consisted in performing a (multi) linear regression of EV2 and log M_{BH} with the confounding variables. In this version, we compute the Spearman correlation coefficients of the residuals of

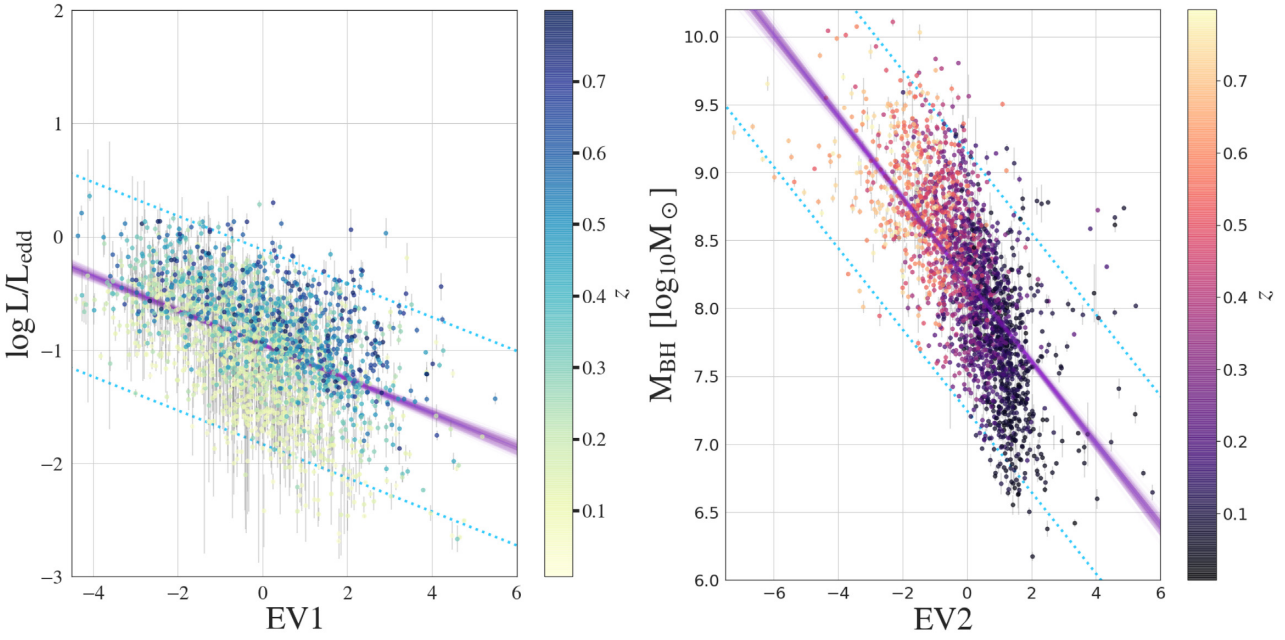

Figure 10. *Left-hand panel*: $L/L_{\rm Edd}$ has an exponential correlation to EV1. The 95 per cent confidence contours are represented by the dashed blue lines. *Right-hand panel*: $M_{\rm BH}$ is correlated to EV2. Again, the 95 per cent confidence contours are shown in blue. The linear regressions were performed with *Linmix* (Kelly 2007).

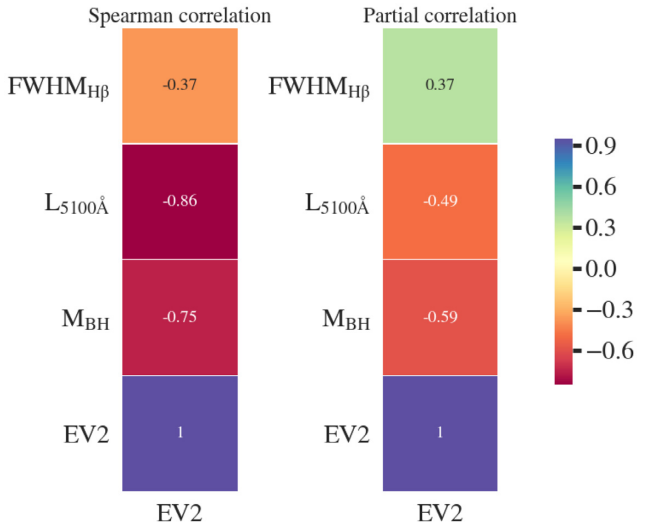

Figure 11. The monotonic correlation coefficients of EV2 with $M_{\rm BH}$, $L_{5100\,\text{Å}}$, and $\rm FWHM_{H\,\beta}$ for our sample are shown: simple Spearman correlation in the left column and partial Spearman correlation in the right column. For the computation of the partial correlation of a parameter pair, we control for the other confounding variables. The $M_{\rm BH}$ versus EV2 anticorrelation slightly drops when we control for $L_{5100\,\text{Å}}$ and $\rm FWHM_{H\,\beta}$, while partial $M_{\rm BH}$ versus EV2 anticorrelation remains significantly strong, $p = -0.59$.

these two multilinear regressions. EV2 is still strongly anticorrelated to $M_{\rm BH}$ ($\rho = -0.56$). This is not surprising since in equation (4), log $M_{\rm BH}$ is a non-linear function of $\rm FWHM_{H\,\beta}$ and $L_{5100\,\text{Å}}$, while the eigenvectors obtained from PCA are obtained through linear orthogonal transformations of the initial parameter. The monotonic anticorrelation measured in the residuals might thus simply arise from the non-linearity of equation (4). More generally, this result implies that one can construct a linear combination of the parameters that dominate the orientation of EV2 (see Fig. 9) in order to estimate $M_{\rm BH}$ at redshifts $z > 0.2$.

We propose that EV2, through its anticorrelation to black hole mass, might be related to the evolution of the broad-line AGN population. Over the observed redshift range, the median $M_{\rm BH}$ increases with decreasing redshift. This trend is a combination of two effects: the *ROSAT*/2RXS X-ray flux limit resulting in the Malmquist bias and the decreasing number density of high black hole masses at lower redshifts. This downsizing of black hole masses across cosmic time remains a matter of debate (cf. downsizing models, e.g. Fanidakis et al. 2012). We conclude that $L/L_{\rm Edd}$ and $M_{\rm BH}$ are related to the principal components, which explain ∼45 per cent of the total variance in our data.

6 ASYMMETRY OF THE BROAD Hβ EMISSION LINE; A MARKER OF TYPE 1 AGN DIVERSITY

6.1 Asymmetries along the EV1 plane

We investigate the role played by H β line asymmetries in this framework. Since they appear to be one of the central sources of variance in our sample, they can be a source of the optical diversity of Type 1 AGNs. Fig. 12 presents the sample in the $\rm FWHM_{H\,\beta}$ versus $r_{\rm Fe\,II}$ plane, also called the EV1 plane (see the previous section). The sources have been separated in two subsets: blue- and red-asymmetric H β emission. More specifically, these two subsamples were constructed according to: $\Delta\lambda_{\rm H\,\beta} > 0.07$ and $\Delta\lambda_{\rm H\,\beta} < -0.07$. These criteria exclude symmetric sources and account for the typical uncertainties of the asymmetry index ($\sigma_{\Delta\lambda_{\rm H\,\beta}} \sim 0.05$). Contours of the bivariate Gaussian kernel density estimates (KDE; e.g. Silverman 1986) for each of the subsets are also indicated. The contours delimit the areas containing 68, 86,

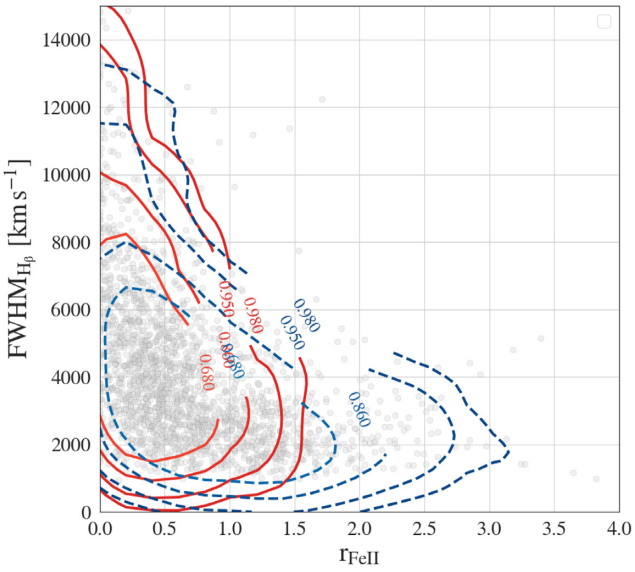

Figure 12. The horizontal trend of the H β asymmetry in the EV1 plane is made clear in the above figure. The contours delimit the areas in which 68, 86, 95, and 98 per cent of the data points are confined. These contours reveal two very different occupation domains for red- and blue-asymmetric Balmer emitters: sources with red-asymmetric H β are confined at low-$r_{\rm Fe\,II}$ values, while sources with blue-asymmetric H β seem to extend over the full EV1 sequence.

95, and 98 per cent of the data points.[9] The sample occupies an L-shaped form in the plane. Using simulations, C19 demonstrate that the absence of high-FWHM$_{H\beta}$ and high-$r_{\rm Fe\,II}$ sources is to an extent due to the limitations of the spectral fitting method. Most of the sources with blue-asymmetric H β emission profiles and those with red-asymmetric profiles appear to occupy different domains of this sequence. While the AGNs with blueward asymmetric H β spread over the full $r_{\rm Fe\,II}$ range and a large portion of the FWHM$_{H\beta}$ range, the redward asymmetric sources are concentrated at lower values of $r_{\rm Fe\,II}$, while dominating the higher segment of FWHM$_{H\beta}$. This result is essentially suggesting that strong accretors, with relatively large $r_{\rm Fe\,II}$ values, show mainly blueward asymmetric H β profiles. We also observe a considerable overlap of red and blue asymmetries for sources with moderate FWHM$_{H\beta}$ and $r_{\rm Fe\,II}$.

In their in-depth study of the H β line profiles of ∼470 low-z SDSS (DR5) quasars, Z10 show that low-accretion rate sources (low $r_{\rm Fe\,II}$) possess a typically red-asymmetric H β profile, while high-accretion rate sources tend to prefer blue-asymmetric H β profiles. The two panels of Fig. 13 can directly be compared to Figs 6(a) and 9(a) presented in their work:

(i) The size of our sample allows us to extend their results by affirming that blue-asymmetric H β profiles can be found in sources with relatively high FWHM$_{H\beta}$ (∼8000 km s^{-1}) and along the full EV1 sequence. There is, however, a clear decrease in blue-asymmetric Balmer profiles at FWHM$_{H\beta}$ ∼ 4000 km s^{-1}, consistent with the low-redshift separation in Population A/B introduced by Sulentic et al. (2000a). An excess of red-asymmetric H β profiles is indeed observed for larger widths. Sources with lower FWHM$_{H\beta}$ do not, however, display more symmetric profiles. We thus argue that there is no evidence for a systematic relationship between FWHM$_{H\beta}$ and $\Delta\lambda$. The excess of red-asymmetric H β emission for larger widths might be the signature of an additional redshifted broad emission component.

