

The Pennsylvania State University

The Graduate School

Department of Astronomy and Astrophysics

A STUDY OF UV, OPTICAL AND X-RAY EMISSION FROM  
QUASARS

A Dissertation in

Astronomy and Astrophysics

by

Jian Wu

© 2011 Jian Wu

Submitted in Partial Fulfillment  
of the Requirements  
for the Degree of

Doctor of Philosophy

August 2011

The dissertation of Jian Wu was reviewed and approved<sup>1</sup> by the following:

Donald P. Schneider  
Professor of Astronomy and Astrophysics  
Dissertation Adviser  
Co-Chair of Committee

Jane C. Charlton  
Professor of Astronomy and Astrophysics  
Co-Chair of Committee

Niel W. Brandt  
Professor of Astronomy and Astrophysics

Yuexing Li  
Assistant Professor of Astronomy and Astrophysics

Paul Sommers  
Professor of Physics

Daniel E. Vanden Berk  
Assistant Professor of Physics, Special Member

Steinn Sigurdsson  
Associate Professor of Astronomy and Astrophysics  
Chair of Graduate Program

---

<sup>1</sup>Signatures on file in the Graduate School.

## Abstract

The rest-frame UV, optical, and X-ray are important wavebands to study quasar activities because they trace ionization phases and dynamical properties of the physical structures surrounding the central engine, including the accretion disk, the broad line region, and the outflow.

This thesis investigates the correlations of spectral components (continua, emission lines and absorption lines) and the physical conditions implied by these components using well-defined samples of spectra from the archives of the Sloan Digital Sky Survey (SDSS), the Hubble Space Telescope (HST), the International Ultraviolet Explorer (IUE), the Swift Space Telescope, and the Keck Telescope. The scientific goal of these studies is to produce a set of well-documented properties of quasar UV, optical, and X-ray spectra that will provide important constraints on models of the quasar phenomenon.

The UV and optical continuum properties are studied using a set of composite quasar spectra created from the SDSS quasar catalog. We discovered a dependency of the continuum shape on both quasar luminosity and redshift, and found that the Baldwin Effect, in which the equivalent width of an emission line is anti-correlated with the continuum, applies for most of the broad emission lines. Focusing on the CIV emission line, we construct a sample which consists of 272 Type 1 AGNs. An analysis of this sample indicates that the changes of the EWs are controlled by both UV and soft X-ray emissions, suggesting that the UV and X-ray emissions are closely related.

We show that variability is an important contributor to the scatter of many important correlations, such as the Baldwin Effect, as well as the quasar spectral energy distribution (SED) between optical and X-ray. In order to eliminate the influence of variability, we compiled a catalog of quasars with simultaneous UV, optical and X-ray observations from the *Swift* gamma-ray-burst explorer. Using the *Swift* data, plus the photometric data from SDSS, we constructed a set of composite quasar SEDs for most of the objects in our catalog. We use this set of SEDs to constrain the amplitude of the big blue bump. We also re-visit the  $l_{\nu}(2500 \text{ \AA}) - \alpha_{\text{ox}}$  relation and demonstrate that the dispersion of this correlation can be reduced by  $\sim 20\%$  using this simultaneously observed sample.

Quasar intrinsic absorption lines are important diagnoses of the quasar outflow. Models of intrinsic narrow absorption line systems towards three high redshift quasars from Keck/HIRES, reveals high ionization levels and super-solar metallicities of these absorbers, indicating vigorous star formation occurred before the quasars become visible.

## Table of Contents

List of Tables . . . . .	vii
List of Figures . . . . .	ix
Acknowledgments . . . . .	xii
Chapter 1. Introduction . . . . .	1
1.1 AGN Taxonomy and Multiple Wavelength Properties . . . . .	1
1.2 Unified Quasar Structure . . . . .	3
1.3 Quasar Surveys . . . . .	5
1.4 Unified Themes and Organizations of This Thesis . . . . .	6
1.4.1 UV/Optical Emission . . . . .	6
1.4.2 X-ray Emission . . . . .	8
1.4.3 Thesis Organization . . . . .	8
Chapter 2. Continuum and Emission Line Properties of Composite Quasar Spectra from SDSS . . . . .	13
2.1 Introduction . . . . .	13
2.2 Generating the Composite Spectra . . . . .	14
2.3 Spectral Fitting . . . . .	19
2.3.1 Continuum Subtraction . . . . .	19
2.3.2 Emission Lines . . . . .	22
2.3.3 Error Analyses . . . . .	22
2.3.4 Sensitivity of Fitting Results to Initial Guess Values . . . . .	23
2.4 Spectral Properties . . . . .	24
2.4.1 Continuum Properties . . . . .	24
2.4.2 Properties of Emission Lines . . . . .	25
2.4.2.1 Baldwin Effect (BEff) . . . . .	25
2.4.2.2 Systematic Emission Line Shift . . . . .	28
2.4.2.3 Full Width at Half Maximum (FWHM) of Emission Lines . . . . .	30
2.5 Discussion . . . . .	31
2.6 Conclusion . . . . .	32
Chapter 3. Probing the Origins of the C IV and Fe K $\alpha$ Baldwin Effects . . . . .	66
3.1 Introduction . . . . .	66
3.2 Sample Construction . . . . .	68
3.2.1 Sample A . . . . .	68
3.2.2 Sample B . . . . .	69
3.2.3 Sample C . . . . .	70
3.2.4 Combined Sample . . . . .	70

3.3	Data Processing . . . . .	70
3.3.1	Bad Pixel Removal and Reddening Correction . . . . .	70
3.3.2	Spectral Fitting . . . . .	71
3.3.2.1	Sample A . . . . .	71
3.3.2.2	Sample B . . . . .	72
3.3.2.3	Sample C . . . . .	72
3.3.3	Error Analysis . . . . .	73
3.4	Drivers of the C IV Baldwin Effect . . . . .	74
3.4.1	Comparison with Previous Work . . . . .	74
3.4.2	The Effects of $\alpha_{\text{ox}}$ on the C IV and Fe K $\alpha$ BEffs . . . . .	75
3.4.3	Effects of AGN Variability on BEff Relation Scatter . . . . .	76
3.4.4	Regressions of EW and Luminosity . . . . .	77
3.5	Relation Between Fe K $\alpha$ and C IV . . . . .	78
3.6	Discussions and Conclusions . . . . .	79
 Chapter 4.	A Quasar Catalog with Simultaneous UV, Optical and X-ray Observations by <i>Swift</i> . . . . .	95
4.1	Introduction . . . . .	95
4.2	Observations and Data Processing . . . . .	97
4.2.1	Instruments . . . . .	97
4.2.2	XRT Data . . . . .	98
4.2.3	Weak X-ray Sources . . . . .	100
4.2.4	UVOT Data . . . . .	102
4.3	Quasar SEDs . . . . .	105
4.3.1	Supplementary Data . . . . .	105
4.3.2	Initial SED Plots . . . . .	106
4.3.3	Emission Line Correction . . . . .	107
4.3.4	Photometric Shift . . . . .	108
4.3.5	Error Analyses . . . . .	109
4.3.6	SED Models . . . . .	109
4.3.7	Bolometric Luminosity and Black Hole Mass . . . . .	111
4.4	Catalog Description . . . . .	112
4.4.1	Comments on Individual Objects . . . . .	113
4.5	Correlation Analyses . . . . .	115
4.5.1	Sample Selection . . . . .	115
4.5.2	The $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$ Relation . . . . .	118
4.6	Conclusions . . . . .	119
 Chapter 5.	The Physical Conditions of the Intrinsic Nv Narrow Absorption Line Systems of Three Quasars . . . . .	162
5.1	Introduction . . . . .	162
5.2	Data . . . . .	165
5.3	Photoionization Modeling . . . . .	167
5.3.1	Voigt Profile Fitting of the Nv Doublets . . . . .	167
5.3.2	Specification of the Ionizing Continuum . . . . .	168

5.3.3	Comparison of Cloudy Model Results to the Data and the Role of the Coverage Fraction . . . . .	169
5.4	Photoionization Model Results . . . . .	170
5.4.1	HE0130-4021 . . . . .	170
5.4.2	Q1009+2956 . . . . .	172
5.4.3	HS1700+6416 . . . . .	172
5.4.4	The O VI Doublet . . . . .	173
5.5	Partial Coverage of Continuum and Broad-Emission Line Regions . .	173
5.6	Discussion . . . . .	176
5.6.1	High Metallicities . . . . .	176
5.6.2	Properties of the Absorbing Gas . . . . .	177
5.7	Summary and Conclusions . . . . .	179
Chapter 6. Summary of Thesis Results . . . . .		201
Appendix A Fitting the Continuum of a Composite SDSS Quasar Spectrum Using CMA-ES . . . . .		203
A.1	Introduction . . . . .	203
A.2	Problem Description . . . . .	204
A.3	Solution Strategy . . . . .	205
A.4	Program Design . . . . .	208
A.5	Optimization Results . . . . .	209
A.5.1	Testing Program . . . . .	209
A.5.2	Results in Default Setup . . . . .	210
A.5.3	Varying Parameters . . . . .	210
A.6	Conclusion and Discussion . . . . .	211
Appendix B Software Design & User Manual for QSFIT . . . . .		225
B.1	Introduction . . . . .	225
B.1.1	Software Availability . . . . .	225
B.1.2	Convention . . . . .	226
B.1.3	Copyright and Maintenance . . . . .	226
B.2	Software Structure . . . . .	226
B.2.1	Data Structure and Algorithms . . . . .	226
B.2.2	Software Parts . . . . .	226
B.3	User Manual . . . . .	230
B.3.1	Power-law Estimation . . . . .	230
B.3.2	Continuum Join-Fit . . . . .	233
B.3.3	Emission Lines Fitting . . . . .	237
Bibliography . . . . .		241

## List of Tables

2.1	Empirical line-free windows used to estimate the power-law component.	59
2.2	Benchmark wavelengths used to evaluate the goodness of fit. . . . .	60
2.3	Underlying continuum parameters in joint fitting phase. . . . .	61
2.4	Quasar emission line list. . . . .	62
2.5	Varying initial values of continuum fitting parameters. . . . .	63
2.6	BEff slopes and intercepts for three emission lines. . . . .	64
2.7	Baldwin Effect slope $k$ for different emission lines as a function of ionization potentials. . . . .	65
3.1	Summary of samples. . . . .	88
3.2	UV properties of the combined sample. . . . .	89
3.3	UV and X-ray properties of Sample A objects. . . . .	90
3.4	Hypothesis and linear fitting results. . . . .	91
3.5	Correlation and partial-correlation analysis results. . . . .	92
3.6	The RMS values of residuals after regression from different variables. .	93
3.7	Correlation and regression analysis for EW, emission line luminosity, and flux data between Fe K $\alpha$ and C IV. . . . .	94
4.1	XRT Data Binning Strategy. . . . .	150
4.2	Models used to fit XRT spectra. . . . .	151
4.3	$F_{\text{PSF}, \text{Xspec}}^0 / \text{CR}$ vs. offset angle $\theta$ . . . . .	152
4.4	UVOT Sky Image Flagging Strategy. . . . .	153
4.5	List of reasons to exclude unacceptable UVOT sky images. . . . .	154
4.6	Four types of objects in the raw catalog. . . . .	155
4.7	Central wavelengths and FWHMs of UV/optical wavebands. . . . .	156
4.8	Photometric shift strategies. . . . .	157
4.9	Parameters for black hole mass calculation. . . . .	158
4.10	Catalog description. . . . .	159
4.11	Classification for the entire SED sample. . . . .	160
4.12	Correlation and regression analysis. . . . .	161
5.1	Summary of Properties of Intrinsic NAL Systems <sup>a</sup>	194
5.2	Results of Decomposition of N V Doublet Profiles <sup>a</sup>	195
5.3	Spectral Indices of Different SED Models Compared to Observed Values	196
5.4	Summary of Preliminary Fitting Results	197
5.5	Best model parameters for the Q1009+2956 absorption systems	198
5.6	Best model parameters for the HS1700+6416 absorption systems	199
5.7	Constraints on Continuum and BELR Coverage Fractions	200
A.1	Spectral components and free parameters. . . . .	220
A.2	Initialization. Values of $\mathbf{m}$ are typical values from previous results. . . . .	221

