RAA 2019 Vol. **19** No. **12**, 169(10pp) doi: 10.1088/1674–4527/19/12/169

http://www.raa-journal.org *http://iopscience.iop.org/raa*

Research in Astronomy and Astrophysics

Spectral principal component analysis of the Hβ region of low-redshift SDSS quasars

Bin Ma[1,2], Zhaohui Shang[3,1] and Michael S. Brotherton[4]

[1] National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China; *zshang@gmail.com*
[2] School of Astronomy and Space Science, University of Chinese Academy of Sciences, Beijing 100049, China
[3] Tianjin Astrophysics Center, Tianjin Normal University, Tianjin 300387, China
[4] Department of Physics and Astronomy, University of Wyoming, Laramie, WY 82071, USA

Received 2019 April 22; accepted 2019 May 24

Abstract Studying the relationships among quasar spectral features is essential to unveil the origins of the emission lines and the quasars' physical processes. Principal component analysis (PCA) is a powerful tool to investigate correlations between variables. Here, we present the results of PCA on the spectra of low-redshift SDSS quasars. The rest-frame wavelength range studied is 4000 – 5500 Å, involving some typical features of quasar spectra, such as Hβ, [O III] and Fe II emission lines. The first principal component is the anti-correlation between [O III] and Fe II, the well-known eigenvector one (EV1). The next six principal components also show clear (anti-)correlations between line strengths and/or velocity widths of various features, which agree well with measured spectral properties. By comparing the weights of these principal components with other quasar properties, we can identify their underlying drivers. We find that the second principal component represents spectral slope, and can quantify quasar host fraction, intrinsic slope and reddening well. The third component exhibits the velocity width variation of Hβ, and may be a proxy for orientation. In addition, we calculate the fractional-contribution spectra to investigate which components dominate the variance at individual wavelength ranges. Our results also indicate that the optical Fe II emission may have distinct origins.

Key words: quasars: emission lines — quasars: general

1 INTRODUCTION

Quasar spectra consist of continua, broad lines, narrow lines and absorption lines, whose properties and relationships reflect fundamental physics involving luminosity, black hole mass, Eddington ratio ($L_{\rm Bol}/L_{\rm Edd}$), orientation and intrinsic dust reddening. Specialists in this field have been compiling large samples and employing various means to study the relationships between emission lines to provide quantitative limits for theory, determining the underlying physical mechanisms and finally understanding the physical nature of quasars.

Boroson & Green (1992) performed a principal component analysis (PCA) on the measurements of spectral properties in the Hβ region of 87 low redshift quasars. They found that the first principal component (also well known as Eigenvector 1 or EV1) was the anti-correlation between Fe II optical emission and the narrow lines. They suspected that this relationship was related to the Eddington ratio. They also first defined the 2D plane by FWHM(Hβ) and $\rm Re_{Fe}$ which is the relative strength of Fe II to Hβ. This plane has been widely used since then, and the distribution of quasars in the plane is sometimes called the quasar main sequence. Many properties in other bands are also correlated with EV1 (e.g., Wang et al. 1996; Laor et al. 1997; Sulentic et al. 2000; Bensch et al. 2015), but the physical driver behind the main sequence remains uncertain. The Eddington ratio as the driver was firstly proposed by Boroson & Green (1992), and has been the frequently favored driver (e.g., Sulentic et al. 2000; Marziani et al. 2001). Furthermore, Shen & Ho (2014) not only inferred the Eddington ratio as the main driver of $\rm Re_{Fe}$ or EV1, but also inferred that orientation contributes to the dispersion of FWHM(Hβ) in the main sequence. This concept is supported by results from various studies (e.g. Sun & Shen 2015; Sun et al. 2018). However, the independent measurement of orientation poses another difficulty. For radio-loud quasars, the radio core dominance is often utilized as a proxy (Wills & Browne 1986), but for radio-quiet quasars, which are the majority class, there is as yet no good indicator for orientation.

Despite many studies supporting the Eddington ratio as the governing physical parameter, it is challenged by other studies. Śniegowska et al. (2018) improved quasar spectral fitting by taking into account several Fe II templates, Balmer continuum and starlight contamination. However, they found that for the quasars in Shen & Ho (2014) with the largest $\mathrm{Re}_{\mathrm{Fe}}$ and small measurement error, the resulting $\mathrm{Re}_{\mathrm{Fe}}$ was significantly lower, and the FWHM(Hβ) differed substantially. Thus they argued that these quasars did not seem to form a uniform group, differing considerably in Eddington ratio. Besides, Panda et al. (2018) physically modeled the distribution of quasars in the 2D plane but concluded that no single simple driver could dominate the main sequence, neither Eddington ratio nor broadband spectrum shape. Also, the orientation is only partially responsible for the dispersion of the main sequence, most of which, however, comes from variations in black hole masses and accretion rates.