(ii) The right-hand panel of Fig. 13 strongly confirms the decreasing trend presented in fig. 5 of Boroson & Green (1992) and fig. 9(a) of Z10. Combined with the Spearman correlation coefficient of $\rho \sim -0.5$ reported in the previous section, we can state that, up to redshift of $z \sim 0.8$, low accretors tend to show more red-asymmetric profiles, with a significant decrease of red-asymmetric H β emitters beyond $r_{\rm Fe\,II} \sim 1.5$. While blue-asymmetric profiles are reported over the full $r_{\rm Fe\,II}$ range, the shifts become more pronounced at higher values of the Fe II–H β flux ratio.

Z10 discuss the VBLR emission as potential origin of typically red asymmetries in H β (2.2). In a layered model of the BLR, the asymmetry index traces the displacement between the H β BC and its VBC. The H β lines often show a broad, redshifted, low-intensity feature in their profiles (Peterson & Ferland 1986; Corbin 1995; Brotherton 1996). The presence of a distinct emission region, the VBLR (Popović et al. 2004; Marziani et al. 2010), has been proposed to explain the characteristic red wing of broad Balmer emission. The VBLR is expected to be located in the inner region of the AGN: The typical widths of the VBC in H β indicate much higher Keplerian velocities than the classical BLR. The presence of a VBC in the profile shapes has to be corrected for when broad low-ionization lines are used as virial broadening estimator (Marziani et al. 2013a). In Appendix B, Fig. B1, three sources of our sample, which have one broad Gaussian component with FWHM > 10 000 km s^{-1}, are presented.

6.2 Tracing broad component displacements with centroid shifts

In order to further characterize the kinematics of two potentially distinct emitting regions, we compute the centroid shifts of the H β profiles at 15 (c_{15}) and 80 per cent (c_{80}) fractional intensity. Fig. 14 displays the correlations between these shifts and the asymmetry index of H β, colour coded according to the FWHM$_{H\beta}$. The shifts with respect to the rest frame at 80 per cent fractional intensity are negatively related to the asymmetry index, i.e. the more blueshifted the top of the profile, the more redward asymmetric the complete profile and vice versa. Symmetric profiles range over the full shift range. For the centroid shifts at c_{15} fractional intensity, an opposite and stronger trend is observed. c_{15} correlates positively with $\Delta\lambda_{H\beta}$: a redward asymmetric H β profile can be associated with a redshift of the broad component. As expected, sources with symmetric H β span over the full c_{15} range. The highest values of FWHM$_{H\beta}$ occur for red-asymmetric H β, consistent with the results in Fig. 13. A clear picture arises from the c_{15} distribution in the EV1 plane (Fig. 15). The absolute value of the shifts at the profile base appears to decrease along the EV1 sequence, while c_{15} shows clear trends with both $r_{\rm Fe\,II}$ and FWHM$_{H\beta}$. This picture is consistent with the presence of a redshifted VBLR. The base shifts of the H β profiles in our sample are tightly correlated to the asymmetry index $\Delta\lambda_{H\beta}$. The largest shifts are observed for the most red-asymmetric profiles. The highest base component redshifts are found at the top of the Type 1 AGN main sequence, i.e. at the largest widths of H β; the broadest components in our sources' H β profiles are preferentially redshifted.

The correlation of c_{15} with EV1 provides additional information: As can be seen in Fig. 16, the positive base centroid shifts appear

[9]The density contours shown throughout the rest of this work are defined in the same way.

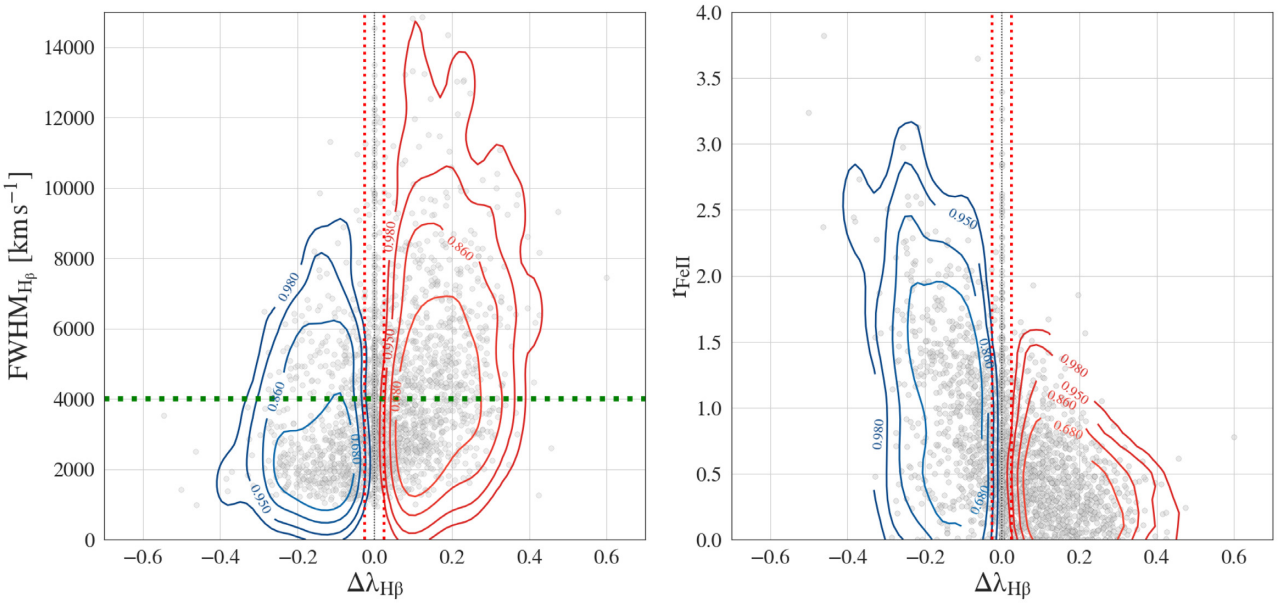

Figure 13. *Left-hand panel:* FWHM of H β as a function of the Balmer asymmetry index. The kernel density contours separate blue- from redward asymmetric H β-emitting sources. The green dotted line indicates the separation in Pop.A/B at FWHM$_{H\beta} \sim 4000\,\mathrm{km\,s^{-1}}$. *Right-hand panel:* $r_{Fe\,II}$ as a function of the Balmer asymmetry index. The kernel density contours separate blue- from redward asymmetric H β-emitting sources. For both the figures, the typical 1σ errors are shown centred at zero by two dotted vertical bars.

to correlate with EV1. Interestingly, the broad redshifted H β components are found for the sources with the highest black hole masses. The broad component redshifts are thus a source of variance in our sample. If the redshifted VBCs are interpreted as the signature of a VBLR, we can argue that the presence or absence of a distinct, inner emitting shell in the BLR (or more precisely its redshift) is a strong source of diversity in our sample of Type 1 AGNs. Furthermore, we report that our results are consistent with Marziani et al. (2009), who find a systematic increase of FWHM$_{H\beta}$ with source luminosity. More precisely, Marziani et al. (2009) find a scaling relation between the luminosity and the contribution of the VBLR to the full H β flux in population B ($> 4000\,\mathrm{km\,s^{-1}}$). The absence of the red-asymmetric H β lines at higher $r_{Fe\,II}$ shown in the right-hand panel of Fig. 13 might be due to strong accretion disc winds preventing the formation of a VBLR.

In Appendix C, we explore a toy model of BLR obscuration, which accounts for the presence of a VBLR.

7 BLUE-ASYMMETRIC Hβ: OUTFLOWS IN A FLATTENED AND STRATIFIED BLR MODEL

7.1 Blue asymmetries, outflows, and self-shielding

If we consider that blue- and redward asymmetric H β profiles are the signatures of two distinct kinematic states of the BLR, we should divide our sample according to the asymmetry indices of the sources and investigate the properties of each subsample. As in the previous section, we formed two subsamples with $\Delta\lambda_{H\beta} < -0.07$ (blue asymmetric) and $\Delta\lambda_{H\beta} > 0.07$ (red asymmetric). We constructed a Spearman rank correlation matrix for the parameter subset, as defined in Section 2.1, of each subsample and directly compared their correlation spaces. A striking difference is observed for the correlation of the optical and X-ray luminosity and the iron emission. For blue-asymmetric H β emitters, L_X and $L_{5100\,Å}$ correlate positively with the equivalent width of the iron emission W(Fe II) (Spearman correlation coefficients: $\rho_S = 0.47$ and 0.55, respectively). The sources with red-asymmetric H β have a much weaker correlation between the source luminosities and W(Fe II) ($\rho_S = 0.023$ for L_X and $\rho_S = 0.097$ for $L_{5100\,Å}$). The correlation of $L_{5100\,Å}$ and W(Fe II) for red-asymmetric sources is $p = 0.53$ and can safely be considered as insignificant. This contrasting behaviour is presented in the left-hand panel of Fig. 17. In order to confirm the different clustering of red- and blue-asymmetric H β emitters in the plane spanned by $L_{5100\,Å}$ and W(Fe II), we binned log W(Fe II) in the range 1.0–1.9 and bootstrapped the subsamples in each bin (10 000 resamples). We obtained the mean $L_{5100\,Å}$ for each bin from the sampling distribution of the means. The 68 per cent confidence intervals were derived using the percentile method (from the 16.0 and 84.0 per centpercentiles). This method should yield relatively good estimates of the error, given the symmetric shape of the sampling distribution of the mean in each bin (shown in the right-hand panel of Fig. 17). The positive correlation between $L_{5100\,Å}$ and W(Fe II) for sources with blue-asymmetric H β is confirmed. No trend is found for sources with red-asymmetric H β. The error bars, showing the standard deviation of the bootstrap samples' means, do not overlap. The correlation of $L_{5100\,Å}$ and W(Fe II) for the blue-asymmetric Hβ population withstood a partial correlation test, where we controlled for the confounding W(H β). Similarly, when we control for the FWHM$_{Fe\,II}$ of the Gaussian kernel used to fit Fe II, $L_{5100\,Å}$ and W(Fe II) remain positively correlated. The dependence of the Fe II flux on the continuum luminosity in blue-asymmetric H β emitters is only marginally related to the broadening of the iron lines. This effect also manifests itself in the correlation of the luminosity parameters with the flux ratio $r_{Fe\,II}$. Blue-asymmetric H β emitting sources show a significant correlation between $r_{Fe\,II}$ and central source luminosities ($\rho_S = 0.20$ for $L_{5100\,Å}$ and $\rho_S = 0.19$ for L_X). For red-asymmetric H β sources, these quantities are anticorrelated ($\rho_S = -0.16$ for $L_{5100\,Å}$ and $\rho_S = -0.25$ for L_X). Another noticeable difference is observed for the $\Delta\lambda_{[O\,III]}$ parameter. In the subsample of blue-asymmetric Balmer emitters, $\Delta\lambda_{[O\,III]}$ and $F([O\,III])/F(H\,\beta)$