A.3	Optimal parameter values using LMA and CMA-ES (under the default setup). . . . .	222
A.4	Changes of termination generation as a function of population size $\lambda$ . Proportion of selected offspring is fixed to be 50%. . . . .	223
A.5	Numbers of generations as a function of the proportion of selected offspring with respect to the fixed total population size $\lambda = 220$ . . . . .	224
B.1	Columns in <i>bicheck.pro</i> output file. . . . .	231
B.2	Flags in <i>bicheck.list</i> . . . . .	232
B.3	Parameters and keywords of <i>bicheck.pro</i> . . . . .	232
B.4	Parameters and keywords in <i>qconfita.pro</i> . . . . .	235
B.5	Parameters and keywords in <i>qconsyn.pro</i> . . . . .	236
B.6	Columns in <i>vmlfit.[linename].gss</i> . . . . .	239
B.7	Columns in <i>vmlfit.[linename].par</i> . . . . .	239

## List of Figures

1.1	A mean RQ quasar SED . . . . .	10
1.2	Unification model of AGNs . . . . .	11
1.3	A quasar composite spectrum. . . . .	12
2.1	Calibration correlations of $l_\lambda$ at multiple wavelengths . . . . .	35
2.2	Redshift-luminosity binning diagram for DR3 . . . . .	36
2.3	Luminosity-redshift binning diagram for DR7 . . . . .	37
2.4	Small Balmer Bump shape . . . . .	38
2.5	Composite spectrum continuum fitting examples. . . . .	39
2.6	Composite spectrum emission line fitting examples . . . . .	40
2.7	Continuum parameter fitting results by varying $\alpha_1$ . . . . .	41
2.8	Continuum parameter fitting results by varying $\beta_1$ . . . . .	42
2.9	Continuum parameter fitting results by varying $A_B$ . . . . .	43
2.10	Continuum parameter fitting results by varying $A_{Fe,u}$ . . . . .	44
2.11	Continuum parameter fitting results by varying $A_{Fe,o}$ . . . . .	45
2.12	Changes of UV spectral index as functions of redshift and luminosity . . . . .	46
2.13	Illustration of Baldwin Effect . . . . .	47
2.14	Baldwin Effects for selected emission lines . . . . .	48
2.15	Distributions of EW uncertainties for multiple emission lines . . . . .	49
2.16	Changes of Baldwin Effect slopes with redshift . . . . .	50
2.17	BEff slopes vs. ionization potentials . . . . .	51
2.18	Emission line shifts histograms relative to reference lines . . . . .	52
2.19	Mean emission line shifts as a function of ionization potentials . . . . .	53
2.20	EW(Mg II) as a function of redshift . . . . .	54
2.21	Mean emission line shifts as a function of luminosity . . . . .	55
2.22	Mean emission line shift as a function of redshift . . . . .	56
2.23	FWHMs as functions of redshift and luminosity . . . . .	57
2.24	FWHMs as functions of redshift and luminosity from another perspective	58
3.1	Luminosity-redshift diagram of 272 objects . . . . .	81
3.2	C IV fitting examples . . . . .	82
3.3	C IV Baldwin Effect . . . . .	83
3.4	EW(C IV) as a function of $\alpha_{ox}$ . . . . .	84
3.5	EW(Fe K $\alpha$ ) as a function of $\alpha_{ox}$ . . . . .	85
3.6	EW(Fe K $\alpha$ vs. EW(C IV)) . . . . .	86
3.7	$f(\text{Fe K}\alpha)$ vs. $f(\text{C IV})$ . . . . .	87
4.1	Absolute magnitude-redshift diagram of raw catalog quasars . . . . .	121
4.2	Distribution of source region radii for reference object candidates . . . . .	122
4.3	Distribution of offset angle $\theta$ for reference objects selected . . . . .	123
4.4	$F_{PSF,Xspec}^0/\text{CR}$ vs. offset angle $\theta$ . . . . .	124
4.5	XRT spectra fitting examples . . . . .	125

4.6	UVOT sky image flagging strategy . . . . .	126
4.7	Examples of removed UVOT sky images . . . . .	127
4.8	Apparent magnitudes vs. redshift diagram . . . . .	128
4.9	UVOT quasar light curve examples . . . . .	129
4.10	A UVOT SED as a function of time . . . . .	130
4.11	Examples of initial SEDs . . . . .	131
4.12	Emission line corrections as a function of redshift . . . . .	132
4.13	Distributions of emission line correction differences . . . . .	133
4.14	Photometric UV slope vs. spectroscopic UV slope . . . . .	134
4.15	Examples of emission line corrected SEDs . . . . .	135
4.16	Photometric shift strategy . . . . .	136
4.17	Examples of photometric shifted SEDs . . . . .	137
4.18	Distributions of $\alpha_{\text{UVX}}$ vs. $\alpha_{\text{OX,ph}}$ . . . . .	138
4.19	Examples of SEDs fitting in EXP and TPL models . . . . .	139
4.20	Black hole mass as a function of bolometric luminosity . . . . .	140
4.21	Black hole mass as a function of redshift . . . . .	141
4.22	Black hole masses obtained using different estimators . . . . .	142
4.23	An SED example of a reddened quasar . . . . .	143
4.24	Relative color vs. redshift diagram . . . . .	144
4.25	Distributions of FWHM for H $\beta$ , Mg II, and C IV . . . . .	145
4.26	Distributions of bolometric luminosities from EXP and TPL models . . . . .	146
4.27	Sample size vs. exposure time for the parent sample . . . . .	147
4.28	X-ray detection rate vs. exposure time for the parent sample . . . . .	148
4.29	$\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$ relation for <i>Swift</i> AGN sample . . . . .	149
5.1	Absorption line constraints for HE0130–4021 . . . . .	180
5.2	Absorption line constraints for Q1009+2956 . . . . .	181
5.3	Absorption line constrains for HS1700+6416 . . . . .	182
5.4	Voigt profile fits to N V $\lambda\lambda$ 1239, 1242 doublets . . . . .	183
5.5	Input SEDs for photoionization model . . . . .	184
5.6	Z-U slide for HE0130–4021 . . . . .	185
7a	Observed and synthesized profiles of Si IV $\lambda$ 1396 for HE0130–4021 . . . . .	186
7b	Observed and synthesized profiles of Ly $\alpha$ for HE0130–4021 . . . . .	187
7c	Observed and synthesized profiles of Ly $\alpha$ for Q1009+2956 . . . . .	188
7d	Observed and synthesized profiles of C III $\lambda$ 977 for Q1009+2956 . . . . .	189
7e	Observed and synthesized profiles of Ly $\alpha$ for HS1700+6416 . . . . .	190
7f	Observed and synthesized profiles of C IV $\lambda$ 1549 for HS1700+6416 . . . . .	191
7g	Observed and synthesized profiles of C IV $\lambda$ 1549 for HS1700+6416 . . . . .	192
7h	Observed and synthesized profiles of Ly $\alpha$ for HS1700+6416 . . . . .	193
A.1	Quasar spectrum fit by LMA. . . . .	212
A.2	Rosenbrock function by varying population size . . . . .	213
A.3	Rosenbrock function by varying DOF . . . . .	214
A.4	Rosenbrock function converge generation distribution . . . . .	215
A.5	Spectral fitting using CMA-ES . . . . .	216

A.6 Evolutionary curves as the population size changes. . . . .	217
A.7 Evolutionary curves for different proportions of selected offspring with respect to population size . . . . .	218
A.8 Distribution of $\chi^2$ values at termination . . . . .	219

## Acknowledgments

I thank Professor Donald Schneider for financially supporting my astronomical research project. I also thank Dr. Daniel Vanden Berk for consistently advising me on research details. I thank Professor Jane Charlton on patiently advising and guiding my first paper for three years at Penn State. I thank Niel Brandt, Mike Eracleous, Toru Misawa and Rajib Ganguly for valuable discussions and comments on my papers. I thank all the *Swift* scientists and engineers at Penn State for helping me process my *Swift* data, in particular Dirk Grupe, Scott Koch, Caryl Gronwall, Jonathan Gelbord, and Jamie Kennea. I would also like to thank the following people who helped me with this thesis. They include David Borrows, Sarah Wesolowski, Blair Porterfield, George Chartas, and Yuexing Li.

Discover the power within yourself.

—Unknown

## Chapter 1

# Introduction

## 1.1 AGN Taxonomy and Multiple Wavelength Properties

An active galactic nucleus (AGN) is a compact region at the center of a galaxy, that is much more energetic than normal galaxies (e.g., the Milky Way) and generates emissions in multiple wavebands, which cannot be attributed to stars (Peterson 1997; Fabian 1999a; Kembhavi & Narlikar 1999). A galaxy that hosts an AGN is called an active galaxy. It is believed that unlike normal galaxies, whose energy is contributed mainly from stellar nuclear synthesis, the power of an AGN is primarily contributed by accretion into a super massive black hole (SMBH).

AGNs exhibit distinct features that are not seen in normal galaxies. For example, an AGN usually has a power-law continuum shape in UV and optical bands with a number of broad emission lines (e.g., Korista et al. 1995; Veilleux 2002). In X-ray band (0.3–20 keV), AGNs have a power-law continuum and often a prominent Fe K $\alpha$  emission line at  $\sim 6.4$  keV (e.g., Page et al. 2004b, 2005). Some AGNs have strong emission in radio (e.g., Laing 1993) and/or infrared band (e.g., Sanders et al. 1989) and numerous emission and absorption features (e.g., Maiolino & Risaliti 2007; Turnshek 1988).

Based on their observational properties, AGNs can be classified into the following categories. Because each category is defined based on multiple waveband properties, an AGN can belong to more than one categories.