EV1 was first derived from PCA on measured spectral parameters, which are fitting dependent. The measurements may differ remarkably as mentioned above. To avoid this problem, Francis et al. (1992) applied the PCA method directly to the spectra of quasars rather than the measured properties, i.e., the input variables were the fluxes at different wavelengths. This method is called spectral principal component analysis (SPCA). They performed SPCA on the spectra in the wavelength range of 1150 – 2000 Å, and found that the first three spectral principal components (SPCs) described 75% of the variance in the sample. The resulting SPCs represent the correlations between the emission-line core strength, continuum slope and broad absorption line characteristics. Moreover, SPCA can be applied to quantitatively classify quasars. In recent years, there have been many studies utilizing SPCA to analyze quasars from X-ray to mid-infrared wavebands (e.g., Shang et al. 2003; Yip et al. 2004; Bian et al. 2016; Pennell et al. 2017; Rochais et al. 2017; Gallant et al. 2018). However, the variation in samples and the wavelength ranges that are used changes the order of the principal components, or even the results of the components. Hu et al. (2012) summarized recent SPCA studies by comparing the sample selection, wavelength range and interpretations of principal components. They also introduced the fractional-contribution spectrum, which is the proportion of variation in a wavelength bin accounted for by an SPC. This can clearly signify which SPCs dominate certain spectral features, and effectively helps to distinguish between artificial "cross talk" and real correlations. In addition to studying spectral properties, SPCA is utilized to decompose the spectra as well. Based on the eigenspectra of quasars and galaxies, the observed spectra can be decomposed into pure-host and pure-quasar spectra (e.g., Vanden Berk et al. 2006; Sun & Shen 2015).

Here, we perform SPCA on low-redshift Sloan Digital Sky Survey (SDSS) quasars. This paper is structured as follows: we describe our sample and methods in Section 2 and Section 3, respectively. We then present the SPCA results in Section 4, and we compare our principal components with quasar properties in Section 5. Finally, we summarize this study in Section 6.

2 SAMPLE

The quasars were selected from the Seventh Data Release of the SDSS Quasar Catalog (Schneider et al. 2010), for which Shen et al. (2011) measured detailed properties of the spectra. We use the improved redshifts by Hewett & Wild (2010). The sample was constructed based on the following criteria:

1. The redshift is smaller than 0.65, because we wanted to investigate the rest-frame wavelength range between 4000 and 5500 Å, involving the emission lines such as Hβ, [O III] and optical Fe II. The wavelength coverage of SDSS spectra is 3800 – 9200 Å,
2. The signal-to-noise ratio in the Hβ region is larger than 10.

These criteria resulted in $10\,614$ quasars for our sample. We retrieved the Galactic-dereddened SDSS spectra of these quasars from the SDSS website[1]. We then resampled the spectra to rest-frame with even $\log\lambda$ intervals of 0.0003, corresponding to $210\,\mathrm{km\,s^{-1}}$ in velocity space. We normalized these spectra by their mean flux between 4000 and 5500 Å, thus the integrals of emission lines represent their equivalent widths (EWs). We calculated the mean spectrum and variance of the whole sample. Because a small fraction of outliers that deviate the most from the mean spectrum contribute significantly to the variance, we rejected the sources with variance, compared to the mean-subtracted spectrum, ten times larger than the median. Consequently, the results will be more reliable for general quasars and will be less biased by outliers. The final sample consists of $10\,241$ spectra, and its mean and standard-deviation spectra are displayed in Figure 1. The standard-deviation spectrum contains both the intrinsic diversity of quasars and noise from observations and instruments. The intrinsic variance is derived by the method in Shang et al. (2003), and it accounts for 88.7% of the total variance.

[1] `http://das.sdss.org/va/qso_properties_dr7/data/dered_spectra/`

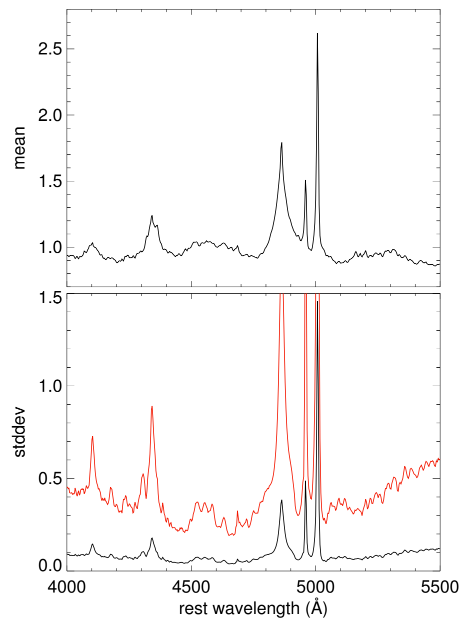

Fig. 1 The mean (*upper*) and intrinsic standard-deviation (*lower*) spectra of our low-z quasar sample. Each spectrum in our sample was normalized by its mean flux between 4000 and 5500 Å. In the lower panel, the spectrum is also enlarged by five times to present the lines clearly and plotted in *red*.

3 METHOD: SPCA

PCA is a dimensionality reduction technique. Assuming there are N quantities for a sample, some of which are affected by the same physical mechanism and thus are related, PCA linearly combines N-quantities to obtain N-new orthogonal quantities. The first new quantity describes the most sample variance, the second quantity describes the next most variance and so on. Only a few quantities are able to describe the vast majority of the sample variance, so the new quantities are called the principal components. In SPCA, the input quantities are the fluxes in wavelength bins, and the output components are called SPCs.

We followed Francis et al. (1992) to perform the SPCA. There are 462 wavelength bins in each rebinned spectrum. For each quasar, the spectrum was normalized by the mean flux between 4000 and 5500 Å, and then the mean spectrum of the sample was subtracted. The SPCA was performed on the 462 fluxes in different wavelength bins to obtain 462 new variables, denoted as SPCi. Each SPC is a linear combination of fluxes, where the coefficients represent the correlation between fluxes at different wavelengths. The fluxes in different wavelengths are positively correlated if their coefficients have the same sign, or anti-correlated if their coefficients have opposite signs. These coefficients are arranged in wavelength order to form an eigenspectrum for each SPC. The order of SPCs is sorted by decreasing eigenvalues, which are their variances, i.e., SPC1 has the largest variance. Usually, a few SPCs contribute the majority of the total variance of the sample, and by linearly combining them, the spectrum of each quasar can be reconstructed very well.