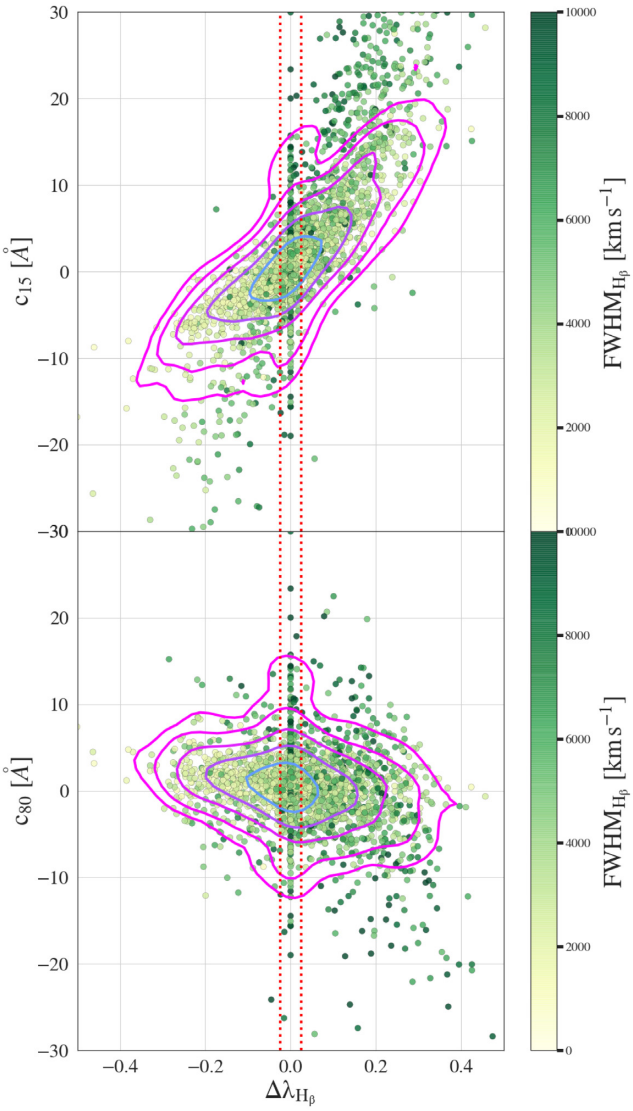

Figure 14. The H β centroid shifts at 15 and 80 per cent fractional intensity as a function of H β asymmetry. c_{15} appears more tightly related to $\Delta\lambda_{H\beta}$ than c_{80}. The kernel density contours show the subsample of sources, for which H β has not been fitted with broad Gaussians (FWHM $<$ 10 000 km s^{-1}). Typical 1σ uncertainties in the measurement of $\Delta\lambda_{H\beta}$ are indicated as vertical red dotted lines.

are anticorrelated ($\rho_S = -0.26$). If blue shifted wings are interpreted as the signature of NLR outflows, this result would imply that outflow velocities increase with increasing $F([O\,III])/F(H\,\beta)$. In the red-asymmetric subsample, these two parameters are positively correlated ($\rho_S = 0.17$).

The blueward asymmetries of low-ionization lines have been related to radiation-driven outflows (Marziani et al. 2013b). They occur in the high-$r_{Fe\,II}$ bins of population A (e.g. Ganci et al. 2019), which contain the highest accretion rate sources along the main sequence (e.g. Sun & Shen 2015; Sulentic et al. 2017; Panda et al. 2019c). The origin of Fe II emission in the BLR has been a long-standing matter of research. Several lines of evidence support an emission of Fe II in the outer parts of the BLR, while H β might be emitted closer to the black hole (e.g. Rodríguez-Ardila et al. 2002; Barth et al. 2013; Marinello et al. 2016). The physical conditions for iron ionization have been investigated in detail. Wills, Netzer & Wills (1985) showed that photoionization models might not suffice to account for the total observed Fe II flux in AGNs. For a review of necessary conditions for the formation of the observed Fe II in photoionization models, see Collin & Joly (2000). Models of collisional excitation (e.g. Baldwin et al. 2004; Joly, Véron-Cetty & Véron 2007), in addition to continuum and line fluorescence (e.g. Sigut & Pradhan 1998, 2003; Marinello et al. 2016), have been considered. These models make predictions on the physical conditions of the Fe II-emitting region, such as shielding from the continuum source and high densities. A more in-depth study of Fe II emission strength should also account for the orientation of the flattened BLR with respect to the observer. This has been recently investigated for narrow-line Seyfert 1 objects by Panda, Marziani & Czerny (2019a) and Panda, Marziani & Czerny (2019b), whose preliminary analysis leads to conclusions on the density of the emitting region.

In a flattened, horizontally stratified cloud distribution in Keplerian motion around the central black hole, the difference in sensitivity to continuum of the Fe II-emitting region might arise from different degrees of exposure to the central continuum source of the dense clouds from which the iron is emitted. Radiation-driven winds produce outflows of H β-emitting gas, thereby exposing the previously shielded Fe II regions more directly to continuum emission. Using the photoionization code CLOUDY (Ferland et al. 2013), Panda, Czerny & Wildy (2017) actually found a decrease of the Fe II-to-H β ratio with increasing disc temperature – a function of accretion luminosity. A similar experiment accounting for BLR stratification and outflowing components would allow us to test the scenario we propose. The scaling between W(Fe II) and the $L_{5100\,Å}$ might be due to the contribution of UV fluorescence to the excitation of Fe+ levels. Investigating Fe II fluorescence in H II regions, Rodríguez (1999) uses the sensitivity to fluorescence of [Fe II] λ4287, as well as the continuum-insensitive [Fe II] λ8617, to investigate the role played by the UV radiation field in the formation of this line. If the presented model of unshielding holds, the intensity ratio I([Fe II] λ4287)/I([Fe II] λ8617) should scale differently with the UV/optical continuum luminosity for the red- and blue-asymmetric H β populations. This may be investigated in further work. In this scenario, the distance of the Fe II region to the SMBH should not depend on the presence of outflows. We performed a two-sample Anderson–Darling test on the distributions of the FWHM$_{Fe\,II}$ of the Gaussian kernel that was convolved with the Fe II template in the red- and blue-asymmetric H β subsets. 55 sources in our sample have no Fe II flux detection and were not considered for this test. The obtained p-value is $p > 0.25$. We thus cannot reject that the FWHM$_{Fe\,II}$ for red- and blue-asymmetric H β sources were sampled from the same distribution. Under the assumption of a Keplerian velocity field, the radial distribution of the Fe II-emitting clouds is not affected by the presence of outflows.

We conclude that our data are consistent with a flat, stratified and self-shielding model of the BLR, in which the Fe II and H β emissions originate at different radii. When the inner clouds are driven out of the plane by outflows during strong accretion events, the previously shielded, neutral and dense Fe II clouds are more exposed to the ionizing continuum radiation, which is clearly seen in the Fe II flux–luminosity scaling.

7.2 Evidence for model degeneracy: Fe II versus H β

We note that C19 fitted the emission of the iron complex over the full H β centred region, using the I Zw 1 template (Boroson &

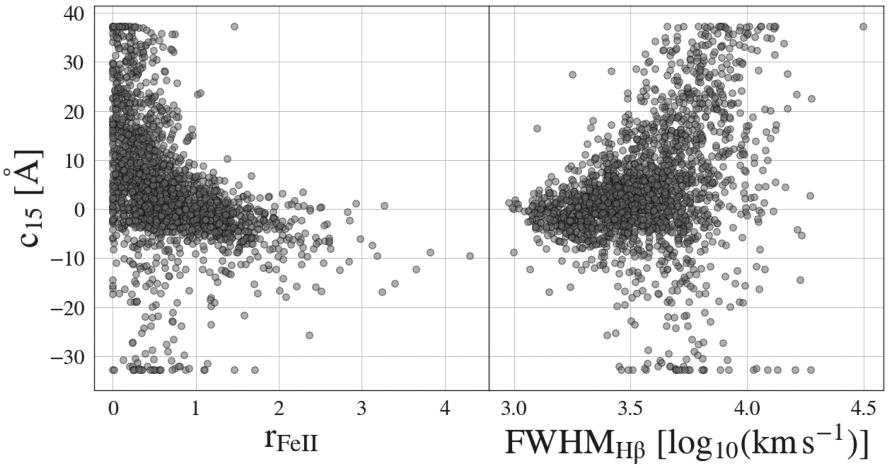

Figure 15. The individual correlations of c_{15} with the two optical 4DE1 dimensions. The absolute centroid shifts decrease along the EV1 sequence.

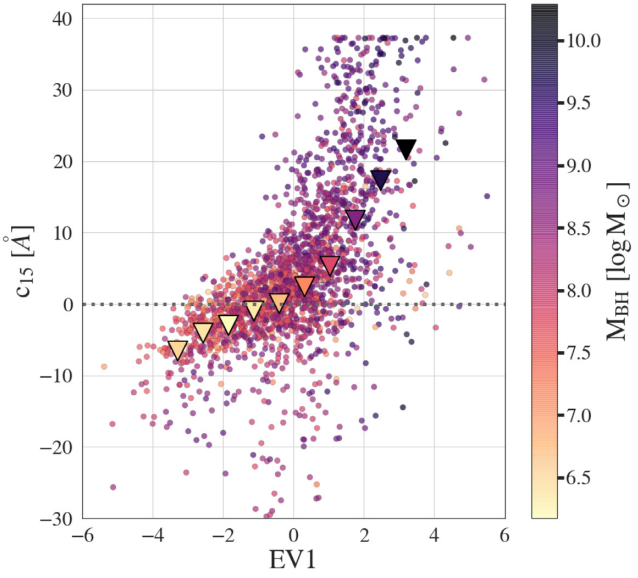

Figure 16. H β base centroid shifts (c_{15}) projected on to EV1. The points are colour coded according to the black hole masses. Positive centroid shifts (redshifts) correlate with EV1. The coloured triangles represent the median of c_{15} in EV1 bins and have been added to improve the visualization of this trend. Their colour represents the median black hole masses in the bins, with darker colours corresponding to higher black hole masses. The highest black hole masses are found for redshifted base components.