- Seyfert galaxies. Seyfert galaxies (Seyfert 1943) are defined based on their relatively low luminosity ( $L_{\text{bol}} \sim 10^{44}$  erg s $^{-1}$ ) and emission line features at UV and optical bands. There are two types of Seyfert galaxies depending on the widths of emission lines. Seyfert 1 galaxies contain both broad ( $\text{FWHM}^1 \gtrsim 1000$  km s $^{-1}$ ) and narrow emission lines ( $\text{FWHM} \lesssim 1000$  km s $^{-1}$ ) while Seyfert 2 galaxies only contain narrow emission lines. There are also intermediate types which are given a fractional type value depending upon the relative strengths of broad and narrow components (e.g., Type 1.5, see Osterbrock & Ferland 2006). A subset of Seyfert 1 galaxies is called narrow line Seyfert 1 (NLS1) galaxy, which is commonly defined as  $\text{FWHM}(\text{H}\beta) < 2000$  km s $^{-1}$  and  $[\text{O III}] \lambda 5007/\text{H}\beta < 3$  (Osterbrock & Pogge 1985; Goodrich 1989). Because of their relatively low luminosity, Seyfert galaxies are currently found at low redshift ( $z \lesssim 0.5$ ) and have visible host galaxies.
- Radio galaxies. These are galaxies with strong radio emission compared with emission in other wavebands. The strength of radio emission is usually parameterized

---

<sup>1</sup>Full width of half maximum

as radio loudness  $R^*$ , which is defined as the ratio of monochromatic luminosities at rest-frame 5 GHz and 2500 Å (e.g., Stocke et al. 1992). Conventionally, radio quiet (RQ) AGNs are defined as  $R^* < 10$  while radio loud (RL) AGNs are defined as  $R^* > 10$ . It is estimated that  $\sim 10\%$  of AGNs are RL (Kellermann et al. 1989). Most radio galaxies are *compact* (i.e., unresolved at  $\sim 1''$  resolution), but it is estimated that  $\sim 10\%$  are *extended* sources (Lu et al. 2007). Due to the sensitivity limits of radio telescopes, most known radio galaxies are at low redshift ( $z \lesssim 1$ ). Radio galaxies can be optically faint or bright. X-ray emission is usually associated with strong radio emission from the jet (Worrall & Birkinshaw 2006; Gelbord et al. 2008). In general, for a given optical luminosity, the X-ray emission from RL quasars is about three times greater than that from their RQ counterparts (Zamorani et al. 1981; Worrall et al. 1987).

- Quasars. Quasars are the most luminous AGNs with  $L_{\text{bol}} \gtrsim 10^{46} \text{ erg s}^{-1}$ . In most UV and optical sky images, quasars are point-like sources. Spectroscopically, they exhibit similar features as Seyfert 1 galaxies with prominent Ly $\alpha$  emission lines and/or other broad emission lines, but with little contribution from host galaxies. The UV and optical spectra of quasars can also exhibit broad absorption lines (BALs) with FWHM between 2,000 km s $^{-1}$  and 25,000 km s $^{-1}$  (Greenstein & Schmidt 1964; Schmidt 1969; Weymann et al. 1991; Trump et al. 2006; Gibson et al. 2009). Besides BALs, there are also narrow absorption lines (NALs). NALs are produced by intervening absorbers such as molecular clouds, interstellar or intergalactic medium (e.g., Ding et al. 2003, 2005; Jones et al. 2010), or by intrinsic absorbers, which are physically associated with the AGN (e.g., Ganguly et al. 1999, 2001; Wu et al. 2010). Similar to Seyfert 2 galaxies, highly obscured quasars are called Type 2 quasars, which are characterized by narrow emission lines in the UV/optical spectra (Zakamska et al. 2003) and/or strong absorption in X-ray band.
- BL Lac objects. In general, AGNs show continuum and emission line variability at all wavelengths at which they have been observed (e.g., Matthews & Sandage 1963; Schmidt 1969; Uttley & Mchardy 2004), from X-ray to radio. A small fraction of AGNs, called BL Lac objects are distinguished by short time-scale variations that are abnormally large (e.g., Burbidge & Hewitt 1992; Wagner & Witzel 1995; Rani et al. 2010). Associated with this large variations are high polarizations (e.g., Fan et al. 2007; Smith et al. 2007), strong radio emission, and absence of emission lines in their UV/optical spectra. Many of them are RL and their emission is found to be relativistically beamed from the jets and non-thermal (Urry & Padovani 1995).
- LINERs (low-ionization nuclear emission regions). They are weak AGNs with bolometric luminosity less than  $10^{44} \text{ erg s}^{-1}$  (Heckman 1980). Seyferts are on average 10 times more luminous than LINERs (Ho 2003). The defining characteristics of LINERs are the relative intensities of their oxygen emission lines, namely  $[\text{O II}] \lambda 3727 / [\text{O III}] \lambda 5007 > 1$  and  $[\text{O I}] \lambda 6300 / [\text{O III}] \lambda 5007 > 1/3$  (Eracleous et al. 2010). Although the spectra of LINERs resembles Seyfert 2 galaxies and H II regions, the two classes can be distinguished by considering the intensity ratios of two pairs of lines, e.g.,  $[\text{O III}] \lambda 5007 / \text{H}\beta$  vs.  $[\text{N II}] \lambda 6583 / \text{H}\alpha$ , which is known

as the “BPT” diagram (Baldwin et al. 1981; Veilleux & Osterbrock 1987). Pioneering work estimated that at least 1/3 of all the spiral galaxies are LINERs (Ho 1996). The energy source powering LINERs remains a subject of debate. It could be a low-luminosity AGN (e.g., González-Martín et al. 2006; Flohic et al. 2006; Eracleous et al. 2010), massive star formation (e.g., Barth & Shields 2000), shock heating mechanism resulting from massive star formation (e.g., Dopita & Sutherland 1995), or some combination of the above.

- Ultra-luminous infrared galaxies (ULIRGs). These galaxies have extremely high infrared (IR) luminosity, typically  $L_{\text{IR}} > 10^{12} L_{\odot}$ . Spectroscopically, they exhibit both black hole accretion and starburst signatures. Current models of ULIRGs suggest that they are powered by both starburst and AGN activity triggered by collision of gas-rich galaxies (Sanders & Mirabel 1996; Lonsdale et al. 2006; Hou et al. 2009).

There are other sub-types of AGNs, such as broad line radio galaxies (BLRG, which is a subset of radio galaxies, see Leighly et al. 1997 for an example), broad line Seyfert 1 (e.g., Vanden Berk et al. 2006), steep spectrum radio quasar (SSRQ, see Bicknell et al. 1997 for an example) and flat spectrum radio quasar (FSRQ, which is a subset of blazars, see Sambruna et al. 1996). In this thesis, we focus our research on quasars and, if not specified, we mean RQ Type 1 quasars.

## 1.2 Unified Quasar Structure

It is believed that objects in the various classes above can be described by a unification model (e.g., Gopal-Krishna et al. 1994; Falcke et al. 1995; Antonucci 1993; Vagnetti & Spera 1994; Lister et al. 1994; Maraschi & Rovetti 1994; Urry & Padovani 1995; Kauffmann & Haehnelt 2000; Lacy 2003; Hopkins et al. 2006; Richards et al. 2011). In this scheme, different types of AGNs have the same (or similar) structure, but appear to be different because of orientation effect, i.e., observers are examining different regions from different viewing angles, and the existence of jets. In general, an AGN consists of the following components (Figure 1.2).

- A central black hole. It is believed that black holes exist at the centers of all massive galaxies (Ho 1999), and the black hole growth is deterministic to AGN structure and evolution (e.g., Marconi et al. 2004). The black hole is rotating (e.g., Wang et al. 2006; Nemmen et al. 2007; Ballantyne 2010) and is accreting surrounding materials. The black hole mass  $M_{\text{BH}}$  and the accretion rate  $\dot{M}_{\text{BH}}$  are two important parameters that drive AGN phenomena. The black hole mass ranges from  $10^6$  to  $10^{10} M_{\odot}$  (Urry 2004). In general, more luminous AGNs have more massive black holes (Blandford & McKee 1982; Peterson 1993; Peterson et al. 2004). Therefore, quasars host the most massive black holes with  $10^7 M_{\odot} < M_{\text{BH}} < 10^9 M_{\odot}$  (e.g., Hopkins & Hernquist 2009). Black holes of Seyfert galaxies have masses of  $M_{\text{BH}} \sim 10^6 M_{\odot}$  (e.g., Denney et al. 2010). If LINERs are powered by AGNs, they host the least massive black holes of the entire AGN family with  $M_{\text{BH}} \lesssim 10^6 M_{\odot}$ .

- An accretion disk. Material surrounding the central black hole, because of loss of angular momentum (Begelman 1985), forms a geometrically thin disk (Peterson 1997; Frank et al. 1992) around the black hole. The accretion disk is the main producer of UV photons (Treves et al. 1988; Armitage 2004). In addition, a corona of hot materials forms above the accretion disk (Poutanen 1998). UV photons generated from the accretion disk can be boosted to X-ray photons via inverse-Compton scattering inside the corona (Reynolds & Nowak 2003).

Because of the quasar outflow which originates from the accretion disk (the accretion disk wind; e.g., Gallagher et al. 2002b; Brandt & Kaspi 2002; Chartas et al. 2009a). Radiation from the accretion disk can thus be shielded or obscured by these outflowing gases, and line/continuum absorptions are seen in UV/optical and/or even X-ray spectra (e.g., Mathur et al. 1995; Reynolds 1997; Giustini et al. 2010; Brandt et al. 2000). These outflows provide feedbacks to AGN host galaxies during the growth of the SMBH in galactic bulges (e.g., Fabian 1999b; Saez et al. 2009). The narrow and broad absorption lines also provide important insights into the dynamics, structure, ionization and metallicity of outflows (e.g., Hamann 1997; Ganguly et al. 2001; Hamann et al. 2002; Hamann & Sabra 2004; Misawa et al. 2007; Ganguly & Brotherton 2008; Wu et al. 2010).

- Broad emission line region (BLR or BLR). The BLR surrounding the accretion disk is where the UV/optical broad emission lines in quasar spectra are generated (see reviews by Sulentic, Marziani & dultzin-Hacyan, 2000). The lines are broadened both because of both thermal and bulk motion of gases. The electron density of this region is typically  $10^{10}$ – $10^{11}$  cm $^{-3}$  (Ferland et al. 1992; Peterson 1997). Because the BLR electron densities are sufficiently high that virtually all forbidden lines are collisionally suppressed, there are no simple temperature and density diagnostics for the this region. It is traditionally believed that the BLR consists of discrete spherically distributed clouds (e.g., Urry 2004; Peterson 1997). However, recent quasar spectral studies of the double-peaked broad line profiles suggest flattening and inclination within the structure of the BLR (e.g., Eracleous 2004; La Mura et al. 2009; Gaskell 2009).
- Narrow line region (NLR). The NLR is outside the BLR and can extend far beyond the host galaxy. Ground-based observations (e.g., Pogge 1989) and subsequent images obtained with the *Hubble Space Telescope* (*HST*; Schmitt & Kinney 1996) reveal that the NLRs in many Seyferts exhibit a bi-conical structure (Kraemer et al. 2011). The electron density in this region is low enough that forbidden and semi-forbidden lines are produced.
- Jets. Jets are not seen in all AGNs. They are usually associated with RL objects, but even in RL quasars jets are not always present. Jets are believed to be perpendicular to the plane of the accretion disk and can extend from the black hole to well beyond the NLR. Jets are rarely visible in UV/optical band, but have strong radio and X-ray emission due to synchrotron radiation (Worrall 2009). Depending on jet shapes, radio quasars are classified as FR 1 and FR 2 (Fanaroff & Riley 1974). FR 1 sources are weaker radio sources which are brightest in the center, with

decreasing surface brightness towards the edges. In contrast, the more luminous FR II sources are limb-brightened, and often show regions of enhanced emission either at the edge of the radio structure or embedded within the structure. Radio jets from AGNs are thought to be an important form of AGN feedback, which can quench star formation and reduce the number of supernova whose progenitors are massive stars (e.g., Vernaleo & Reynolds 2006).