Although SPC1 contributes the most sample variance, it is not always the largest contributor to variance at every wavelength bin. Given that the variance of the samples at some wavelengths is significantly larger (e.g., the variance of [O III] accounts for more than half of the sample variance), the first principal component consequently reflects the correlation between these lines. For other wavelengths where the fluxes vary less, the largest contribution is not necessarily in the first principal component. In addition, some coefficients in the eigenspectrum are introduced due to the effect of cross talk, rather than the true correlation. Therefore, we also adopt the fractional-contribution spectrum proposed by Hu et al. (2012) to clarify which spectral feature is dominated by which eigenspectra.

4 RESULTS

We have applied SPCA on 10 241 quasars in the wavelength range of 4000 – 5500 Å. The first seven SPCs account for 53.3, 22.1, 7.7, 3.9, 2.9, 1.4 and 1.0 percent of the total intrinsic variance in the sample, respectively. Their eigenspectra and fractional-contribution spectra are shown in Figure 2. The wavelengths of the H Balmer lines, Fe II and [O III] are also labeled. In eigenspectra, features with the same sign are correlated while those with opposite signs are anti-correlated. The first three components are very similar to the second, the third, and the fourth components in Yip et al. (2004) (their SPC1 is our mean spectrum). To better understand the meaning of these SPCs, we made composite spectra to see how quasar spectra vary with each SPC. We sorted the quasars by their weights in each SPC, and selected five sets of spectra at the minimum (0%) of the weights, three quartiles (25%, 50% and 75%), and the maximum (100%), respectively. Each composite spectrum was built with 100 spectra, and is plotted in Figure 3 and Figure 4. In addition, we calculated the Spearman rank correlation coefficients (ρ) and associated probabilities (P) between the weights of SPCs and the measurements of spectral features from Shen et al. (2011), such as the velocity width and EW of emission lines. The results are listed in Table 1, and the (anti-)correlations are very consistent with the interpretation of those SPCs. The spectral relationships in these SPCs are described as follows.

(i) SPC1 is dominated by the anti-correlation between the narrow [O III] lines and the optical Fe II emission blends, which is the well-known EV1. Unlike

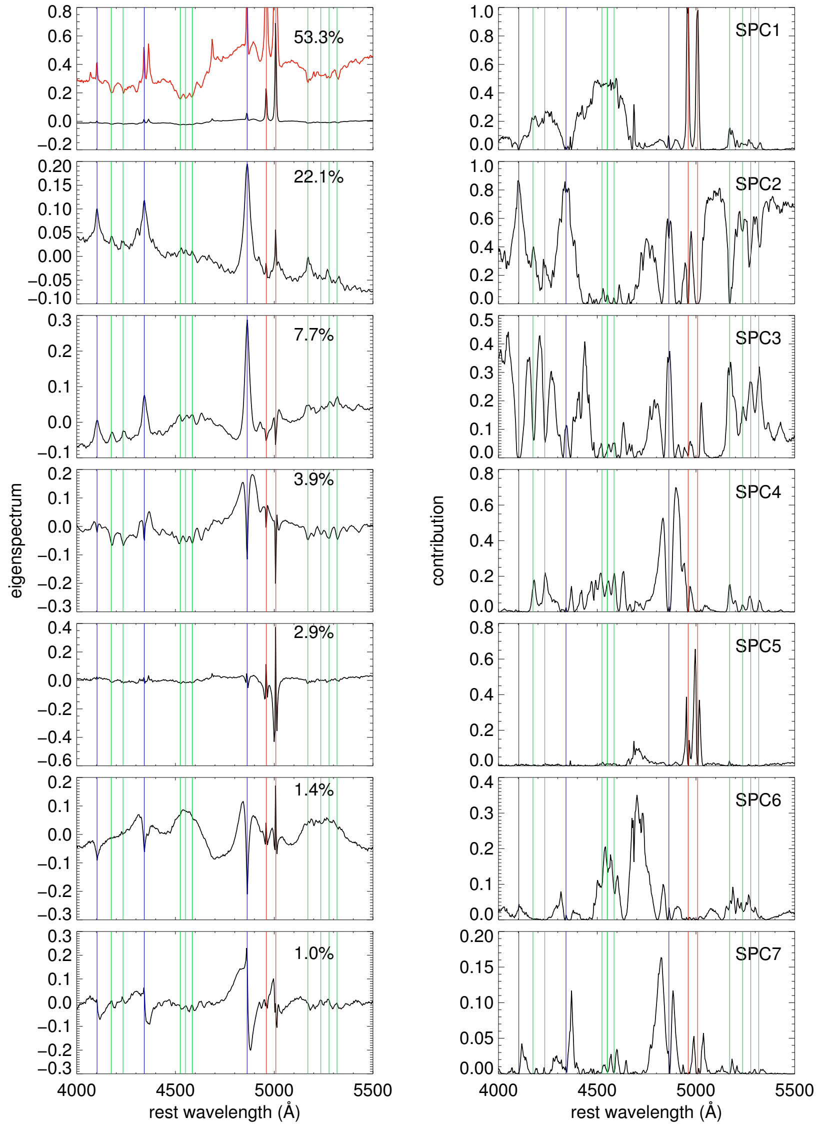

Fig. 2 The SPCA results for our sample. From top to bottom are the first seven SPCs. *Left panel*: the eigenspectra, where features with the same sign are correlated while those with opposite signs are anti-correlated. The percentages denote the fraction of total intrinsic variance contributed by each corresponding SPC. For SPC1, the eigenspectrum is also enlarged by 10 times and shifted up by 0.4 to present Fe II clearly and plotted in *red*. *Right panel*: the fractional-contribution spectra, showing the fraction of variance in individual wavelength bins contributed by corresponding SPCs. The colored vertical lines signify the positions of emission lines, with *blue* for H Balmer series, *red* for [O III] and *green* for the strongest multiplets of Fe II.