Green 1992). Since all the emission features of Fe II in the spectral window were taken into account during the fits, the effect of contamination by H β should be marginal on the total Fe II flux, i.e. W(Fe II) should not be strongly affected by increasing H β line flux. However, we can investigate if the Fe II contaminates the H β fit on its blue side. For blue-asymmetric H β emitters, we seek to test if the W(Fe II) correlates with $\Delta\lambda_{\mathrm{H}\beta}$ asymmetry index, i.e. we want to determine if the excess flux on the blue side of H β is due to contamination by iron. For the subsample with $\Delta\lambda_{\mathrm{H}\beta} < -0.07$, we measure the Spearman rank correlation coefficient of $|\Delta\lambda_{\mathrm{H}\beta}|$ and W(Fe II). We obtain $\rho = 0.087$ with a p-value of $p = 0.08$. Even if we do not correct the significance threshold for multihypothesis testing, one can state that there is no significant correlation between $\Delta\lambda_{\mathrm{H}\beta}$ and W(Fe II). We note, however, that W(H β) is a confounding variable for both $\Delta\lambda_{\mathrm{H}\beta}$ ($p = -0.26$) and W(Fe II) ($p = 0.29$). We thus have to once again perform a partial correlation analysis, marginalizing over W(H β) in order to obtain the real correlation behaviour of the blue asymmetries and the iron strength. The partial correlation coefficient is $p_P = 0.17$ with a p-value of $p \sim 10^{-4}$. There is a significant correlation between $\Delta\lambda_{\mathrm{H}\beta}$ and W(Fe II) once we control for W(H β). We thus cannot exclude that the iron complex contaminates the blue wing of H β.

8 CONCLUSIONS

In this work, we have performed a statistical analysis of a high signal-to-noise ratio (S/N > 20) subsample of the SDSS-IV/SPIDERS DR14 VAC of X-ray-selected Type 1 AGN. The sample included 2100 sources spanning an X-ray luminosity range of $L_{\mathrm{X}} = 1.9 \times 10^{41}$–$9.9 \times 10^{45}$ erg s^{-1}, up to redshift $z = 0.80$. It is the largest sample of high signal-to-noise ratio sources used for the PCA-based, statistical study of Type 1 AGN spectral properties to date.

(i) We used PCA as a central tool to determine the source of variance in our data and have mapped a correlation space that is remarkably consistent with previous studies of the optical EV1 (e.g. Boroson & Green 1992; Sulentic et al. 2000a; Grupe 2004; Shen & Ho 2014), while probing a larger cosmological volume.

(ii) We confirm that the Eddington ratio and the black hole mass are significantly related to the observed diversity of Type 1 AGNs, through their correlation to EV1 and EV2.

(iii) Studying H β line shapes in this context, we find blue-asymmetric emission profiles for the full Type 1 AGN main sequence, while red asymmetries only appear for low accretors. Z10, investigating H β profile asymmetries in the 4DE1 context, have suggested that profiles with lower FWHM$_{\mathrm{H}\beta}$ tend to be more symmetric while profiles with larger FWHM$_{\mathrm{H}\beta}$ are preferentially red asymmetric. The larger number of sources in our sample enabled us to complete this picture: While we do observe a larger portion of sources with red asymmetry index at high FWHM$_{\mathrm{H}\beta}$, lower width profiles appear to cover the full range of asymmetries observed in our sample, i.e. lower FWHM$_{\mathrm{H}\beta}$ do not lead to more symmetric profiles. We can, however, confirm a strong trend between $r_{\mathrm{Fe\,II}}$ and $\Delta\lambda_{\mathrm{H}\beta}$. In their discussion of the physical origins of H β profile

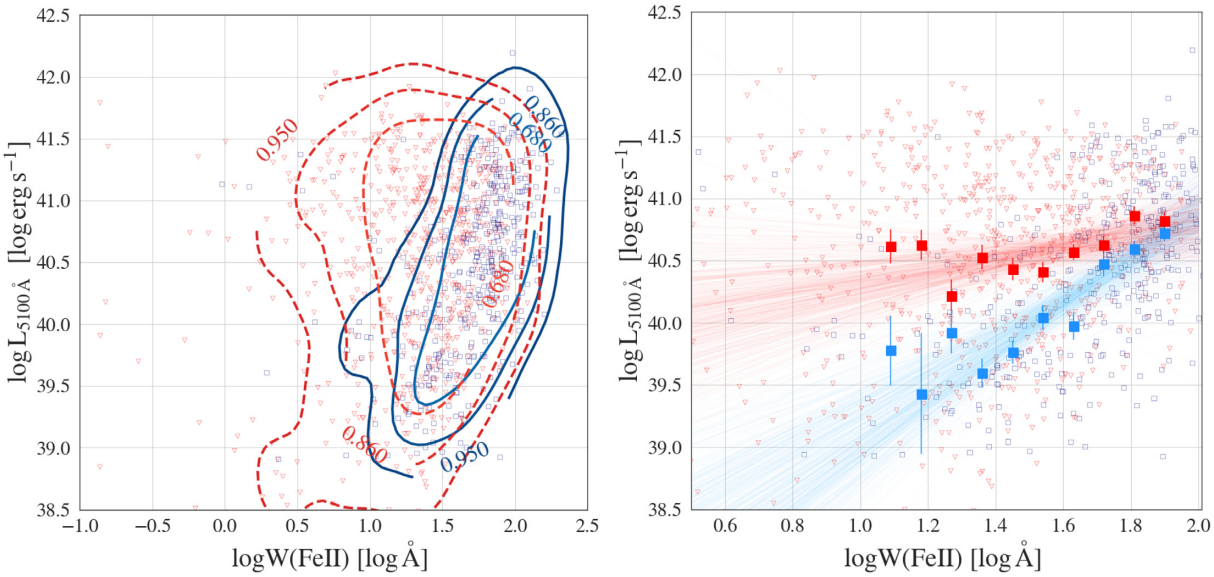

Figure 17. *Left-hand panel*: The monochromatic luminosity at 5100 Å as a function of the equivalent width of the Fe II-blend at 4570 Å. The red triangles and blue squares distinguish red- ($\Delta\lambda > 0.07$) from blue-asymmetric ($\Delta\lambda < -0.07$) H β emitters. Kernel density contours have been included. A tighter relation is observed for sources with blue-asymmetric Balmer emission. A similar picture arises in the case of $L_{0.1-2\,\mathrm{keV}}$–W(Fe II). *Right-hand panel*: The L_{5100}–W(Fe II) plane binned in 10 log$W$(Fe II) bins between 1.0 and 1.9. For each bin, 10 000 bootstrap samples were obtained, for which the mean $L_{5100\ \text{Å}}$ was derived. The points and error bars represent the mean of the sampling distribution of the mean. The error bars correspond to the 1σ error derived from the sampling distribution of the mean, using the percentile method. The linear regression was performed with *Linmix* (Kelly 2007).

shapes, Z10 identify VBLR emission and disc winds as main candidates.

(iv) A subclass of our sources does indeed show very shifted, broad components in their H β emission, possibly due to the presence of a distinct emitting gas distribution in the inner regions of the BLR. The redshift of this VBLR correlates with EV1 and might thus be related to Type 1 AGN diversity.

(v) Exploring parameter correlations for blue- and redward asymmetric H β emitters separately, we observed that Fe II line flux correlates differently with source luminosity for red- and blueward asymmetric H β emitters. We discussed this effect in the light of a flattened, self-shielding BLR, in which the Fe II-emitting clouds are located at larger radii than the ones emitting the Balmer lines. The outflows, which we associate H β blueward asymmetries with, might in such a configuration strip the Fe II region of its shield, i.e. the innermost parts of the BLR. However, we find evidence for interline contamination between Fe II and H β, which might play a confounding role in this effect.

ACKNOWLEDGEMENTS

We thank the anonymous referee for comments that significantly improved the quality of this paper.

We would like to thank Hagai Netzer for insightful and stimulating conversations on the physics of the BLR. We also thank Claudia Scarlata for her precious input on machine learning techniques and Tom Dwelly for helpful advice.

FJC acknowledges financial support through grant AYA2015-64346-C2-1P (MINECO/FEDER).

FJC also acknowledges funding from the European Union's Horizon 2020 Programme under the AHEAD project (grant agreement no. 654215).

Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU)/University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatório Nacional/MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

Funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) under Germany's Excellence Strategy – EXC-2094 – 390783311.