In addition, quasars have host galaxies. Quasars can provide feedback to their host galaxies in forms of outflows and affect their metallicity, star formation rates, and gas fractions. Observationally, host galaxies are usually blended with quasar emission in optical band, while in UV band the galaxy contribution is relatively less. In nearby Seyfert galaxies, AGN host galaxies are visible and could be resolved in some cases. Host galaxies of quasars are usually not visible because the emission from the nuclei vastly outshines the stellar light. In quasar spectra, narrow Ca II absorption lines are usually used to identify the contribution from host galaxies. (e.g., Bennert et al. 2011). The spectral decomposition technique can be used to separate host galaxy component from the AGN components (e.g., Vanden Berk et al. 2006; Shen et al. 2008a).

### 1.3 Quasar Surveys

There have been many multi-waveband surveys targeting at finding and observing quasars since they are discovered in the 1960s. We summarize three surveys that are related to this thesis study.

- 3C radio survey. The Third Cambridge Catalogue of Radio Sources (3C) is a compilation of radio sources detected originally at 159 MHz (Edge et al. 1959). It contains 471 sources with a flux limit of 8 Jansky (Jy). Subsequently, Bennett (1962) created a revised version of the 3C at 178 MHz, formally named as 3CR, which contains 328 sources with a flux limit of 9 Jy. This catalog was further revised by Laing et al. (1983) and 25% of the entries are optically identified as quasars. Subsequent similar radio surveys include PKS survey (Ekers 1969), 4C (Pilkington & Scott 1965; Gower et al. 1967), until 9C (Waldrum et al. 2003). Quasars in these “C” catalogs are, of course, strong radio sources.
- Sloan Digital Sky Survey (SDSS). The SDSS uses a dedicated 2.5-meter telescope located at Apache Point Observatory, New Mexico, to obtain images in five bands (*ugriz*; Fukugita et al. 1996) over approximately  $10,000 \text{ deg}^2$  of high Galactic latitude sky. The survey also obtains spectra from  $\sim 3800 \text{ \AA}$  to  $\sim 9200 \text{ \AA}$  at a spectral resolution of  $\simeq 2000$  (York et al. 2000). Between 2000 and 2008 (the Sloan Legacy Survey), the SDSS imaged over  $11,600 \text{ deg}^2$  of sky and obtained spectra of galaxies and quasars from  $9,380 \text{ deg}^2$  of that imaging. There are in total seven data releases (Stoughton et al. 2002; Abazajian et al. 2003, 2005; Adelman-McCarthy et al. 2006, 2007, 2008; Abazajian et al. 2009) and five editions of quasar catalogs (Schneider et al. 2002, 2003, 2005, 2007, 2010). The latest catalog contains 105,783 spectroscopically confirmed quasars with absolute luminosities  $M_i \leq -22.0$  and highly reliable redshifts.

- The Very Large Array (VLA) Faint Images of the Radio Sky at Twenty-cm (FIRST) Survey was designed to produce the radio equivalent of the Palomar Observatory Sky Survey over 9,900 deg<sup>2</sup> of the North Galactic Gap (Becker et al. 1995), corresponding to the sky regions covered by SDSS. The survey uses the National Radio Astronomy Observatory VLA in its B-configuration to acquire snapshots centered at  $\sim 1.4$  GHz. This survey, with flux density sensitivity of  $\sim 1$  mJy, is much deeper than the C series radio surveys. The FIRST radio data are used to select quasars in the SDSS quasar catalogs; unresolved objects brighter than  $i = 19.1$  that lie within 2.0'' are identified as primary quasar candidates (Richards et al. 2002b). The 20 cm peak flux density (encoded as an AB magnitude) is recorded in the SDSS data release 7 (DR7) quasar catalog (Schneider et al. 2010).

## 1.4 Unified Themes and Organizations of This Thesis

The quasar SED covers a wide range of the electro-magnetic spectrum (Wilkes 2004; Figure 1.1). The main theme of this thesis is to provide observational constraints of quasars in the rest-frame UV/optical and X-ray bands using both spectroscopic and photometric observations. These constraints include the continuum shape, (e.g., Vanden Berk et al. 2011, in preparation), spectral energy distribution (SED; e.g., Wu et al. 2011, in preparation), and important continuum and line correlations (e.g., Wu et al. 2009). These relationships provide deep insights into the structure and physical conditions of quasars (e.g., Wu et al. 2010).

### 1.4.1 UV/Optical Emission

A composite quasar spectrum from the SDSS is presented in Figure 1.3 (Vanden Berk et al. 2001). This spectrum represents the “canonical” shape of a quasar spectrum in UV/optical bands. It contains a UV power-law continuum from Ly $\alpha$  to  $\sim 5600$  Å, and an optical power-law from  $\sim 5600$  Å to  $\gtrsim 8000$  Å with a shallower slope. Through out this thesis, unless we specify the continuum waveband, we refer to the rest-frame UV continuum by default. This component is formed by the thermal emission on the accretion disk (Peterson 1997). The UV power-law shape reflects the thermal status of the accretion disk which is affected by the accretion process (Treves et al. 1988; Frank et al. 1992). The observed spectral slope  $\alpha_\lambda$  can be shallower or even positive due to reddening effects (e.g., Richards 2001; Richards et al. 2003). Changes of spectral slopes over redshift provide information on the quasar evolution, and changes of spectral slopes over luminosity reveal differential accretion process from faint to bright quasars as well as insights into the fundamental parameter that controls the accretion process.

AGN spectra usually contain a set of prominent narrow and broad emission lines. Permitted lines such as Ly $\alpha$ , C IV, Mg II, H $\alpha$  and H $\beta$  are usually broad with FWHM  $\gtrsim 1000$  km s<sup>-1</sup>. Semi-forbidden lines can be broad (e.g., C III] 1908) or narrow (e.g., Al II] 2669). Forbidden lines such as [Ne v 3346], [O III 5008] and [N II 6549] are usually narrow with FWHM  $\lesssim 1000$  km s<sup>-1</sup>. These lines cannot be formed in the BLR because the high density in this region well exceeds the critical densities.

Some broad emission lines, such as C IV, exhibit notable “emission line shifts”, i.e., blue-shift with respect to the quasar’s rest-frame (Gaskell 1982; Wilkes 1986; Espey et al. 1989; Zheng & Sulentic 1990; Corbin 1990; Weymann et al. 1991; Tytler & Fan 1992; Brotherton et al. 1994; Laor et al. 1995; McIntosh et al. 1999; Vanden Berk et al. 2001; Richards et al. 2002a). This phenomenon indicates that the broad emission line gases are involved or associated with the quasar outflow (Richards et al. 2002a, 2011). Narrow emission lines exhibit little or no shift because they are created at considerable distance from the accretion disk and are much less influenced by quasar outflow. Some lines, such as H $\beta$ , can be decomposed of a broad and a narrow component; their central wavelengths can be offset because they are formed in different regions. The correlation between emission line shifts and the ionization potentials of the ions (e.g., Tytler & Fan 1992; McIntosh et al. 1999; Vanden Berk et al. 2001) implies a stratified BLR structure.

In some cases, a small enhancement between 2000 Å and 4000 Å is seen in a quasar spectrum which is conventionally named the 3000 Å bump (Grandi 1982; Oke et al. 1984). The exact origin of this component is not well understood, and it should primarily be comprised of Balmer line complex and a thermal emission component it must be taken into consideration in spectral fitting.

Absorption, especially intrinsic absorption, is very useful in diagnosing physical conditions of outflowing materials driven by the ionization flux from the background continuum source as well as the broad line emitters (e.g., Hamann & Ferland 1999). Broad absorption lines ( $\text{FWHM} \gtrsim 500 \text{ km s}^{-1}$ ) are produced when the line-of-sight passes along the outflow and thus frequently have large ejection velocities. Strong broad absorption in UV is usually associated with X-ray weakness, which is in agreement with the models in which the strong X-ray absorption shields the outflow from over-ionization and enable radiative acceleration (e.g., Gibson et al. 2009). Narrow absorption lines (NAL) could be either intervening (e.g., Rauch 1998; Churchill et al. 2000; Charlton 2003; Ding et al. 2003) or intrinsic (e.g., Foltz et al. 1986; Hamann et al. 1997; Crenshaw et al. 1998; Misawa et al. 2007). If determined to be intrinsic, the NALs allow one to separate each individual absorption line doublet, and to build models that probe the physical conditions of intrinsic absorbers, including the outflow velocity, metallicity, and ionization status. For high redshift quasars  $z \sim 3$ , these modeling results are informative on the evolutionary phases of quasars in the early universe in terms of the star formation and accretion activity.

Although the extreme ultraviolet (EUV) and infrared (IR) are not the focus of this thesis, for completeness, we present a brief summary. The EUV region on the blue side of Ly $\alpha$  suffers from severe intrinsic (from quasar) and intervening (from interstellar medium) absorptions. It is extremely difficult to determine the true continuum due to these absorptions. The emission lines (e.g., O VI  $\lambda 1032$ ) are usually strongly affected by absorption and the intrinsic absorption lines are usually blended with Ly $\alpha$  forest lines. For  $z \gtrsim 6$  quasars, the absorption is so strong that it forms the “Gunn-Peterson Trough” (Gunn & Peterson 1965; also see Fan et al. 2001, 2003, 2004, 2006). The near infrared (IR) and mid IR spectra of AGNs (e.g., Weedman et al. 2005) are characterized by diversity, such as strong silicon absorption (Mrk 231, and Cen A), strong silicon emission (I ZW 1 and NGC 1275), and strong PAH emissions. The IR emission is

produced via the radiative transfer in the torus (Nenkova et al. 2008). The EUV and IR spectral regions are not the focus of the thesis.

### 1.4.2 X-ray Emission

Quasar X-ray emission is relatively simpler than the UV band. The spectrum features a power-law with an average power-law photon index  $\Gamma \approx 2$  (e.g., Young et al. 2009). This X-ray continuum is a result of inverse comptonization of UV photons produced from the inner part of the hot accretion disk (Longair 1994). At  $\sim 6.4$  keV is the Fe K $\alpha$  line (Reynolds & Nowak 2003); because of its high fluorescent yield and large cosmic abundance, it is the strongest emission line appearing in the X-ray band. Although intrinsically narrow, the iron lines are broadened by Doppler shifts and strong gravitational effects (Krolik & Kallman 1987; Fabian et al. 2000).

X-ray and UV emission for a given quasar are correlated (Tananbaum et al. 1979; Zamorani et al. 1981; Vignali et al. 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007). Namely, a non-BAL RQ quasar bright in UV (represented by flux at 2500 Å) is very likely to be bright in X-ray band as well (Gibson et al. 2008). The C IV and Fe K $\alpha$  line flux are also positively correlated (Wu et al. 2009).

In addition, there is a soft X-ray excess discovered below  $\sim 0.5$  keV (e.g., Arnaud et al. 1985; Page et al. 2004b), which is believed to be a thermal component at a constant temperature over a wide range of AGN properties (Crummy et al. 2006).

X-ray absorption, primarily caused by photoelectric effect, is almost always seen in X-ray spectra of quasars. This continuum absorption weakens the X-ray flux and causes a quick decay towards soft-X-ray band. Neutral hydrogen absorption causes an exponential decay in form of  $f_\nu \propto \nu^{1-\Gamma} \exp[-(h\nu/kT)^{-8/3}]$  (Longair 1992). In most cases, the Galactic absorption dominates, but intrinsic X-ray absorption can be significant (see Chapter 4).