Boroson & Green (1992), there is little correlation with FWHM(Hβ), which is consistent with the little correlation ($\rho = 0.11$, $P = 1.5 \times 10^{-30}$) between EW([O III]) and FWHM(Hβ). This difference between our SPC1 and EV1 of Boroson & Green (1992) is due to our sample selection, and also because our sample is much larger than Boroson & Green (1992) by two orders of magnitude.

(ii) SPC2 represents the correlation between the strength of the hydrogen Balmer lines: Hβ, Hγ and Hδ, which are also anti-correlated with the continuum slope α. The continuum is bluer when Hβ is stronger. Although it is not clear in the eigenspectrum, SPC2 and FWHM(Hβ) have a moderately significant Spearman rank correlation coefficient of -0.46. Indeed, Hβ in the composite spectrum tends to be narrower when SPC2 increases, the continuum is bluer or EW(Hβ) increases.

(iii) The “W” shape of Hβ in SPC3 indicates the velocity width variation of Hβ and it is correlated with α and Fe II. The continuum is bluer when Hβ is broader,

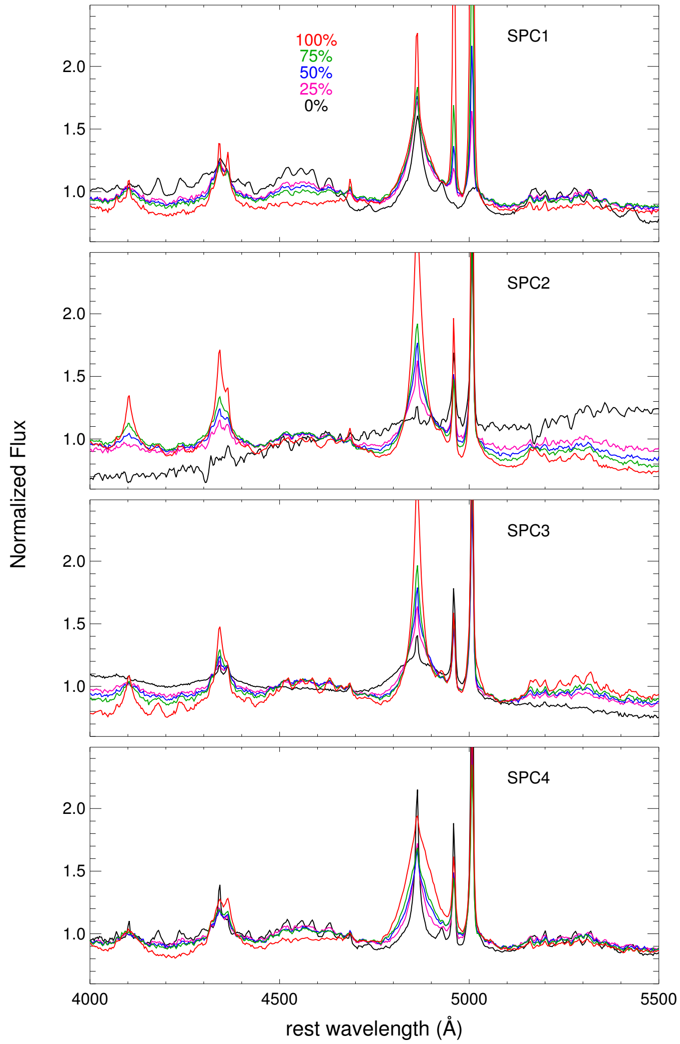

Fig. 3 The composite spectra are presented with increasing weights for each SPC. The quasars were sorted by the weights along each SPC, and 100 spectra were selected to build each composite spectrum from the minimum weight to the maximum, indicated by colors. From top to bottom are the first four SPCs, respectively. The changes in composite spectra clearly represent the spectral relationship of individual SPC.

which, however, is contrary to their relationship in SPC2.

(iv) SPC4 is similar to SPC3, both of which contain features contributing to the velocity width variation of Hβ. However, SPC3 represents the change in the Hβ core region, while SPC4 represents the broader component. This difference also appears in the fractional-contribution spectra: SPC3 peaks at the central wavelength of Hβ, while SPC4 dips at the core of Hβ but rises in the wings. In addition, SPC4 has little correlation with α, while SPC3 has little correlation with EW(Hβ). Besides, compared to SPC2, which has a positive correlation with FWHM(Hβ) but an anti-correlation EW(Hβ), SPC4 has positive correlations with both of them.

(v) SPC5 manifests the velocity width variation of [O III], consistent with the significant correlation with the full width at half maximum (FWHM) of the narrow lines.

(vi) SPC6 shows an anti-correlation between components of Fe II and He II λ4686 emission. Moreover, in the composite spectra, Fe II is broader with almost absent multiplets when stronger, which is not the case in SPC1. This is consistent with the fractional-contribution spectrum of SPC6, where peaks in the Fe II region are not located at the wavelengths of the strongest multiplets (as in the case of SPC3).