REFERENCES

Abolfathi B. et al., 2018, ApJS, 235, 42

Adhikari T. P., Różańska A., Czerny B., Hryniewicz K., Ferland G. J., 2016, ApJ, 831, 68

Ameijeiras-Alonso J., Crujeiras R. M., Rodríguez-Casal A., 2019, TEST, 28, 900
Antonucci R., 1993, ARA&A, 31, 473
Assef R. J. et al., 2011, ApJ, 742, 93
Baba K., Shibata R., Sibuya M., 2004, Aust. N.Z. J. Stat., 46, 657
Bahcall J. N., Kozlovsky B.-Z., Salpeter E. E., 1972, ApJ, 171, 467
Baldwin J. A., 1977, ApJ, 214, 679
Baldwin J. A., Burke W. L., Gaskell C. M., Wampler E. J., 1978, Nature, 273, 431
Baldwin J., Ferland G., Korista K., Verner D., 1995, ApJ, 455, L119
Baldwin J. A., Ferland G. J., Korista K. T., Hamann F., LaCluyzé A., 2004, ApJ, 615, 610
Baron D., 2019, preprint (arXiv:1904.07248)
Barth A. J. et al., 2013, ApJ, 769, 128
Baskin A., Laor A., 2018, MNRAS, 474, 1970
Bentz M. C., Katz S., 2015, PASP, 127, 67
Bentz M. C. et al., 2009, ApJ, 705, 199
Blandford R. D., McKee C. F., 1982, ApJ, 255, 419
Blanton M. R. et al., 2017, AJ, 154, 28
Boller T., Freyberg M. J., Truemper J., Haberl F., Voges W., Nandra K., 2016, A&A, 588, 103
Bon N., Bon E., Marziani P., 2018, Frontiers Astron. Space Sci., 5, 3
Boroson T. A., Green R. F., 1992, ApJS, 80, 109
Brotherton M. S., 1996, ApJS, 102, 1
Brotherton M. S., Wills B. J., Francis P. J., Steidel C. C., 1994, ApJ, 430, 495
Brotherton M. S., Singh V., Runnoe J., 2015, MNRAS, 454, 3864
Calderone G., Nicastro L., Ghisellini G., Dotti M., Sbarrato T., Shankar F., Colpi M., 2017, MNRAS, 472, 4051
Chen K., Halpern J. P., 1989, ApJ, 344, 115
Chen K., Halpern J. P., Filippenko A. V., 1989, ApJ, 339, 742
Coffey D. et al., 2019, A&A, 625, A123
Collin S., Joly M., 2000, New Astron. Rev., 44, 531
Comparat J. et al., 2019, preprint (arXiv:1912.03068)
Corbin M. R., 1995, ApJ, 447, 496
Czerny B., Hryniewicz K., 2011, A&A, 525, L8
Czerny B. et al., 2017, ApJ, 846, 154
Dietrich M., Hamann F., Shields J. C., Constantin A., Vestergaard M., Chaffee F., Foltz C. B., Junkkarinen V. T., 2002, ApJ, 581, 912
Du P., Wang J.-M., Hu C., Ho L. C., Li Y.-R., Bai J.-M., 2016, ApJ, 818, L14
Dwelly T. et al., 2017, MNRAS, 469, 1065
Elitzur M., 2012, ApJ, 747, L33
Eracleous M., Halpern J. P., 1994, ApJS, 90, 1
Eracleous M., Halpern J. P., 2003, ApJ, 599, 886
Fanidakis N. et al., 2012, MNRAS, 419, 2797
Ferland G. J. et al., 2013, RMxAA, 49, 137
Fischer N., Mammen E., Marron J., 1994, Comput. Stat. Data Anal., 18, 499
Ganci V., Marziani P., D'Onofrio M., del Olmo A., Bon E., Bon N., Negrete C. A., 2019, A&A, 630, A110
Gaskell C. M., 2009, New Astron. Rev., 53, 140
Gaskell C. M., Harrington P. Z., 2018, MNRAS, 478, 1660
Gebhardt K. et al., 2000, ApJ, 543, L5
Georgakakis A., Nandra K., 2011, MNRAS, 414, 992
Gravity Collaboration, 2018, Nature, 563, 657
Grogin N. A. et al., 2005, ApJ, 627, L97
Grupe D., 2004, Astron. J., 127, 1799
Grupe D., 2011, Proc. Sci., Statistical Analysis of an AGN Sample with Simultaneous UV and X-ray Observations with Swift. SISSA, Trieste, PoS#4
Grupe D., Beuermann K., Mannheim K., Thomas H. C., 1999, A&A, 350, 805
Gunn J. E. et al., 2006, AJ, 131, 2332
Hall P., York M., 2001, Stat. Sin., 11, 515
Haynes W., 2013, in Dubitzky W., Wolkenhauer O., Cho K.-H., Yokota H., eds, Encyclopedia of Systems Biology. Springer, New York, NY, p. 154
Heckman T. M., Miley G. K., van Breugel W. J. M., Butcher H. R., 1981, ApJ, 247, 403
Hickox R. C., Alexander D. M., 2018, ARA&A, 56, 625
Jarvis M. J., McLure R. J., 2006, MNRAS, 369, 182
Jolliffe I. T., Cadima J., 2016, Phil. Trans. R. Soc. London Ser. A, 374, 20150202
Joly M., Véron-Cetty M., Véron P., 2007, in Ho L. C., Wang J.-W., eds, ASP Conf. Ser. Vol. 373, The Central Engine of Active Galactic Nuclei. Astron. Soc. Pac., San Francisco, p. 376
Kaspi S., Maoz D., Netzer H., Peterson B. M., Vestergaard M., Jannuzi B. T., 2005, ApJ, 629, 61
Kelly B. C., 2007, ApJ, 665, 1489
King A., 2016, MNRAS, 456, L109
Kraft R. P., Burrows D. N., Nousek J. A., 1991, ApJ, 374, 344
Kuraszkiewicz J., Wilkes B. J., Schmidt G., Smith P. S., Cutri R., Czerny B., 2009, ApJ, 692, 1180
Laird E. S. et al., 2009, ApJS, 180, 102
Lewis K. T., Eracleous M., Storchi-Bergmann T., 2010, ApJS, 187, 416
McLure R. J., Jarvis M. J., 2002, MNRAS, 337, 109
Mao Y.-F., Wang J., Wei J.-Y., 2009, Res. Astron. Astrophys., 9, 529
Marinello M., Rodríguez-Ardila A., Garcia-Rissmann A., Sigut T. A. A., Pradhan A. K., 2016, ApJ, 820, 116
Marziani P., Sulentic J. W., Zwitter T., Dultzin-Hacyan D., Calvani M., 2001, ApJ, 558, 553
Marziani P., Sulentic J. W., Zamanov R., Calvani M., Dultzin-Hacyan D., Bachev R., Zwitter T., 2003a, ApJS, 145, 199
Marziani P., Zamanov R. K., Sulentic J. W., Calvani M., 2003b, MNRAS, 345, 1133
Marziani P., Sulentic J. W., Stirpe G. M., Zamfir S., Calvani M., 2009, A&A, 495, 83
Marziani P., Sulentic J. W., Negrete C. A., Dultzin D., Zamfir S., Bachev R., 2010, MNRAS, 409, 1033
Marziani P., Sulentic J. W., Plauchu-Frayn I., del Olmo A., 2013a, A&A, 555, A89
Marziani P., Sulentic J. W., Plauchu-Frayn I., del Olmo A., 2013b, ApJ, 764, 150
Marziani P., Negrete C. A., Dultzin D., Martínez-Aldama M. L., Del Olmo A., D'Onofrio M., Stirpe G. M., 2017, Frontiers Astron. Space Sci., 4, 16
Marziani P. et al., 2018, Frontiers Astron. Space Sci., 5, 6
Mejia-Restrepo J., Trakhtenbrot B., Lira P., Netzer H., Capellupo D., 2016, MNRAS, 460, 187
Mejía-Restrepo J. E., Trakhtenbrot B., Lira P., Netzer H., 2018, MNRAS, 478, 1929
Negrete C. A. et al., 2018, A&A, 620, A118
Netzer H., 2008, New Astron. Rev., 52, 257
Netzer H., 2015, ARA&A, 53, 365
Netzer H., Trakhtenbrot B., 2007, ApJ, 654, 754
Oke J. B., 1987, in Zensus J. A., Pearson T. J., eds, Superluminal Radio Sources. Cambridge University Press, Cambridge and New-York, p. 267
Padovani P. et al., 2017, A&AR, 25, 2
Panda S., Czerny B., Wildy C., 2017, Frontiers Astron. Space Sci., 4, 33
Panda S., Czerny B., Adhikari T. P., Hryniewicz K., Wildy C., Kuraszkiewicz J., Śniegowska M., 2018, ApJ, 866, 115
Panda S., Marziani P., Czerny B., 2019a, preprint (arXiv:1912.03118)
Panda S., Marziani P., Czerny B., 2019b, preprint (arXiv:1912.03119)
Panda S., Marziani P., Czerny B., 2019c, ApJ, 882, 79
Pedregosa F. et al., 2011, J. Mach. Learn. Res., 12, 2825
Perrotta S., Hamann F., Zakamska N. L., Alexand roff R. M., Rupke D., Wylezalek D., 2019, MNRAS, 488, 4126
Peterson B. M., 1993, PASP, 105, 247
Peterson B., 2006, in Alloin D., Johnson R., Lira P., eds, Physics of Active Galactic Nuclei at all Scales. Springer, Berlin, p. 77
Peterson B. M., Ferland G. J., 1986, Nature, 324, 345
Popovic L. Č., Mediavilla E. G., Kubičela A., Jovanovic P., 2002, A&A, 390, 473
Popović L. Č., Mediavilla E., Bon E., Ilić D., 2004, A&A, 423, 909
Rakshit S., Woo J.-H., 2018, ApJ, 865, 5
Rees M. J., Netzer H., Ferland G. J., 1989, ApJ, 347, 640
Richards G. T., Vanden Berk D. E., Reichard 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., 2006, ApJS, 166, 470
Rodríguez M., 1999, A&A, 348, 222
Rodríguez-Ardila A., Viegas S. M., Pastoriza M. G., Prato L., 2002, ApJ, 565, 140
Salvato M. et al., 2018, MNRAS, 473, 4937
Saxton R. D., Read A. M., Esquej P., Freyberg M. J., Altieri B., Bermejo D., 2008, A&A, 480, 611
Shang Z., Wills B. J., Robinson E. L., Wills D., Laor A., Xie B., Yuan J., 2003, ApJ, 586, 52
Shen Y., 2013, Bull. Astron. Soc. India, 41, 61
Shen Y., Ho L. C., 2014, Nature, 513, 210
Shen Y., Liu X., 2012, ApJ, 753, 125
Shen Y. et al., 2015, ApJ, 805, 96
Sigut T. A. A., Pradhan A. K., 1998, ApJ, 499, L139
Sigut T. A. A., Pradhan A. K., 2003, ApJS, 145, 15
Silverman B. W., 1981, J. R. Stat. Soc. Ser. B (Methodological), 43, 97
Silverman B. W., 1986, Density Estimation for Statistics and Data Analysis. Chapman & Hall, London
Singh K. P., 2013, Bull. Astron. Soc. India, 41, 137
Smee S. A. et al., 2013, AJ, 146, 32
Storchi-Bergmann T., Schimoia J., Peterson B., Elvis M., Denney K., Eracleous M., Nemmen R., 2016, ApJ, 835, 236
Sulentic J. W., Marziani P., 2015, Frontiers Astron. Space Sci., 2, 6
Sulentic J. W., Marziani P., Dultzin-Hacyan D., 2000a, ARA&A, 38, 521
Sulentic J. W., Zwitter T., Marziani P., Dultzin-Hacyan D., 2000b, ApJ, 536, L5
Sulentic J. W., Marziani P., Zamanov R., Bachev R., Calvani M., Dultzin-Hacyan D., 2002, ApJ, 566, L71
Sulentic J. W., Dultzin-Hacyan D., Marziani P., 2007a, in Kurtz S., ed. Revista Mexicana de Astronomia Conference Series, 28. Rev. Mex. Astron. Astrofis., Mexico City, p. 83
Sulentic J. W., Bachev R., Marziani P., Negrete C. A., Dultzin D., 2007b, ApJ, 666, 757
Sulentic J., Marziani P., Zamfir S., 2011, Balt. Astron., 20, 427
Sulentic J. W. et al., 2017, A&A, 608, A122
Sun J., Shen Y., 2015, ApJ, 804, L15
Tang B., Shang Z., Gu Q., Brotherton M. S., Runnoe J. C., 2012, ApJS, 201, 38
Trakhtenbrot B., Netzer H., 2012, MNRAS, 427, 3081
Urry C. M., Padovani P., 1995, PASP, 107, 803
Vestergaard M., 2002, ApJ, 571, 733
Vestergaard M., Peterson B. M., 2006, ApJ, 641, 689
Voges W. et al., 1999, A&A, 349, 389
Wanders I., Peterson B. M., 1996, ApJ, 466, 174
Wang J., Xu D. W., Wei J. Y., 2018, ApJ, 852, 26
Wills B. J., Netzer H., Wills D., 1985, ApJ, 288, 94
Wills B. J., Browne I. W. A., 1986, ApJ, 302, 56
Winkler H., Chauke T., 2014, preprint (arXiv:1412.4498)
Yuan M. J., Wills B. J., 2003, ApJ, 593, L11
Zakamska N. L. et al., 2016, MNRAS, 459, 3144
Zamanov R., Marziani P., Sulentic J. W., Calvani M., Dultzin-Hacyan D., Bachev R., 2002, ApJ, 576, L9
Zamfir S., Sulentic J. W., Marziani P., Dultzin D., 2010, MNRAS, 403, 1759