X-ray and UV absorption are also correlated (e.g., Brandt et al. 2000). Particularly, BAL quasars are also commonly detected to be X-ray weak as a result of high intrinsic column densities ( $N_{\text{H}}$ ) typically in range of  $(1\text{--}50) \times 10^{22}$  cm $^{-2}$  (e.g., Gallagher et al. 2002a, 2006; Saez et al. 2009).

Because X-ray photons are produced closer to the black hole than UV photons, investigating quasars in X-ray allows us to examine the physical processes closer to the central engine than UV/optical radiation. The X-ray-UV connection indicates that the UV and X-ray emission regions should be considered jointly. If quasars are similar and their spectra are remarkably so scalable, there should be a few fundamental parameters, such as the black hole mass  $M_{\text{BH}}$  or the accretion rate  $\dot{M}_{\text{BH}}$ , that drive the entire central engine and many of the phenomena we observe.

### 1.4.3 Thesis Organization

The thesis is organized as follows. In the next chapter, we study the average quasar spectral properties as functions of redshift and luminosity using a set of composite quasar spectra. In Chapter 3, we focus on the Baldwin Effect using a sample of 272 RQ Type 1 AGNs and quasars. In Chapter 4, we utilize simultaneous UV/optical

and X-ray observations from *Swift* Gamma-ray Explorer (here after *Swift*) to create a quasar catalog. We attempt to constrain the bolometric luminosity and re-visit the  $\alpha_{\text{ox}} - l_\nu(2500 \text{ \AA})$  relation using a “clean” sample selected from this catalog. In Chapter 5, we turn our attention to absorption lines. By modeling intrinsic NAL systems of three quasars, we attempt to understand the physical conditions of the intrinsic absorbers of quasars at high redshift. Through out this paper, we adopt the same cosmology as the SDSS quasar catalog (Schneider et al. 2010), assuming an  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$ , and  $\Omega_\Lambda = 0.7$ .

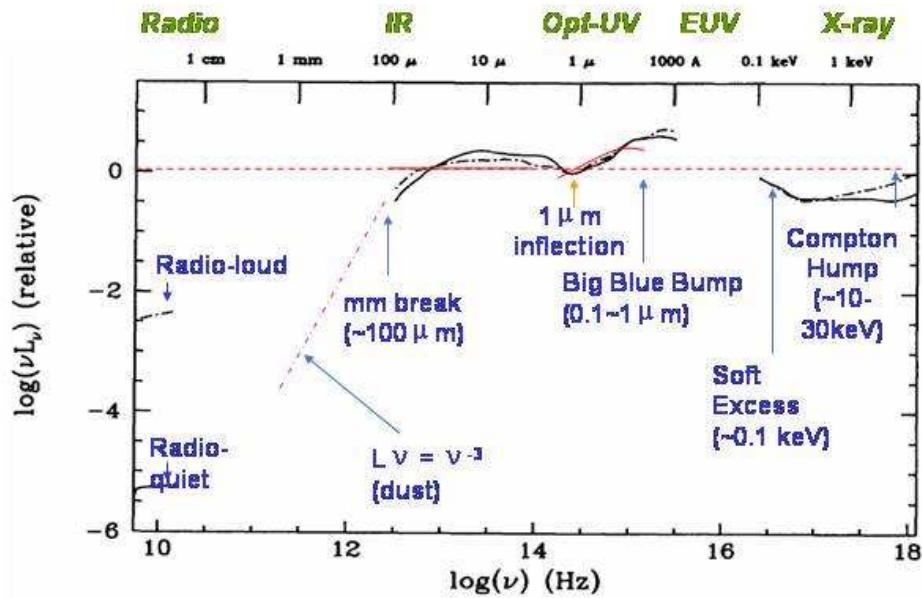


Fig. 1.1 A mean Type 1 RQ quasar SED (Elvis et al. 1994) covering the radio to the X-ray band with important features indicated.

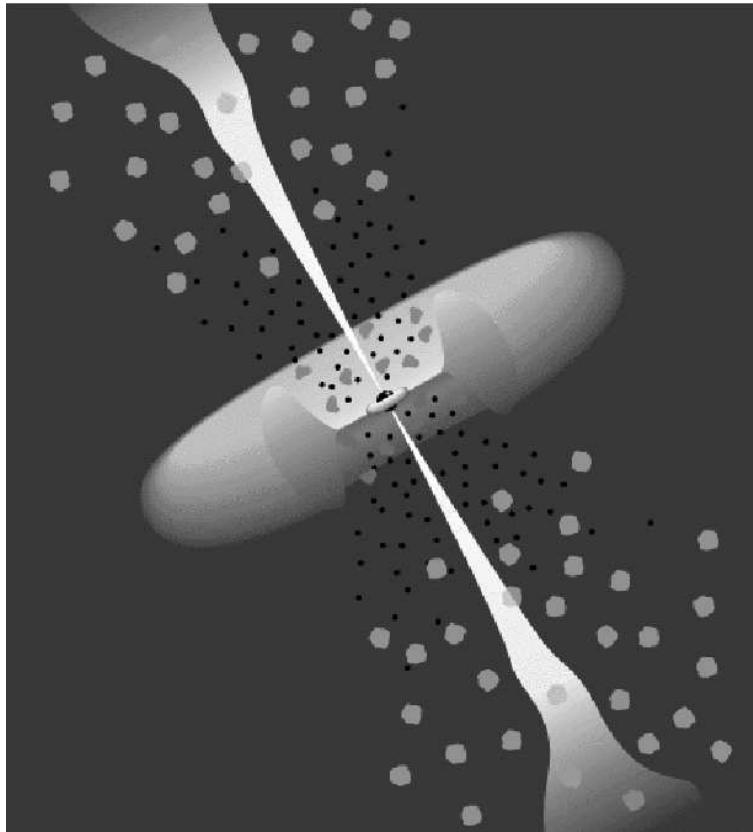


Fig. 1.2 Unification model of AGNs (Torres & Anchordoqui 2004). In this model, different types of AGNs are due to line-of-sight effects. Blazars are those AGNs for which the jets are close to line of sight. A regular quasar or a Seyfert 1 galaxy is observed if the orientation angle is  $\sim 30^\circ$ , where the narrow line and broad line regions are visible. At larger angular offsets, the BLR will be hidden by torus, the corresponding class being Seyfert 2 galaxies. Perpendicular to the jet axis, the full extent of the jets may be seen particular at low frequencies, giving rise to a morphology typical radio galaxies.

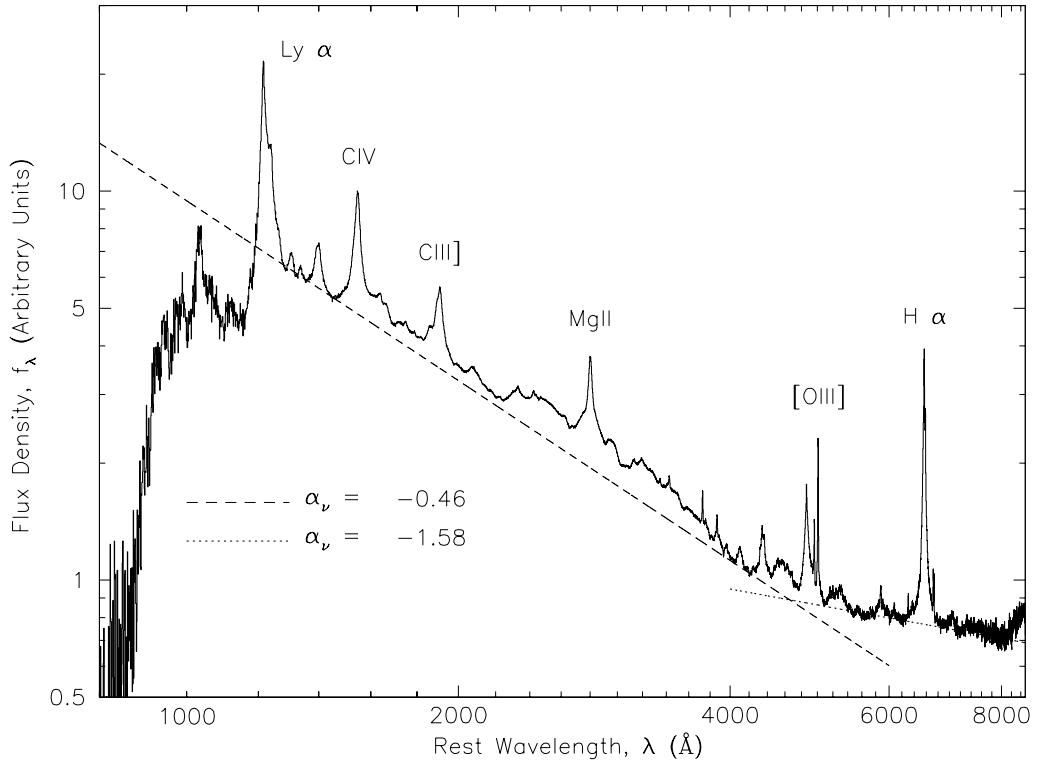


Fig. 1.3 The composite spectrum of quasars generated using  $\gtrsim 2000$  SDSS quasar spectra (Vanden Berk et al. 2001). The spectrum represents two power-law continuum, the iron emission forest and several prominent emission lines.

## Chapter 2

# Continuum and Emission Line Properties of Composite Quasar Spectra from SDSS

### 2.1 Introduction

Two of the most fundamental parameters for any class of extragalactic objects are redshift and luminosity. Redshift provides an estimate of the distances to the objects, as well as the age of the universe at the time the observed light was emitted. Luminosity is directly related to the energy generation mechanism of the objects. Empirically derived parameters from observations of the objects may change with either redshift or luminosity, providing information about the evolution of the population over cosmic time, and the changes that occur when the energy output varies.

There are several observable UV/optical quasar parameters that are known to correlate with luminosity. These include an anti-correlation of emission line equivalent width with ultraviolet continuum luminosity (known as the Baldwin effect; Baldwin 1977; see also, e.g., Wu et al. 2009; Richards et al. 2010), shifts of emission positions relative to systemic positions (e.g., Gaskell 1982; Tytler & Fan 1992; Richards et al. 2002a), UV spectral slope, optical/UV variability amplitude (e.g. Vanden Berk et al. 2004), and the relative amplitudes of eigenspectra reconstructions of quasar spectra (Yip et al. 2004). Many of these parameters appear observationally to be related to one another (e.g. Richards et al. 2011). Including X-ray observations, there is also an anti-correlation between the ultraviolet 2500Å continuum luminosity and its strength relative to the 2 keV X-ray luminosity – the so-called  $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$  relation (e.g. Avni & Tananbaum 1982; Just et al. 2007).

In contrast, a number of important quasar parameters, such as metallicity of the broad line region, do *not* appear to depend upon redshift (e.g. Hamann & Ferland 1999; Jiang et al. 2007; Juarez et al. 2009). However, as a population, the luminosity function of quasars evolves strongly over cosmic time (e.g. Schmidt 1968, 1970; Schmidt & Green 1983; Croom et al. 2009), and it may be expected that some global spectral properties of quasars evolve over time as well.