(vii) SPC7 manifests a shift of Hβ, consistent with the significant correlation with offset velocity of Hβ. The composite spectrum of quasars with the minimum

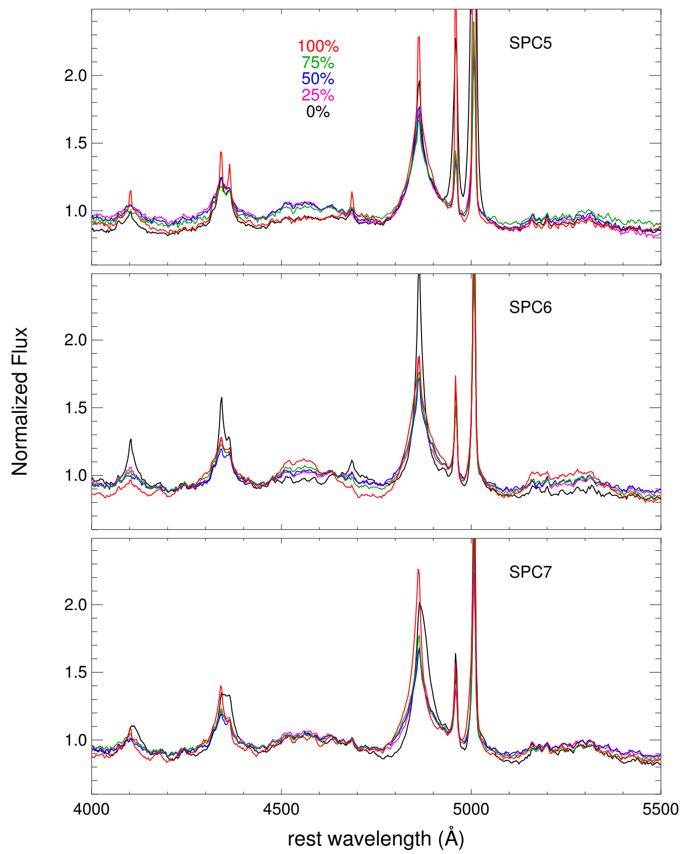

Fig. 4 The same as Fig. 3 but for SPC5 to SPC7.

Table 1 Spearman Rank Correlation Coefficients between the First Seven SPCs and Spectral Fitting Results

Fractional contribution (%)	SPC1 53.3	SPC2 22.1	SPC3 7.7	SPC4 3.9	SPC5 2.9	SPC6 1.4	SPC7 1.0
EW([O III] λ5007)	**0.92**	0.14	–0.05	0.11	–0.14	–0.02	0.06
EW(Fe)	**–0.49**	0.26	0.25	–0.20	–0.32	**0.56**	–0.09
α	0.22	**–0.81**	**0.40**	0.01	0.11	0.04	–0.01
FWHM(H$\beta_{\rm broad}$)	0.18	**–0.46**	**–0.55**	**0.42**	0.08	0.09	0.04
EW(Hβ)	0.22	**0.57**	–0.03	**0.59**	–0.09	0.24	0.02
FWHM(H$\beta_{\rm narrow}$)	–0.21	0.16	0.08	0.11	**–0.44**	–0.07	–0.16
$v_{\rm offset}$(Hβ)	0.05	–0.18	–0.06	–0.03	0.06	0.13	**0.43**

The numbers below the SPCs are fractional contributions to the total intrinsic variance. The spectral fitting results are referenced from Shen et al. (2011). The correlation coefficients with absolute values greater than 0.4 are expressed in bold. The associated probabilities that this distribution of data occurs by chance are not presented because they are all very close to zero, except for those with very small absolute coefficients.

(maximum) SPC7 also exhibits an obvious red(blue)-shift.

The fractional-contribution spectra can clearly show the significance of each SPC at different wavelengths. However, there are some complications because we cannot separate emission lines from continuum in each wavelength bin and the standard deviation is based on the total flux. This is not a problem for strong emission lines because they always dominate the continuum flux, but for the wavelength bins in the Fe II regions where emission lines are relatively weak and vary less than the continuum, the values of the variance of these lines are relatively small, but the actual fractional-contribution of these lines is still sometimes remarkable. Moreover, although in some cases there are obvious features in the eigenspectrum, the fractional-contribution of this SPC at these wavelengths is almost zero. There are two reasons for this. The first, described by Hu et al. (2012), is that the features in eigenspectra may be caused by cross talk rather than real correlations (e.g., the features around [O III] in SPC2 to SPC4).

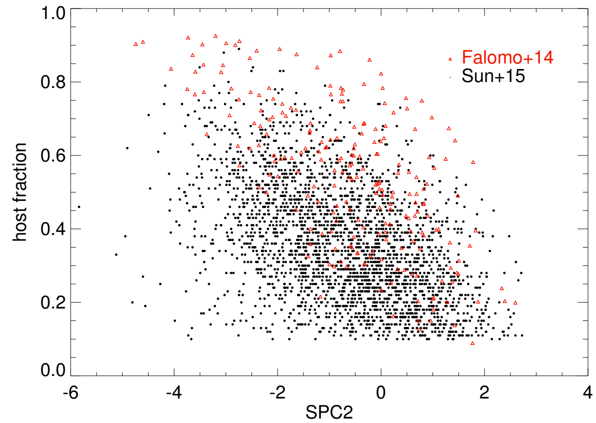

Fig. 5 Host galaxy fraction is plotted against SPC2. The *black dots* utilize the host fraction at 4185 Å taken from Sun & Shen (2015), and the *red triangles* apply the host fraction inferred from i-band magnitude by Falomo et al. (2014). Both host fractions are significantly anti-correlated with SPC2.