APPENDIX A: EQUIVALENT WIDTH ESTIMATE

The equivalent width offers a measure of an emission line's strength. Equivalent widths were not included in C19 and had to be derived.

The equivalent width of a spectral line is defined as

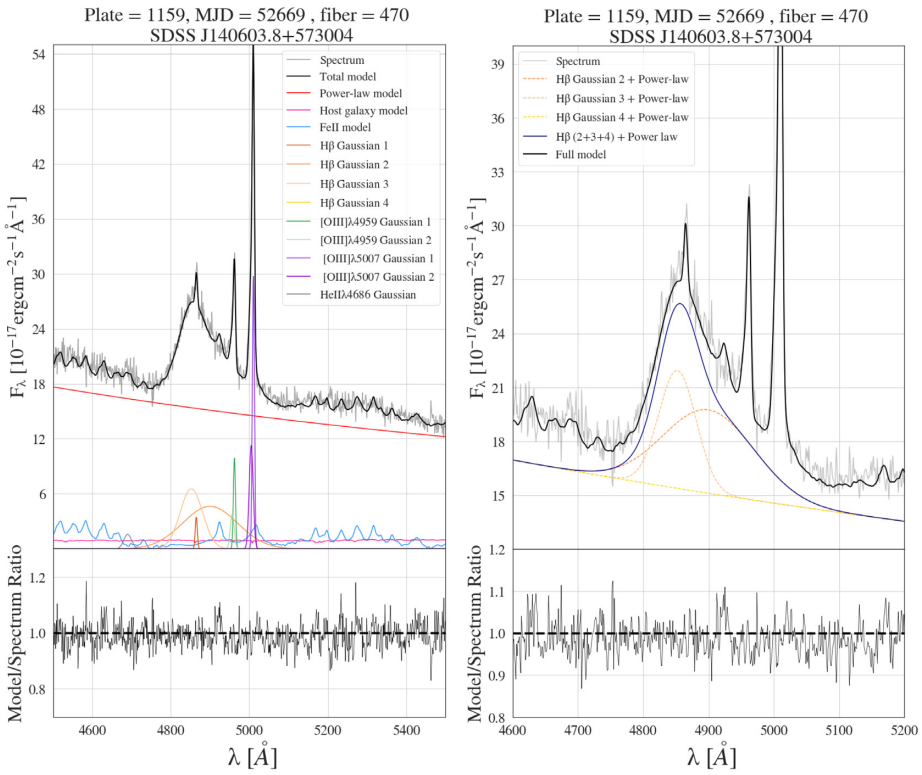

Figure A1. *Left-hand panel*: Reconstructed H β fitting region using the model used by C19. The [O III], He II, and H β model components are labelled. The components of the continuum model (a power law, a host galaxy model, and the iron pseudo-continuum) are also displayed. The original observed spectrum is shown in light grey. *Right-hand panel:* A zoom on to the H β Gaussian line decomposition. In order to improve the visualization, the power-law contribution was added to the Gaussian components. Only two broad components were required to fit H β.

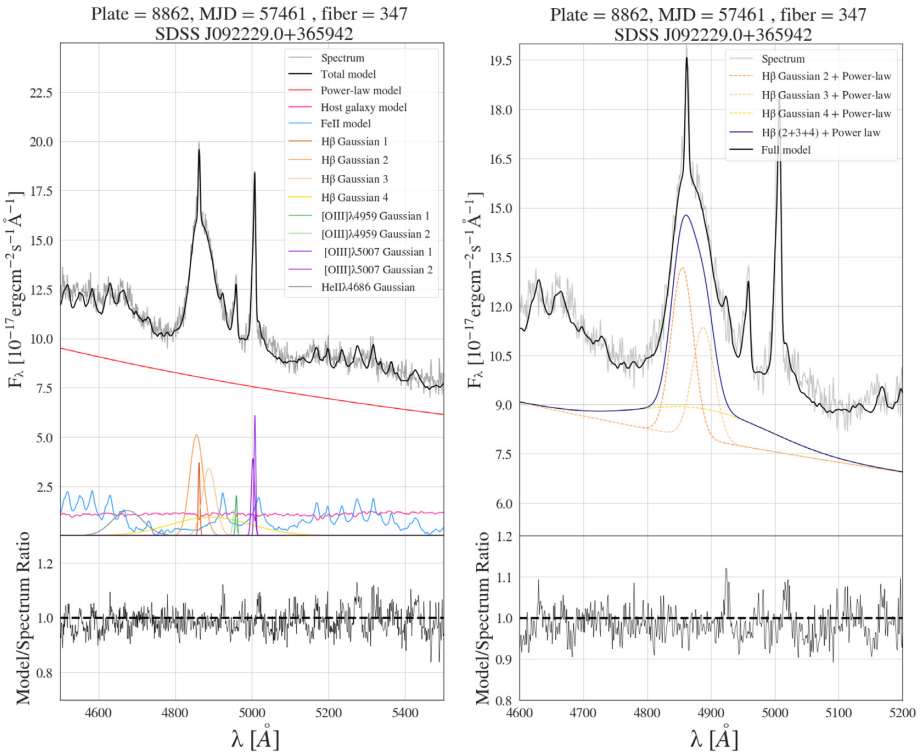

Figure A2. Same as in Fig. A1 but for another source in our sample. Here, three broad Gaussian components were needed to fit H β.

$$W_\lambda = \int \frac{S_1(\lambda) - S_c(\lambda)}{S_c(\lambda)} d\lambda, \tag{A1}$$

where S_1 denotes the total flux density at a given wavelength and S_c the continuum flux density.

To obtain the interpolated S_1 and S_c fluxes, the lines were reconstructed from the catalogue's fitting parameters, the associated Gaussian fitting functions, the continuum power law, and the templates for the iron and galaxy contributions. Each Gaussian component $G(\lambda)$ was reconstructed according to

$$G_{\rm line}(\lambda) = N_{\rm line}\ \exp\left[-\frac{1}{2}\left(\frac{\lambda - P_{\rm line}}{W_{\rm line}}\right)^2\right]. \tag{A2}$$

Here, $N_{\rm line}$, $P_{\rm line}$, and $W_{\rm line}$ denote the norm, peak wavelength, and width for the fit as listed in C19, respectively. For multiple Gaussian components $G_j(\lambda)$, the total flux in the line is then obtained from $G_{\rm total}(\lambda) = \sum_j G_j(\lambda)$. In C19, the continuum model consists of a power law, a host galaxy component, and the iron emission.

The power-law model is reconstructed from

$$PL(\lambda) = N_{\rm PL}\,\lambda^{S_{\rm PL}} \tag{A3}$$

$N_{\rm PL}$ and $S_{\rm PL}$ are the norm and the slope of the power law for each spectrum as listed in the catalogue, respectively. Following the catalogue's author's prescription, a Gaussian filter was applied to the Fe II emission to reproduce the blending of the multiplets. The continuum emission is calculated by

$$S_c(\lambda) = PL(\lambda) + FEII(\lambda) + GAL(\lambda). \tag{A4}$$

The total flux in a given line is: $S_1(\lambda) = S_c(\lambda) + G_{\rm total}(\lambda)$.

Figs A1–A2 display two fits centred around the H β/[O III] complex reconstructed using the above model. Interpolating D_1 and S_c over the fitting range of H β or Mg II and performing the integration in (A1) yield the equivalent width for the chosen lines.

APPENDIX B: EXAMPLES OF Hβ PROFILES: RED- AND BLUE-ASYMMETRIC AND VERY BROAD COMPONENTS

In the top (middle) three panels of the figure in Appendix B, we present examples of red(blue)-asymmetric broad H β profiles and the Gaussian fits performed in C19. The corresponding asymmetry indices are indicated in the caption. Similarly, in the bottom three panels of this same figure, the H β lines with very broad Gaussian components are presented. The VBC is highlighted in red.