The large sample of quasars from the SDSS spans a wide range of luminosity and redshift, and allows the isolation of the dependence of quasar properties on both parameters. Statistical properties of quasars in small bins of luminosity and redshift can be determined with high precision. In this study, we examine a number of ensemble quasar properties, including spectral slope, emission line equivalent width, and emission line shifts, as a function of both luminosity and redshift.

In this work, we construct a set of composite quasar spectra from the Sloan Digital Sky Survey (SDSS; York et al. 2000) Data Release 3 quasar catalog (DR3; Schneider et al. 2005) to study the changes of spectral parameters as functions of redshift and

luminosity. We follow the method to construct the composite spectra as described by Vanden Berk et al. (2001), who constructed only a single quasar spectrum based on 2200 spectra from SDSS early data release (Stoughton et al. 2002).

The SDSS used a dedicated 2.5-meter telescope located at Apache Point Observatory, New Mexico, to obtain images in five bands (*ugriz*; Fukugita et al. 1996) over approximately 10,000 deg<sup>2</sup> of high Galactic latitude sky. The survey also obtains spectra from  $\sim 3800 \text{ \AA}$  to  $\sim 9200 \text{ \AA}$  at a spectral resolution of  $\approx 2000$  (York et al. 2000). The DR3 quasar catalog consists of 46,420 objects in the SDSS DR3 that have luminosities larger than  $M_i = -22$ . The quasar redshifts range from 0.08 to 5.41. The rest-frame wavelength in the SDSS spectra ranges from  $\sim 610 \text{ \AA}$  to  $\sim 9000 \text{ \AA}$ . These data were used to construct 186 composite spectra.

An updated work, which is based on the SDSS DR7 quasar catalog (Schneider et al. 2010), is in progress. Because the number of quasars in DR7 is over twice that in DR3, the number of quasars per composite increases (see the new luminosity-redshift binning diagram in Figure 2.3). This improves the signal-to-noise ratio ( $S/N$ ) for each composite spectrum. Although the total number of bins increases as well, this does not provide a significant number of new bins with high-quality spectra. Most new quasar spectra are of intermediate luminosity and fractions of faint and very luminous quasars are small. In addition, bins with very few (typically less than 20) quasar spectra are dropped because of low  $S/N$ . As a result, the new quasar catalog does not significantly extend the redshift and luminosity ranges. However, a preliminary study of the DR7 composite spectra indicates that the general trends of spectral parameter do not change. Therefore, in this work, we present our work based on the DR3 quasar catalog. The work based on the DR7 quasar catalog will form the basis of the published study.

This paper is organized as follows. In Section 2.2, we introduce the methods and procedures to generate the composite spectrum set; in Section 2.3, we fit the quasar spectra using our custom software package *QSFIT*; in Section 2.4, we study the spectral properties in terms of spectral parameters as functions of luminosity and redshift; In Section 2.6, we present our conclusions. Throughout this paper, we apply the following cosmological parameters:  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 2.2 Generating the Composite Spectra

The quasar sample is drawn from the SDSS DR3 quasar catalog (Schneider et al. 2007) and the spectra of those quasars were taken from the SDSS data release 5 (DR5) database. Not all of the quasars from the catalog should be used for this study, since the continuum shapes and emission lines of the sample can depend upon quantities other than redshift and luminosity, but those quantities may not be representatively sampled due to observing limitations. For example, quasars selected *solely* by their X-ray flux in the ROSAT All Sky Survey (Voges et al. 1999, 2000) should not be included, because those quasars are detected almost exclusively at low redshift, and their high-redshift counterparts are not represented in the SDSS quasar catalog. Any unique redshift or luminosity dependence of quasar spectral properties in such X-ray selected quasars could impart artificial trends on the larger sample. Therefore, only quasars selected as candidates by their optical colors were included in the sample. X-ray or radio selected

quasars were included only if they could also have been selected by their optical colors (see Richards et al. 2002b for description of SDSS quasar target selection).

Quasars with broad absorption line systems (BALQSOs) were also excluded from the sample. BALQSOs are statistically redder than their non-BAL counterparts (e.g. Weymann et al. 1991), likely due to reddening by the absorber material (e.g. Reichard et al. 2003). Also, measurement of the continuum luminosities of BAL quasars is problematic because the continuum is often heavily attenuated by absorption lines (e.g. Hall et al. 2002). All quasars identified in the SDSS DR5 BALQSO catalog of Gibson et al. (2009) were therefore excluded from the sample. Applying the X-ray/radio/BALQSO exclusion criteria to the DR3 quasar catalog produced of sample of 44,969 that was used to construct our composite spectra.

The spectra of individual quasars were shifted into their rest frames by dividing the observed wavelengths by  $1 + z$ , then were re-binned onto a common wavelength scale by extending the SDSS wavelength scale to shorter wavelengths. The individual spectral pixels of rest wavelength and flux density  $f_\lambda$  were re-binned onto the common wavelength scale using cubic spline interpolation. The same procedure was applied to the error arrays of each spectrum.

Quasar redshift is not an unambiguous quantity, because the value of the redshift frequently depends upon the emission line used for the measurement (e.g., Richards et al. 2002a). The redshift differences imply relative emission line velocity shifts that can be as high as hundreds of  $\text{km s}^{-1}$ . One of the goals of this study is to determine how relative velocity shifts are related to other quantities. Therefore, we have measured multiple redshifts for each quasar based on a number of strong lines: H $\beta$ , Mg II  $\lambda 2798$ , and C IV  $\lambda 1549$ . The SDSS spectrum of every quasar with a redshift less than 4.9 will cover at least one of those lines. The wavelength at the peak of an emission line was determined by fitting a region around the line with a set of quasar eigenspectra generated by Yip et al. (2004). The wavelength at the peak flux density of the fit was taken to be the wavelength of the emission line, and used to determine the emission line redshift. When the composite spectra were constructed, for each set of spectra to be combined, only the redshifts based upon the same single emission line were used. The H $\beta$  redshifts were used for  $0.0000 < z \leq 0.7378$ , the Mg II redshifts were used for  $0.7378 < z \leq 1.7542$ , and the C IV redshifts were used for  $1.7542 < z \leq 4.2481$ . The use of a single emission line for redshifts allows the emission line velocity shifts of every other line in the composite spectrum to be determined relative to the same line.

Recent work by Hewett & Wild (2010) systematically investigate the relationship between different redshift estimation schemes for more than 91,000 quasars in the SDSS DR6 (Adelman-McCarthy et al. 2008). Using a cross-correlation technique with a master-quasar template, they achieved a redshift accuracy down of  $\Delta z/(1 + z) \leq 10^{-4}$  ( $30 \text{ km s}^{-1}$ ) per unit redshift. They presented the consistency of redshifts using their methods and the redshifts determined by H $\beta$  and Mg II  $\lambda 2798$  centroid, and the systematic redshift biases when measuring redshifts using C IV  $\lambda 1549$  (e.g., Figure 9 of their paper). The emission line shift study of our work focuses on the *relative* shifts with respect to reference lines. The redshift values of our composites are based on the eigen-spectrum technique and can achieve at least comparable accuracy with that of the SDSS spectroscopic pipeline ( $\Delta z/(1 + z) \lesssim 0.002$ ), or even better, which is sufficient to study

emission line shift trends. For quasars with redshifts greater than 4.25, the SDSS quasar catalog redshift was used when generating the composite spectra. In most of those high-redshift cases, the C IV emission line is either not covered by the spectrum, very weak, or contaminated by night sky line residuals. The high-redshift spectra usually also lack any other strong and unabsorbed emission lines, so emission line shifts measured from the composite spectra of those high-redshift quasars will be quite uncertain.

In order to compare the properties of quasars as a function of luminosity, a consistent measurement of luminosity is required. This is difficult to achieve for quasars, since the continuum is heavily contaminated with emission lines at most wavelengths, and a fixed observed frame wavelength window covers different rest frame wavelength ranges at different redshifts. We have selected the specific luminosity at 2200 Å as the common luminosity measure for our study for the following reasons: 1) the rest frame 2200 Å wavelength is directly accessible in almost 80% of the quasar spectra. 2) the region near 2200 Å is relatively free of contaminating emission lines – at this wavelength blended iron emission lines contributes only approximately 5% to the total quasar specific luminosity. 3) 2200 Å is close to 2500 Å, which is a wavelength at which quasar specific luminosities are often measured (e.g., Vignali et al. 2003; Strateva et al. 2005), although at this wavelength the continuum is sufficiently more heavily contaminated by emission lines.

When accessible, the flux density at the rest frame 2200 Å wavelength was measured directly. For each spectrum this was performed by finding the median flux density value of the dereddened spectrum for all the pixels within the wavelength range 2160–2240 Å. When the 2200 Å region was not accessible, scaled measurements of the specific luminosity at other continuum wavelengths were used instead. The scale factors were determined by finding the average ratio of the specific luminosities at a set of continuum wavelength to the specific luminosities at 2200 Å, determined from quasars for which both continuum regions were accessible. The other continuum wavelengths used were 5100 Å for redshifts less than 0.77, 1450 Å for redshifts between 3.09 and 5.13, and 1350 Å for redshifts greater than 5.13. The specific luminosity scale factors in each case are 4.427 ( $\lambda = 5100 \text{ \AA}$ ), 0.542 ( $\lambda = 1450 \text{ \AA}$ ), and 0.481 ( $\lambda = 1350 \text{ \AA}$ ).

In Figure 2.1, we plot the 2200 Å luminosity vs. monochromatic power-law continuum luminosities at 5100 Å, 3000 Å, and 1450 Å for all composite spectra. The panel at bottom right plots  $l_\lambda$  (3000 Å) vs.  $l_\lambda$  (5100 Å), which are also strongly correlated. These wavelengths are often used in black hole mass estimators as benchmark wavelengths to calculate bolometric luminosities (e.g., Shen et al. 2008b). The tight linear correlations in all panels indicate that it is reasonable to use these luminosities to represent quasar brightness.

The sample we used to generate the composite spectrum set is comprised of 44,969 quasars. The redshift of the sample ranges from 0.06 to  $\sim 5.41$  and the luminosity range is  $39.30 < \log l_\lambda < 44.26^1$ . All of the spectra, including the flux density and error array components, were corrected for Galactic extinction using the reddening map constructed by Schlegel et al. (1998) and the average Milky Way extinction curve described by Fitzpatrick (1999).

---

<sup>1</sup>In the following paragraphs, unless specified, we use  $\log l_\lambda$  to represent  $\log l_\lambda(2200 \text{ \AA})$ .

The quasar sample was divided into small redshift and luminosity bins to construct the composite spectra. The redshift bin limits were set at equal logarithmic intervals,  $\Delta \log(1 + z) = 0.04$ , beginning at  $z = 0$ . The logarithmic spacing makes each redshift interval roughly equivalent to a constant interval of cosmological time up to the highest redshifts in the sample. The luminosity density bin limits were set at equal intervals of  $\Delta \log l_\lambda = 0.25$ . The redshift and luminosity bin sizes were selected to be small enough that there is little difference in the composite spectral properties in adjacent bins, but sufficiently large to contain a statistically meaningful number of spectra in bins at the sample limits. A plot of the number of quasars in each of the redshift-luminosity bins is shown in Fig. 2.2; 186 bins are populated (Figure 2.3 shows the distribution for the DR7 quasars).