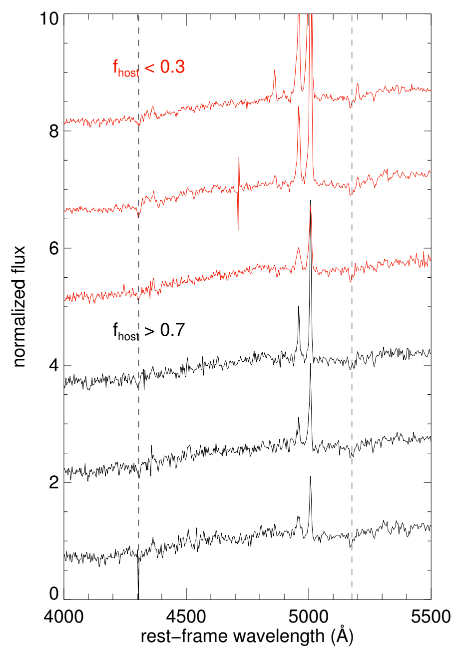

Fig. 6 Example spectra of quasars with different host fractions but similar SPC2. The spectra with SPC2 < -4 are expected to exhibit large host fraction, some of which manifest host fraction larger than 0.7 in Sun & Shen (2015) (*black lines*), while others have host fraction smaller than 0.3 (*red lines*), which are inconsistent with SPC2. However, the galaxy absorption lines at wavelengths marked by *dashed lines* are also clearly present in the red spectra of small host fractions, which are very similar to the black spectra with large host fractions. This implies that the host fractions for the red ones were underestimated but SPC2 traces the host fraction more accurately.

The other reason is that the coefficients in eigenspectra are nearly zero, like the Hδ feature in SPC3, because the emission lines are anti-correlated with the continuum at these wavelength bins.

We conclude that the major contributors of the SPCs for the variance of individual lines are as follows.

(i) [O III]: SPC1 contributes almost all the variance in the core region, and SPC5 affects the wing of the profile.

(ii) Hβ: SPC2 and SPC3 contribute the most in the core region while SPC4 dominates the wing region.

(iii) Fe II: This is more complicated than other lines. There are three main groups of Fe II multiplets between 4000 and 5500 Å (Kovačević et al. 2010), which approximately correspond to the blue, central and red parts, respectively. The central part is the most variable, the blue part is moderate, while the red part is marginal compared to the contin-

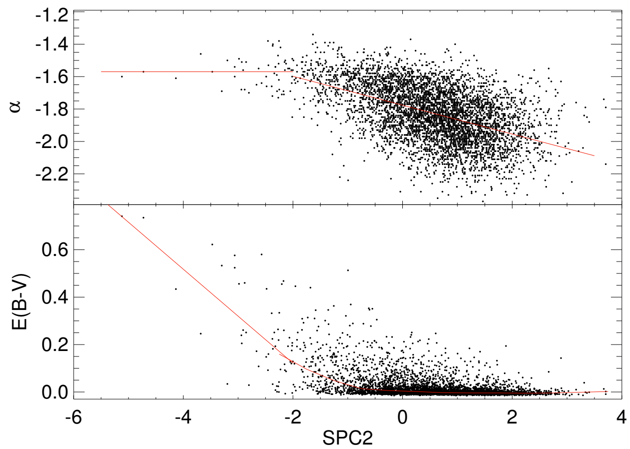

Fig. 7 The quasar intrinsic continuum spectral index α (*upper*) and $E(B-V)$ for the Small Magellanic Cloud (SMC) reddening law (*lower*) are plotted against SPC2. The α and $E(B-V)$ are referenced from Krawczyk et al. (2015). In the upper panel, the *solid line* is the median value for SPC2 < -2 and the best linear-fitting for SPC2 > -2. In the lower panel, the *solid line* signifies the best linear-fitting for SPC2 < -2 and the median values for SPC2 bins with interval of 0.5 for SPC2 > -2.

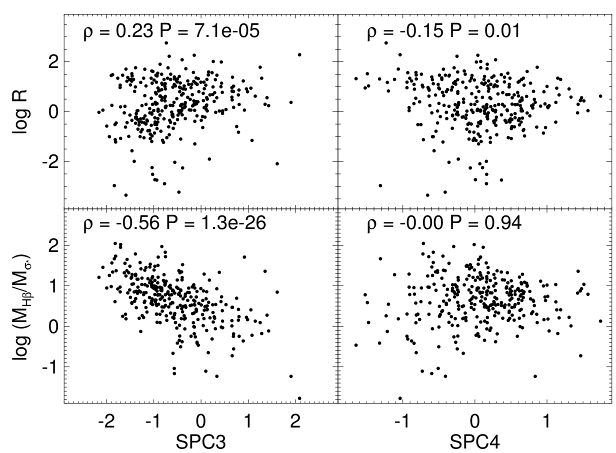

Fig. 8 The quasar orientation proxies from Brotherton et al. (2015) and Runnoe & Boroson (2019, submitted to ApJ) are plotted against SPC3 (*left*) and SPC4 (*right*), respectively. $\log R$, the radio core dominance, is displayed in the upper panels, and $\log(M_{\mathrm{H}\beta}/M_{\sigma_*})$ is depicted in the lower panels. The Spearman rank correlation coefficient ρ and probability P are provided for each panel.

uum (see standard deviation spectrum in Fig. 1). However, the contributions from individual SPCs are diverse. From Table 1, the EW(Fe II) of the central part mainly correlates with SPC1 and SPC6, while SPC2 to SPC4 also contribute some as seen from fractional-contribution spectra, but the importance of individual SPC varies for these groups. SPC1 dominates the blue and central parts while SPC3 dominates the red part. In addition, SPC6 contributes significantly to the central and red parts, but little to the blue part. The little contribution of SPC2 to the red part, and the little contribution of SPC3 to the blue part are due to the anti-correlation between Fe II and continuum. These differences indicate that these Fe II parts may have different origins.