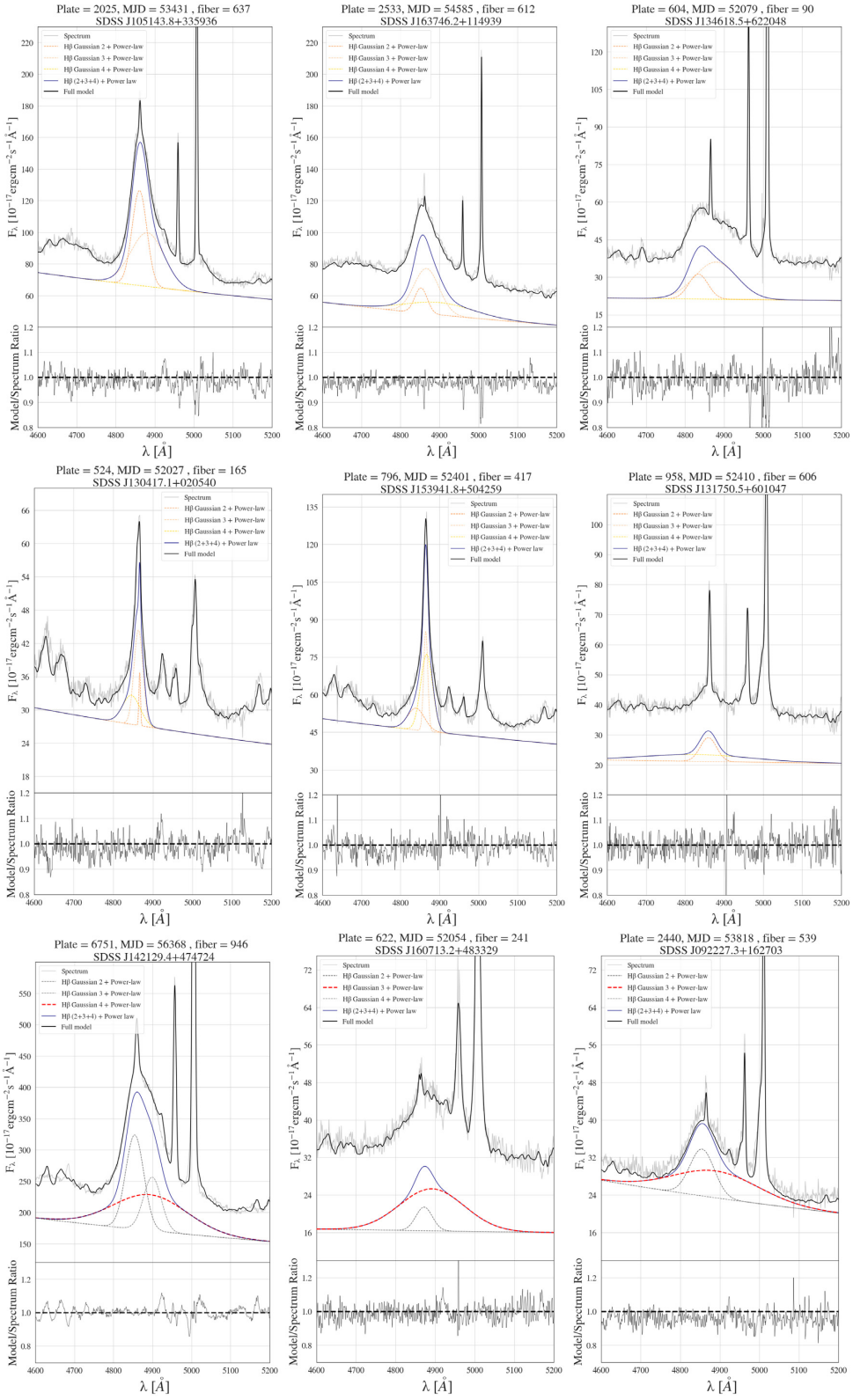

Figure B1. *Top:* The Gaussian line decomposition and spectra of the broad H β from three sources with red-asymmetric H β profiles. From left to right, the asymmetry indices $\Delta\lambda_{\mathrm{H}\,\beta}$ are: +0.22, +0.22, and + 0.27. *Middle:* The Gaussian line decomposition and spectra of the broad H β from three sources with blue-asymmetric H β profiles. From left to right, the asymmetry indices $\Delta\lambda_{\mathrm{H}\,\beta}$ are: −0.23, −0.20, and −0.24. *Bottom:* The Gaussian line decomposition and spectra of the broad H β from three sources that have one very broad Gaussian component (FWHM $>$ 10 000 km s^{-1}. The very broad component is marked in red.

APPENDIX C: TESTING A SIMPLE OBSCURATION SCENARIO

Richards et al. (2002) proposed that the spatial configuration resulting from the BLR's inclination and obscuration might explain the asymmetries in the observed profile of C IV λ1549 Å (see also 4.3 in Z10). In this picture, the asymmetry arises from the attenuation of redshifted photons by obscuring material. Similarly, Gaskell & Harrington (2018) developed a model of outflowing clouds of dust, blocking the line of sight to the inner disc-shaped BLR, which predicts broad emission profiles similar to those observed. An alternative view of the effect of the BLR and outflows is developed by Czerny et al. (2017) (see also Czerny & Hryniewicz 2011). In this model, the dusty clouds are driven out of an optically thick disc, and are eventually exposed to radiation by the central source. The dust evaporates, removing the matter/radiation interaction and results in the fallback of gas clouds. This scenario is referred to as Failed Radiatively Accelerated Dusty Outflow (FRADO). The physics of dust sublimation and the consequences for the BLR disc structure were further investigated by Baskin & Laor (2018).

Fig. 14 allows us to investigate a simple obscuration scenario in which the observed asymmetry of the line shapes of H β would result from the partial obscuration of a flattened BLR in Keplerian motion around the black hole. If the receding or incoming region (with respect to our line of sight) of a disc-like structure is obscured by optically thick material, the emitted broad lines would be, respectively, red or blue asymmetric (see, for example, fig. 2 in Gaskell & Harrington 2018), since one peak of the double-peaked emission profile would be attenuated. In this scenario, the low-intensity, 'excess' emission measured on blue (or red) side of the line would correspond to the attenuated blue (or red) horn of the disc-like emission.

Furthermore, if we assume no further kinematics in our BLR model, the attenuation of the red of blue peak of the disc-like emission would lead to a centroid shift configuration:

(i) Blueward asymmetric H β ($\Delta\lambda_{\mathrm{H}\beta} < 0$): Redshift of the peak component ($c_{80} > 0$) and blueshift of the base component ($c_{15} < 0$);

(ii) Redward asymmetric H β ($\Delta\lambda_{\mathrm{H}\beta} > 0$): Blueshift of the peak component ($c_{80} < 0$) and redshift of the base component ($c_{15} > 0$).

Fig. 14 supports this scenario. The results in the previous section suggest that the broad Balmer shapes might be affected by the presence of a distinct emitting region: the VBLR. The next step is to test if the obscuration scenario is compatible with the presence of a very broad component related to the stratification of the BLR. To this effect, we identify sources for which the broad H β has been fitted with broad Gaussians.

Fig. C1 presents the peak wavelength of all the Gaussians that have been used to fit the components of H β as a function of their FWHM. We clearly identify a sequence of very broad Gaussians at velocities $>10\,000\,\mathrm{km\,s^{-1}}$, which are mostly redshifted with respect to systemic redshift. We investigated the reduced χ^2 distribution of the fits of the sources that have at least one Gaussian component of FWHM $>10\,000\,\mathrm{km\,s^{-1}}$. The sources that have a very broad component are not biased towards higher reduced χ^2 (and thus higher BIC) values, indicating that they are not preferentially found in less secure fits. The statistical significance of a second mode in the distribution of the Gaussian FWHM in our sample at FWHM$>10\,000\,\mathrm{km\,s^{-1}}$ was assessed with a Silverman test of multimodality (Silverman 1981, for its calibration see e.g. Hall & York 2001; Ameijeiras-Alonso, Crujeiras & Rodríguez-Casal 2019). We could confirm that this second mode of very broad Gaussian components is indeed significant. The test is presented in detail in Appendix D. We empirically identify this sequence as the subclass of sources that contain a VBLR. As expected, the broadest Gaussians have the largest contributions to the full line luminosity. Fig. 14 includes the density contours containing the sources for which H β has not been fitted with a very broad Gaussian (FWHM $< 10\,000\,\mathrm{km\,s^{-1}}$, i.e. the sources to the left of the vertical dashed line in Fig. C1). The trend observed for all objects in our sample is preserved for this subset, supporting the possibility of a partially obscurated flattened BLR + VBLR model. The red- and blueshifts of the H β base and peak components are symmetrically distributed, which is consistent with the obscuration scenario.

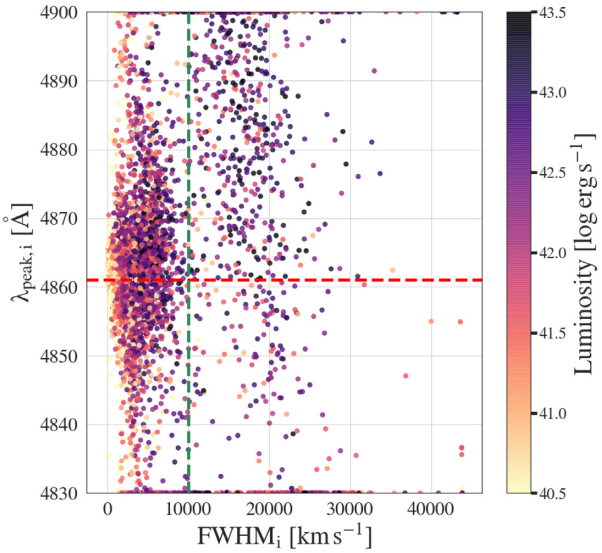

Figure C1. The peak wavelength of the Gaussians used to fit the H β components as a function of their FWHM. The red-dotted line indicates the systemic wavelength of H β. The green dotted line shows an empirical separation of broad and very broad Gaussian components at $10\,000\,\mathrm{km\,s^{-1}}$. The right side of this separation clearly reveals a sequence of very broad, mainly redshifted Gaussians. The points are colour coded according to the luminosity from each Gaussian. Each source in our sample has four points in this figure (one for each of its Gaussian).

In a simple model in which wing attenuation via partial obscuration is the sole source of asymmetry of the broad emission lines, we expect the H β line fluxes to be symmetrically distributed for $\Delta\lambda_{\mathrm{H}\beta} < 0$ and $\Delta\lambda_{\mathrm{H}\beta} > 0$, as no wing of the emission-line profile should be preferentially attenuated. The left-hand panel in Fig. C2 shows the equivalent width of H β plotted against its asymmetry index for our full sample. We observe an offset in $W(\mathrm{H}\,\beta)$ between the population of red- and blue-asymmetric emitters. We argue that this offset might once again be the signature of the presence of a VBLR in some of our sources. If we display the kernel density contours for the subset of sources that were not fitted with a very broad Gaussian, a more symmetric distribution of equivalent widths is produced (i.e. line fluxes), as shown in the right-hand panel of Fig. C2.

These simple tests have demonstrated that our sample is consistent with a model of partial obscuration coupled to a stratified BLR (one that might or might not contain a VBLR).