The spectra of the quasars in each redshift-luminosity bin were combined to form a single composite spectrum for each bin. There are a number of ways of combining spectra into composites, depending upon their intended use. To investigate the continuum spectral shape of quasars, the geometric mean is an appropriate statistic as the continuum spectra of most quasars is well-described by a power law with an index particular to each quasar. The geometric mean of a sample of power law spectra will result in a power law spectrum with the mean index of the sample. This result will not necessarily be true for the arithmetic mean of the same sample. For analysis of composite emission lines, the arithmetic mean of the sample preserves the average flux and equivalent width of the emission lines in the sample. Both geometric mean and arithmetic mean composite spectra were constructed for each redshift-luminosity bin.

The SDSS spectroscopic pipeline flags spectroscopic pixels that are potentially unreliable (Stoughton et al. 2002). The flags are given as a binary mask array in the primary header and data unit of the spectra. Some flags indicate more serious problems than others. By inspection of a large number of individual spectra, and examination of results of composite spectrum construction, we have identified a set of flags that should cause the rejection of a pixel for the construction of composite spectra. These flags are (each preceded by ‘SP\\_MASK\\_’): NOPLUG, BADTRACE, BADFLAT, BADARC, LOWFLAT, FULLREJECT, SCATLIGHT, BRIGHTSKY, NODATA, COMBINEREJ, BADFLUX-FACTOR, BADSKYCHI, and REDMONSTER. In addition, pixels were masked from composite construction if their flux density values are unrealistically negative, defined here as  $f_\lambda < -5\sigma_\lambda$ ; most of these pixels are also rejected based on the mask array flags.

Pixels at the location of strong, narrow absorption lines — except those in the Ly $\alpha$  forest — were also excluded, since these locations do not represent the underlying flux of the quasar itself. Correction for the effects of the Ly $\alpha$  forest have not been made in this study, so any attempt to measure quasar intrinsic properties from the composite spectra below 1216 Å is unreliable. Absorption line features were identified by finding sets of at least three contiguous pixels that fall below the local continuum of a quasar by at least the 95% confidence level. The continuum of a spectrum was estimated by smoothing the spectrum with a low-pass Fourier transform filter between the peaks of the major emission lines. The 95% lower confidence limit was found from the distribution of the flux density differences.

This procedure is intended to be fast (since it must be applied to all of the spectra), but it is not designed to identify all of the statistically significant absorption features.

Comparison between the results of this procedure and those of a much more thorough algorithm show that the vast majority of the obvious absorption features are detected by our adopted method.

Host galaxy spectra were *not* removed from individual spectra before combining. This approach was adopted because the host galaxy estimates are less reliable for spectra with lower  $S/N$ , such as is the case with many individual spectra. Eigenspectrum fitting is sufficiently reliable to estimate the luminosity bin of individual quasars, but often lacks the required accuracy for determining the overall host galaxy spectrum shapes (Shen et al. 2008a). The resulting  $S/N$  of the composites is usually high enough that a reasonable mean host galaxy spectrum can be estimated and removed, at least at redshifts of  $z < 0.7$  (Shen et al. 2008a). Retaining the host galaxy components in the individual spectra will introduce no problems with the arithmetic mean composites, as the composite spectra will be composed of the arithmetic mean spectra of both the quasars and their host galaxies. The geometric mean composites, however, may be distorted in unpredictable ways with the inclusion of host galaxy components. For example, if the same host galaxy spectrum were added to every quasar spectrum in a sample, the geometric mean of the sums would be larger than the sum of the host galaxy and the geometric mean of the quasar spectra (Hoehn & Niven 1985). Caution should be taken then when interpreting the spectral slopes of geometric mean composites in bins where the host galaxy component is likely to be significant. We will highlight such cases in the analysis of the composite spectra.

While constructing the composite spectra, individual spectra were scaled so that the flux density of a quasar best matched that of the running composite spectrum in the overlapping wavelength regions. Because the logarithmic luminosity density bin widths are set to 0.25, the scale factor for an individual spectrum can have a maximum value of 1.78. Because of the relatively small redshift bin widths, the overlapping wavelength region will cover at least 93% of the wavelengths range of an individual spectrum.

The individual spectra can be weighted, e.g. by the inverse variance at each pixel, in order to improve  $S/N$  of the composite. However, this approach may overly weight the higher luminosity quasars in each bin, as they will tend to have spectra of higher  $S/N$  than their lower luminosity counterparts. The effect may be negligible in our analysis because all of the quasars in a given bin have approximately the same luminosity, and the higher  $S/N$  is advantageous in some cases. We have generated both weighted and unweighted composite spectra for each bin for both the geometric and arithmetic composites.

In the unweighted spectrum case, arithmetic mean and geometric mean values of the scaled flux densities were calculated at each rest wavelength pixel. The geometric mean composite spectra were generated by constructing the arithmetic means of the logarithms of the luminosity density spectra. Because the geometric mean of sequences that contain negative values is not defined, pixels with negative flux density values were masked.

## 2.3 Spectral Fitting

We have constructed four sets of composite spectra corresponding to two ways of averaging spectra (geometric and arithmetic) and whether the averaging process is weighted or unweighted. As proved by Vanden Berk et al. (2001), the geometric mean spectra preserves the spectral slope. Therefore, in this work we analyze the 186 geometric weighted mean composite spectra in the rest-frames with redshift ranging from 0.000 to 4.754 and logarithmic luminosity densities ( $\log l_\lambda$ , in  $\text{erg s}^{-1} \text{\AA}^{-1}$ ) ranging from 38.25 to 44.00. There are 37 composites that were created from less than ten individual spectra (see Figure 2.2). These composites, while available for public use, are not included in our analysis. In the following discussion, the 149 composites formed from at least ten spectra (blue and green boxes in Figure 2.2) were used.

The underlying continuum of each spectrum is fit using a two-component power-law (here after *double power-law*), the small Balmer bump, and the iron emission forest in the UV and optical bands. After this continuum is subtracted from the composite spectrum, the emission lines are fit using single or multiple Gaussian profiles. We developed a software package in IDL, named *QSFIT*, to perform the spectral fitting process. In the subsections below, we describe the algorithms used in this package. The architecture and technical details are described in the software manual (see Appendix B).

### 2.3.1 Continuum Subtraction

The major component of the underlying continuum in quasar is the power-law, defined here as  $f_\lambda = 10^\beta \lambda^{\alpha_\lambda}$ . For quasar spectra with  $z < 0.7$ , a two component power-law is required (see Figure 3 of Vanden Berk et al. 2001) in the UV and optical bands. The two power-laws have different indices and scaling factors and intersect at  $\approx 5600 \text{\AA}$  in most cases. The redshifts of over 67% spectra in our sample are greater than 0.7 where only a UV power-law is needed.

The second continuum component is the small Balmer, or  $3000 \text{\AA}$ , bump (Grandi 1982; Oke et al. 1984) between  $2000 \text{\AA}$  and  $4000 \text{\AA}$ . A definitive conclusion on the origin of this feature has not yet been reached. In general, this feature is believed to be produced by a combination of a low ionization iron emission (Fe II and Fe III) forest, the lower order Balmer lines, and Balmer continuum (Oke et al. 1984; Wills et al. 1985). It has been shown that the shape of the Balmer continuum can be fitted by either a small-optical-depth, high-temperature continuum or a  $T \sim 10000 \text{ K}$ ,  $\tau \sim 1\text{--}2$  continuum (Wills et al. 1985). We follow the fitting model adopted in Vanden Berk et al. (2001) for this continuum component. This feature has a continuous shape but a connection point at  $\lambda_B = 3663 \text{\AA}$ . The component with  $\lambda < \lambda_B$  represents an attenuated black body spectrum and the  $\lambda \leq \lambda_B$  is an exponential decay (e.g., Grandi 1982).

$$f_\lambda = \begin{cases} A_B \cdot \frac{1 - e^{\tau_B}}{e^{hc/(k\lambda_B T_B)} - 1} \cdot e^{-(\lambda - \lambda_B)/100} & , \quad \lambda \geq \lambda_B = 3663 \text{\AA} \\ A_B \cdot \frac{1 - e^\tau}{e^{hc/(k\lambda T_B)} - 1} \cdot \left(\frac{\lambda}{\lambda_B}\right)^{-5} & , \quad \lambda \leq \lambda_B, \tau = \tau_B \left(\frac{\lambda}{\lambda_B}\right)^3 \end{cases} \quad (2.1)$$

When  $\lambda \geq \lambda_B$ ,  $f_\lambda$  is an exponential decay which simulates the lower order Balmer line forest emission. The expression when  $\lambda \leq \lambda_B$  is a modified black-body spectrum. We fix the value of  $T_B = 10^4$  K and  $\tau_B = 1$ , leaving the scale factor  $A_B$  as the only free parameter. A typical 3000 Å bump profile is presented in Figure 2.4.

The third contributor to the continuum is the low ionization iron emission forest, which is far more complicated than the previous two components. Wills et al. (1985) simulated the Fe II emission using photoionization modeling of  $\sim 3000$  individual lines. Although we attempted to fit the iron emission forest using individual lines by a generic algorithm, it is very difficult to converge to a stable solution, i.e., there are many possible solutions which can produce equally good spectral fits. Because it is almost impossible to handle each individual iron emission profile, we decide to use iron templates, which are widely employed in quasar spectral fitting (e.g., Trump et al. 2009; Wu et al. 2009; Shen et al. 2010). We adopt the UV iron emission template from Vestergaard & Wilkes (2001) ranging from 1075 – 3090 Å, and the optical iron emission template from Véron-Cetty et al. (2004) that covers from 3535 – 7534 Å. We convolve each template with Gaussian profiles and create a series of 36 iron emission spectra with velocity dispersions ranging from 1,250 – 10,000 km s<sup>-1</sup> spaced at 250 km s<sup>-1</sup>.

The iron emission forest is represented as

$$f_{\lambda,uv} = A_{uv} N_{uv} f_{\lambda,uv}(\sigma_{uv}) \quad (2.2)$$

$$f_{\lambda,o} = A_o N_o f_{\lambda,o}(\sigma_o) \quad (2.3)$$

in which the subscripts “uv” and “o” represent “UV” and “optical” regions. The overall normalization factors for the iron templates are  $N_{uv} = 10^{16}$  and  $N_o = 1$ . The quantities  $\sigma_{uv}$  and  $\sigma_o$  are velocity dispersions of UV and optical templates, respectively. By default, the entire UV (optical) template has a universal scale factor  $A_{uv}$  ( $A_o$ ). However, either template can be subdivided into a number of sections during the manual/automatic fitting processes so that the user/program can individually set the scaling factors for each section of the iron template. We do not change the velocity dispersion for each section of the UV (or optical) template because in most cases, we can find acceptable fits using the same value of the dispersion. Visual inspection of the fits revealed a few cases where the iron complex was clearly redshifted or blueshifted. For these composites, a small offset ( $\Delta z \sim 0.001$ ) was applied to the iron template in the fitting process.

We first estimate the power-law component by connecting two relatively “line-free” windows tabulated in Table 2.1. Each of these regions is assigned two attributes: a priority and a decremental factor. Wavelength windows with higher priority are preferentially used because these regions are relatively less contaminated by emission lines. When these high priority wavelength windows are not covered in the spectrum, we use wavelength windows with lower priorities, which may reduce power-law estimation accuracy. A fine-tuning of decremental factors in the next step can potentially helpful to increase the accuracy. A “decremental factor” is used to “correct” the emission line contamination in these “line-free” windows; a default value is 0.97 was chosen.