(iv) Continuum: SPC2 contributes the most, and SPC3 supplements the blue parts.

We also note that in Table 1, the signs of coefficients for EW(Fe II) and FWHM(Hβ) are nearly always opposite. This effect is apparent in Figure 3 and Figure 4 in that when Fe II is stronger, Hβ tends to be narrower for SPC1, 3, 4 and 6. It agrees with the idea that Hβ originates in two regions: a broad-line region and an intermediate-width region, the latter of which is close to the gas emitting Fe II (e.g., Brotherton 1996; Hu et al. 2012).

5 DISCUSSION

To reveal the underlying physical meaning of the principal components, we studied their correlations with the inferred physical properties of quasars, such as host galaxy contribution, reddening and orientation. The measured values are obtained from the literature, and we calculated the Spearman rank correlation coefficients between the weights of principal components and these parameters. The results are discussed below in detail.

5.1 Host Galaxy Fraction

As seen in Figure 3, when SPC2 is at its maximum, almost no broad quasar emission lines are visible, while the galaxy absorption lines, such as Ca G-band 4304 Å and Mg 5175 Å, are obvious, indicating that SPC2 reflects the proportion of host galaxy contribution in the spectrum. We took the host galaxy fractions from Sun & Shen (2015) and Falomo et al. (2014), which separate the quasar and galaxy parts from spectra and images, respectively. There are 2592 and 225 matched quasars with our sample, respectively. We then compared the SPC2 and host fraction in Figure 5. It can be seen that SPC2 correlates clearly with both host fractions, with $\rho = -0.46$ and -0.62, respectively. The host fraction in Sun & Shen (2015) is smaller on average, because the spectra were only taken from the center of targets, while Falomo et al. (2014) also took into account the extended flux of galaxies from the images.

For instance, quasars with SPC2 < -4 are expected to be heavily contaminated by the host, and there is clear and consistent evidence that some of the quasars do have large host fraction values in Sun & Shen (2015). However, it is surprising that others have very small values. We visually inspected these spectra with small host fraction in Sun & Shen (2015) in Figure 6, and found they also showed significant galaxy absorption lines but weak quasar broad lines, indicating that they actually have large host fractions. Accordingly, it is likely that Sun & Shen (2015) underestimated the host fractions for these quasars, but our SPC2 is more reliable to trace the host contamination.

Although SPC3 also contains the continuum slope, it has much weaker correlation with host fraction, with $\rho = 0.24$ and 0.18, respectively. In the composite spectrum of minimum SPC3, the spectrum is reddest but has little galaxy absorption lines, implying that the continuum slope in SPC3 is mostly quasar intrinsic continuum slope.

5.2 Red or Reddened

The variation of continuum slope is caused by both intrinsic colors of quasars and reddening by dust extinction. Krawczyk et al. (2015) distinguished the two parts and derived intrinsic slope α and extinction $E(B-V)$ from the spectral energy distribution (SED) between 1216 Å and 1 μm for a large sample. We find that SPC2 has similar degrees of correlation with both α and $E(B-V)$ for 4889 matched quasars, with $\rho = 0.49$ and -0.41, respectively (see Fig. 7).

Furthermore, the data points in Figure 7 can be roughly divided into two groups, whose trends are marked by red lines. In the left region (SPC2 < -2), as SPC2 increases, $E(B-V)$ decreases, while the quasar slope described by the parameter α is almost a constant, indicating that these quasars are intrinsically the reddest. In the right region, as SPC2 increases, α decreases linearly despite large scatter, while $E(B-V)$ is mostly very small with decreasing scatter. As a result, the reasons that the observed continuum is bluer as SPC2 increases are different. This is caused by decreasing extinction for SPC2 < -2, but mainly caused by intrinsically bluer color for SPC2 > -2.

Therefore, SPC2 can be utilized as a simple but efficient proxy to quantitatively distinguish between intrinsic color and extinction, although it is derived from only a much narrower wavelength range than Krawczyk et al. (2015). Again, SPC3 has much weaker correlation with α or $E(B-V)$, with $\rho = -0.05$ and 0.12, respectively, implying that SPC3 cannot distinguish between intrinsic color and reddening.

5.3 Orientation

FWHM(Hβ) is one of the primary parameters in the quasar main sequence, and its observed values are affected by viewing angle or orientation. In our analysis, SPC3 and SPC4 both present the velocity width variation of Hβ; thus, we studied their relationships with orientation. The data on orientation are referenced from Brotherton et al. (2015) and Runnoe & Boroson (2019, submitted to ApJ), who provided two proxies to characterize the orientation: the radio core dominance, R, the ratio of radio flux densities between core and lobe, and the black hole mass ratio $M_{\mathrm{H}\beta}/M_{\sigma_*}$, which were calculated based on single-epoch spectra and the $M-\sigma_*$ relationship, respectively. The quasar tends to be face-on if its log R is larger or $M_{\mathrm{H}\beta}/M_{\sigma_*}$ is smaller. The plots of orientation proxies against SPC3 and SPC4 are shown in Figure 8. We can see that the correlation of SPC3 with $\log R$ is marginal ($\rho = 0.23$), not as strong as the correlation with $M_{\mathrm{H}\beta}/M_{\sigma_*}$ ($\rho = -0.56$). However, SPC4 has little correlation with orientation regardless of its significant correlation with FWHM(Hβ). Therefore, SPC3 is more likely to be a proxy of orientation.