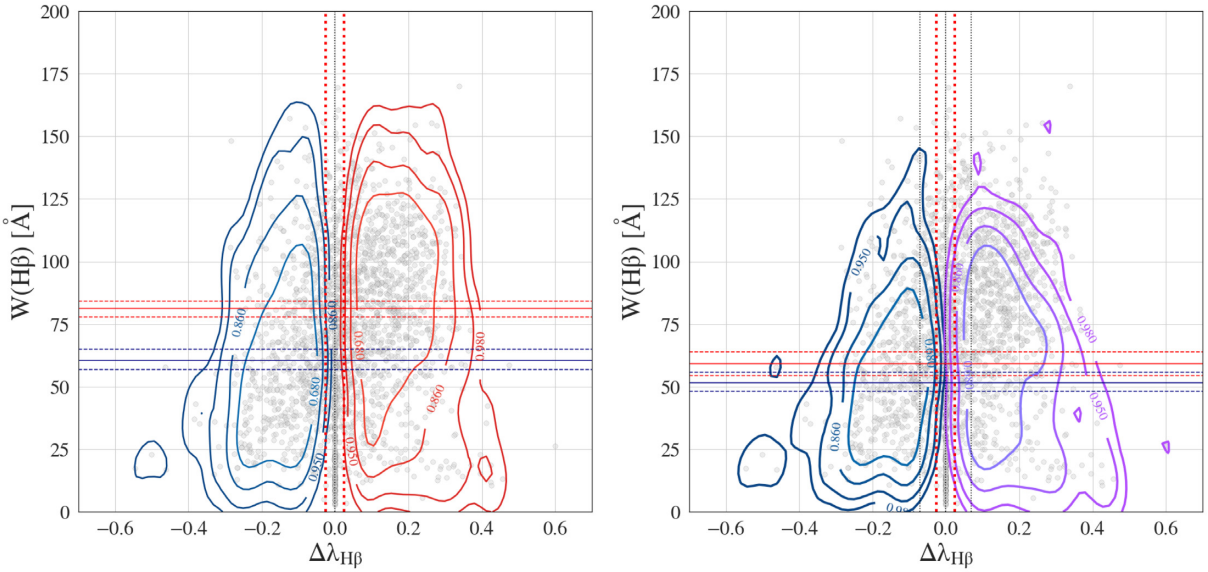

Figure C2. *Left-hand panel*: The equivalent width of the broad H β emission as a function of its asymmetry Index. The blue and red density contours are included to improve the visualization of the red- and blue-asymmetric subsamples. An offset between the two populations is observed. The horizontal red (blue) line marks the mean of the bootstrapped mean $W(\mathrm{H}\,\beta)$ of the red(blue)-ward asymmetric sources. The red and blue, horizontal dashed lines correspond to the 3σ confidence contours, as derived from the percentile method at 0.13 and 99.87 per cent of the sampling distribution of the mean. *Right-hand panel*: Same format as on the left. The blue and purple contours, respectively, represent the sources with blue- and red-asymmetric H β, for which the Gaussian models had FWHM $< 10\,000\,\mathrm{km\,s^{-1}}$.

APPENDIX D: VBC MODE SIGNIFICANCE WITH A SILVERMAN TEST OF MULTIMODALITY

The FWHMs of broad Gaussian components ($>800\,\mathrm{km\,s^{-1}}$) that were fitted to the H β emission lines show a bimodal distribution. The two modes appear to distinguish broad components ($800\,\mathrm{km\,s^{-1}} <$ FWHM $< 10\,000\,\mathrm{km\,s^{-1}}$) and very broad components (FWHM $> 10\,000\,\mathrm{km\,s^{-1}}$). Out of the 8400 Gaussian components used to fit the 2100 sources of the sample, 4936 (58 per cent) can be considered broad, with FWHM $> 800\,\mathrm{km\,s^{-1}}$, and 1080 (13 per cent) are very broad components with FWHM $> 10\,000\,\mathrm{km\,s^{-1}}$. The significance of the second mode was assessed using the Silverman test of multimodality (Silverman 1981). We present the outline of this test here.

(i) We extract all the Gaussians used in the fits of our sample with the condition: FWHM $> 800\,\mathrm{km\,s^{-1}}$ and $\lambda_{\mathrm{peak}} \neq 4830$ Å and $\lambda_{\mathrm{peak}} \neq 4900$ Å. This cut removes narrow Gaussians, as well as Gaussians which had their . This cut removes narrow Gaussians, as well as Gaussian Gaussianswhich 4275 Gaussian components remain. We show the resulting density histogram of FWHM in Fig. D1.

(ii) For a range of bandwidths $h = [0.01, 2.00]$ (step size $s = 100$), we performed a Gaussian KDE of the distribution of FWHM (centred at zero and scaled to unit variance). An example probability density function (pdf) from KDE is displayed in Fig. D1 (red line). For n observations in our sample, the bandwidth h smooths the Gaussian kernel estimate as follows:

$$\hat{f}(x,h) = \frac{1}{nh}\sum_{i=1}^{n}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-x_i)^2}{2h^2}}. \quad \text{(D1)}$$

(iii) We then detect the critical bandwidths $h_{c,j}^{\mathrm{DATA}}$, at which new local maxima appear in the pdf, when h is decreased. For $1 < j < n$ (where n corresponds to the number of observations), the critical bandwidths $h_{c,j}^{\mathrm{DATA}}$ are formally defined as the minimum h, at which the Gaussian density estimate has no more than j maxima. In Fig. D2, the bandwidth h and the corresponding number of detected local maxima j are displayed. Critical widths are symbolized by vertical lines.

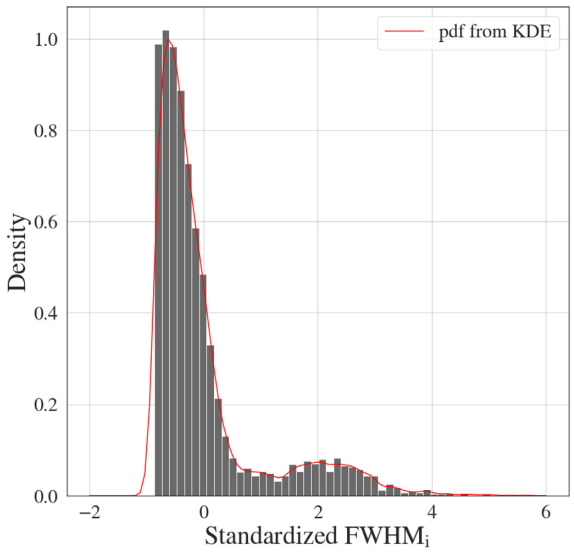

Figure D1. Density of standardized Gaussian component FWHM.

(iv) In order to assess the significance of a mode detection, we follow Fischer, Mammen & Marron (1994) and set up a null-hypothesis test, which is based on the observation that the number of modes of $\hat{f}(x)$ decreases with increasing bandwidth h. If k is the true number of modes:

Null hypothesis H_0: $k = j$
Alternative, H_1: $k \geq j + 1$ for $j = 1, 2...$

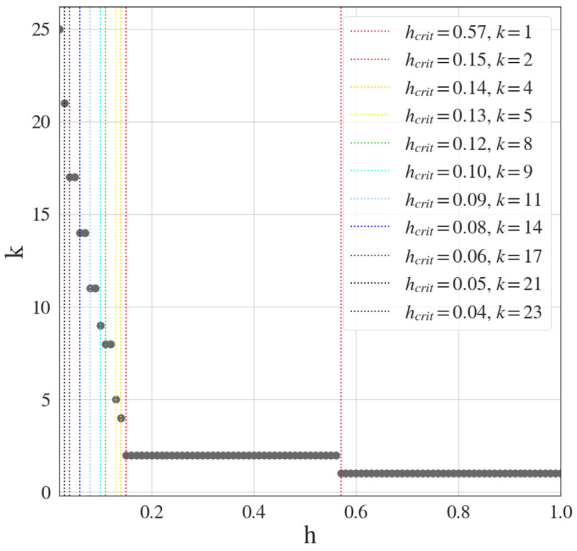

Figure D2. The number of local maxima k of the estimated pdf as a function of the smoothing h, the bandwidth of the Gaussian KDE. The critical values $h_{c,j}^{\rm DATA}$ are marked as vertical lines.

Table D1. p-values for mode detections for decreasing critical bandwidths.

$h_{\rm j}^{\rm DATA}$	Number of local maxima	p-value
0.57	1	0.100
0.15	2	0.001
0.14	3	0.008
0.13	5	0.041
0.12	8	0.095
...	...	...

For each $h_{c,j}^{\rm DATA}$ detected in our data, we perform a Gaussian KDE on 100 smoothed bootstrapped samples, using $h_{c,j}^{\rm DATA}$ as bandwidth. These resamples are obtained, by sampling smoothed FWHM$_{\rm s}$, using bootstrapping (with replacement). The exact computation of FWHM$_{\rm s}$ is as follows:

$$\mathrm{FWHM_s} = \frac{1}{\sqrt{1 + {h_{c,j}}^2/\sigma^2}} \left(\mathrm{FWHM_r} + h_{c,j}\epsilon\right) \tag{D2}$$

where FWHM$_{\rm r}$ are resampled FWHM, σ is the standard deviation of the bootstrapped sample, and ϵ is a random standard variable.

For the null-hypothesis test, Silverman (1981) motivates a p-value defined for a given number of maxima j as

$$p = P\left\{h_{c,j}^{\rm NEW} > h_{c,j}^{\rm DATA}\right\} = P\{\hat{f}\left(x, h_{c,j}^{\rm DATA}\right) \text{ has } > j \text{ modes}\}, \tag{D3}$$

where $h_{c,j}^{\rm NEW}$ is a the critical value for j maxima derived from the resamples, $h_{c,j}^{\rm DATA}$ is the initial critical width from the data, and $\hat{f}(x, h_{c,j}^{\rm DATA})$ is the KDE of the resamples using $h_{c,j}^{\rm DATA}$ as critical density. The null hypothesis is rejected if $P\{h_{c,j}^{\rm NEW} > h_{c,j}^{\rm DATA}\} < \alpha$. For each $h_{c,j}^{\rm DATA}$, and thus each mode detection, we compute a p-value. We assess how many times the density estimates with fixed bandwidths of the 100 resamples have at most j detected modes. The ratio of these realizations and the total number of simulations is taken as p-value. The resulting p-values are listed in Table D1.

(v) Following Silverman (1981), we read this table as hierarchical successive significance test. As soon as, for a new j, the p-values drop below an α significance threshold, we interpret the detection of j modes as significant. From Table D1, if we adopt a significance threshold $\alpha = 0.05$, it appears that the bimodality of the distribution of our Gaussian components FWHM is significant, i.e. the very broad mode of Gaussian components is significant.

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