6 SUMMARY

We performed SPCA for a sample of 10 241 low redshift quasars for the wavelength interval 4000 – 5500 Å, involving Hβ, [O III] and Fe II spectral features. The first seven principal components contribute 92.3% of the intrinsic variance in the sample, and they have some clear correlations with spectral properties including the continuum slope, as well as the intensity and velocity width of the narrow and/or broad lines. Consequently, SPCA can be an efficient method to select quasars with special spectral properties without model-dependent spectral fitting or measurements. In several SPCs, EW(Fe II) and FWHM(Hβ) are almost always anti-correlated, consistent with the scenario that there is an intermediate-width component of broad Hβ, which is correlated with Fe II (e.g., Brotherton 1996).

Additionally, we interpreted the underlying drivers of some SPCs by comparing them with measured quasar spectral properties. The second component is a combination of host contamination, intrinsic color and dust extinction, and is able to quantify them well compared with other methods. The third component is possibly a proxy of orientation.

We summarize our first seven principal components as follows:

(i) The anti-correlation between EW([O III]) and EW(Fe II), which is the traditional EV1;
(ii) The correlation between continuum slope and EW(Hβ), which indicates host contamination, intrinsic color or dust extinction;
(iii) The correlation between continuum slope and Hβ velocity width, which is a possible proxy of orientation;
(iv) The correlation between FWHM and EW of Hβ, which has little correlation with orientation;
(v) The velocity width variation of [O III];
(vi) The anti-correlation between EW(Fe II) and EW(He II λ4686), and Fe II tends to be broader when stronger;
(vii) The shift of Hβ.

The results are a little complicated because the strength and/or velocity width of some lines may appear in several SPCs. To better understand the importance of each SPC on individual lines, we also calculate the fractional-contribution spectra. Any spectral feature is dominated by two or three SPCs. Based on this, we note that there are different groups of Fe II multiplets, implying complicated emitting processes and origins for quasar optical Fe II emissions.

Acknowledgements This paper has used data from the SDSS. Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The SDSS Web Site is `http://www.sdss.org/`.

References

Bensch, K., del Olmo, A., Sulentic, J., Perea, J., & Marziani, P. 2015, Journal of Astrophysics and Astronomy, 36, 467
Bian, W.-H., He, Z.-C., Green, R., et al. 2016, MNRAS, 456, 4081
Boroson, T. A., & Green, R. F. 1992, ApJS, 80, 109
Brotherton, M. S. 1996, ApJS, 102, 1
Brotherton, M. S., Singh, V., & Runnoe, J. 2015, MNRAS, 454, 3864
Falomo, R., Bettoni, D., Karhunen, K., Kotilainen, J. K., & Uslenghi, M. 2014, MNRAS, 440, 476
Francis, P. J., Hewett, P. C., Foltz, C. B., & Chaffee, F. H. 1992, ApJ, 398, 476
Gallant, D., Gallo, L. C., & Parker, M. L. 2018, MNRAS, 480, 1999
Hewett, P. C., & Wild, V. 2010, MNRAS, 405, 2302
Hu, C., Wang, J.-M., Ho, L. C., et al. 2012, ApJ, 760, 126
Kovačević, J., Popović, L. Č., & Dimitrijević, M. S. 2010, ApJS, 189, 15
Krawczyk, C. M., Richards, G. T., Gallagher, S. C., et al. 2015, AJ, 149, 203
Laor, A., Fiore, F., Elvis, M., Wilkes, B. J., & McDowell, J. C. 1997, ApJ, 477, 93
Marziani, P., Sulentic, J. W., Zwitter, T., Dultzin-Hacyan, D., & Calvani, M. 2001, ApJ, 558, 553
Panda, S., Czerny, B., Adhikari, T. P., et al. 2018, ApJ, 866, 115
Pennell, A., Runnoe, J. C., & Brotherton, M. S. 2017, MNRAS, 468, 1433
Rochais, T., Singh, V., Chick, W., et al. 2017, MNRAS, 464, 553
Schneider, D. P., Richards, G. T., Hall, P. B., et al. 2010, AJ, 139, 2360
Shang, Z., Wills, B. J., Robinson, E. L., et al. 2003, ApJ, 586, 52
Shen, Y., & Ho, L. C. 2014, Nature, 513, 210
Shen, Y., Richards, G. T., Strauss, M. A., et al. 2011, ApJS, 194, 45
Śniegowska, M., Czerny, B., You, B., et al. 2018, A&A, 613, A38
Sulentic, J. W., Marziani, P., & Dultzin-Hacyan, D. 2000, ARA&A, 38, 521
Sun, J., & Shen, Y. 2015, ApJ, 804, L15
Sun, M., Xue, Y., Wang, J., Cai, Z., & Guo, H. 2018, ApJ, 866, 74
Vanden Berk, D. E., Shen, J., Yip, C.-W., et al. 2006, AJ, 131, 84
Wang, T., Brinkmann, W., & Bergeron, J. 1996, A&A, 309, 81
Wills, B. J., & Browne, I. W. A. 1986, ApJ, 302, 56
Yip, C. W., Connolly, A. J., Vanden Berk, D. E., et al. 2004, AJ, 128, 2603