When estimating a power-law, we first select the wavelength windows covered by a spectrum. If more than two windows are selected, the program generates a power-law for each pair of windows and ranks each power-law fitting based on goodness-of-fits,

which is the reduced  $\chi^2$  evaluated for all pixels in all “line-free” windows. This ranking process excludes “line-free” windows which are so close in wavelength that the power-law fluxes predicted at other “line-free” windows significantly deviate from observed fluxes. This process selects a pair of “line-free” windows which produces the best power-law continuum among all possible “line-free” window combinations.

After the optimal window pair is selected, the program adaptively adjusts the decremental factors for each window until three conditions are satisfied: (1) mean flux densities at central wavelengths of all emission lines are above the power-law component; (2) mean flux densities at all line-free windows are above the power-law component; (3) mean flux densities at all line-free windows are no more than  $3\sigma$  above than the power-law component. The mean flux is evaluated by averaging flux densities of  $\pm 2.5$  Å centered around the emission line central wavelengths. The optical power-law ( $\lambda > \lambda_p$ ) is usually much easier to fit than the UV counterpart. Connecting two “line-free” windows usually gives a good power-law estimation.

For composites formed from relatively few spectra (e.g.,  $N < 20$ ), the composite may exhibit absorption features in the line-free regions. We have not identified an effective method to detect and avoid these troughs automatically, so we visually inspect the power-law estimation results and customize the windows to estimate the power-law for these spectra.

In the next phase, we perform a joint fitting process by considering all the three components. This process can be performed either interactively or automatically using *QSFIT*. Our method first automatically fits the spectra using the Levenberg-Marquardt least-squares technique (More & Wright 1993) and then allows manually adjustment of the parameters. The fitting parameters, their initial values, and their limits are tabulated in Table 2.3. To evaluate the goodness-of-fit, we compare the original spectra and the synthesized continua within a set of benchmark wavelengths (Table 2.2). Finally, the continuum complex is subtracted from the original spectra, and the residuals are used for emission line measurements. Figure 2.5 shows two examples of the underlying continuum fitting results.

The uncertainties in this fit arise from the following sources:

1. Original observational random uncertainties that propagate to the composite spectra. This effect is especially conspicuous for pixels with less than 15 contributing quasars.
2. Redshift uncertainties caused by using different emission lines to calculate redshifts.
3. The iron templates. For example, the UV iron template is extracted from the narrow line Seyfert galaxy 1 ZW 1 (Vestergaard & Wilkes 2001); it is unlikely that this iron emission template represents the iron emission for all quasars.
4. The model for the 3000 Å bump, which is not well parameterized.
5. Our model to fit the spectra. Although a global power-law provides a reasonable match to the general shape of the quasar spectrum in the UV band, the situation in reality is undoubtedly more complicated, particularly at the intersection of the UV and optical power-law components. In addition, not all the quasar UV spectra

can be well fit by a single power-law. As we will see below we must occasionally use a local power-law when fitting an emission line.

### 2.3.2 Emission Lines

We fit the each emission line by a series of Gaussian profiles. The number of profiles used to fit an emission line is empirical, and is dependent on the goodness-of-fit. If an emission line is relatively separated or weakly blended with other lines, we fit it independently; otherwise, we fit the blended lines simultaneously. If the *global* continuum is obviously below or above the local continuum level around a certain emission line or an emission line group, we measure the line flux after subtracting the total flux around this emission line (group) with a *local* power-law continuum. We then apply a low pass band filter to smooth the spectrum around the line fitting region. We recognize the existence of an emission line if the peak signal strength around the emission line rest-frame wavelength is stronger than three times of the amplitude of the spectral noise at the peak signal pixel.

We fit an emission line using a single or multiple Gaussian profiles using the Levenberg-Marquardt least-squares fit. Generally, a broad emission line is fit using a broad component at the base and a narrow component for the core. However, some lines require three Gaussian profiles, e.g., H $\beta$ . Most narrow lines are well fit by a single Gaussian profile. If a weak line is blended with a strong line, the fitting procedure may neglect the weak line and treat it as noise, i.e., the weak line is not fit essentially. To preserve this weak line, we must reduce the degree of freedom by sample a piece of spectrum around the weak line peak and fitting these points with a single Gaussian to estimate of the weak line dispersion. We then fix the dispersion a constraint when performing joint fitting with the strong line. If a strong broad emission line is blended with multiple weak narrow emission lines, we mask these weak narrow emission lines and focus on fitting the strong lines first. We then fit the weak narrow lines after subtracting the broad line contribution. For example, the H $\alpha$  line is blended with two [N II] lines on either side close to its peak, and two [S II] lines beyond its red wing. In this situation, we first mask [N II] and [S II] lines and fit H $\alpha$ . Then, we subtract H $\alpha$  and fit the weak lines we masked. In some cases, narrow absorption features can appear in a broad emission line; when this occurs, we manually mask the absorption line region.

Table 2.4 lists all the emission lines we fit. The last column gives the default type of continuum subtracted. We present some emission line fitting examples Figure 2.6.

### 2.3.3 Error Analyses

The uncertainties of both continuum and emission line fitting parameters are estimated by artificially adding noise to the model spectra obtained in spectral fitting process. The noise amplitude for each pixel in the simulations follows a normal distribution, with the width of the normal distribution at each pixel equal the flux density uncertainty at that pixel. For each model spectrum, we artificially produce 100 noisy spectra and fit each one automatically in the same manner we fit the original composite.

The statistical uncertainty of each spectral parameter is then measured from its distribution (usually a reasonable facsimile Gaussian). This method counts only random uncertainties, which are dependent on the composite  $S/N$ .

### 2.3.4 Sensitivity of Fitting Results to Initial Guess Values

We use the Levenberg-Marquardt algorithm (LMA) (Levenberg 1944; Marquardt 1963; Gill & Murray 1978) in both of underlying continuum and emission line fitting processes. The LMA interpolates between the Gauss-Newton algorithm (Björck 1996) and the method of gradient descent (Avriel 2003). The LMA is more robust than GNA, which means that in many cases it finds a solution even if the initial guess is far off the final minimum. This algorithm has been programmed in many programming languages and is part of the Curve Fitting and Function Minimization package of the Markwardt IDL Library<sup>2</sup>. The LMA method has been used in quasar spectral fitting in Shen et al. (2008a). In our continuum fitting process, we found that the fitting results based on this algorithm changes if we use a different set of initial guess values. To investigate the initial guess impact on the final fitting results, we vary the initial guess parameters and explore how the final results vary correspondingly. We perform this experiments to spectra at different redshifts, and we found that the variation of continuum fitting results corresponding to initial value changes is negligible or not significant. The results of emission line fitting results are even more stable since the fitting process uses analytical expressions (multiple Gaussian profiles) as templates. A representative case is presented in this section using a quasar spectrum at  $z = 1.089$  and  $\log l_\lambda = 42.00$ . This spectrum covers a wavelength range between 1656 Å and 4420 Å in its rest-frame, which includes both of UV and optical iron templates and the small blue bump feature. There are five parameters that are the most important in fitting the underlying continuum shape (see Table 2.3),  $\alpha_1$ ,  $\beta_1$ ,  $A_B$ ,  $A_{Fe,u}$  and  $A_{Fe,o}$ . The initial values of these parameters are tabulated in Table 2.5. In each experiment, I change the initial guess of one of the five parameters and obtain the fitting results of all five parameters. The fitting outcomes are presented in Figures 2.7, 2.8, 2.9, 2.10 and, 2.11.

From these plots, we can see that although changes of initial values cause changes of fitting results, in most of cases, the change is small and negligible. For example, a change of  $\alpha_1$  by  $-0.1$  only leads to a change of its fitting results by about 0.01. In my experiments, the iron template scale factors influence the fitting results the most (Figures 2.10 and 2.11), but in these cases, at least two parameters (e.g.,  $A_{Fe,u}$ ,  $A_{Fe,o}$  and  $A_B$ ) change in the opposite way so that the overall spectral shape does not exhibit significant change.

The experiments above indicate that it is in general safe to use the initial parameters in Table 2.3 to perform continuum spectral fitting. In contrast to LMA, the evolutionary algorithm provides a most robust and stable method to solve the multi-variable optimization problem. With proper setups, it can find the global minimum and the fitting results are not sensitive to initial parameter values. An attempt to fit quasar continuum spectrum using this method is presented in Appendix A.

---

<sup>2</sup><http://www.physics.wisc.edu/~craigm/idl/idl.html>

## 2.4 Spectral Properties

### 2.4.1 Continuum Properties

In the spectral fitting process, the underlying continuum is comprised of the UV and optical power-law continua, the small blue bump, and low ionization iron emissions. In this section, we focus on the UV power-law spectral index as this is the dominant component of the quasar continuum shape at UV band. In addition, the optical PL is not covered in SDSS spectra with  $z > 0.3$ .

In the upper panel of Figure 2.12, we plot the UV power-law spectral index  $\alpha_\lambda$  as a function of luminosity for different redshifts. It is clear that most spectral indices lie between  $-1.9$  and  $-1.3$  with a median around  $\alpha_\lambda \approx -1.6$ ; these values correspond to  $\alpha_\nu (= -2 - \alpha_\lambda)$  between  $-0.1$  and  $-0.7$  with a median of  $-0.4$ . This is in agreement with previous work (e.g., Richstone & Schmidt 1980; Boroson & Green 1992; Vanden Berk et al. 2001; Dietrich et al. 2002; Telfer et al. 2002). For quasars with luminosity  $39 \lesssim \log l_\lambda \lesssim 42$  and  $0.000 \lesssim z \lesssim 1.512$ , more luminous quasars tend to have smaller  $\alpha_\lambda$  (bluer). Host galaxy contributions have been subtracted for low redshift ( $z < 0.7$ ) composite quasar spectra. This trend most likely reflects intrinsic reddening from quasars and/or an reddening due to intervening dust (Hopkins et al. 2004; Richards 2006, private communication) Above  $\log l_\lambda \sim 42$ , the slopes do not show strong dependence on luminosity. The highest redshift quasars in our sample ( $4.248 \leq z \leq 4.754$ ) do not appear to follow the above trend, but the result must be given low weight because composite spectra at these redshifts are generated by relatively small number of real quasar spectra (< 100; see Figure 2.2) and individual differences may be subject to small number variations.

The *upper* panel of Figure 2.12 also suggests a redshift dependence of  $\alpha_\lambda$ . In the *bottom* panel of Figure 2.12, we plot  $\alpha_\lambda$  as a function of redshift. We divide the entire redshift range into four regions corresponding to four sources of redshift measurements. The power-law indices of luminous and high redshift quasars with  $42 \lesssim \log l_\lambda \lesssim 44$  and  $1.6 \lesssim z \lesssim 3.8$  do not exhibit significant change over redshift. The rise at  $z \sim 1.512$  with respect to  $z \sim 1.291$  is because the line-free windows used to estimate the continua of these two spectra are different. As redshift increases, the long wavelength cut-off of the rest-frame spectra moves farther away from 4200 Å, which is a key benchmark wavelength for power-law estimation. Between 2200 Å and 4200 Å is a complex which consists of low ionization iron emission forest and the small blue bump whose flux level is very unpredictable. Although we can use the 1360 Å and 2200 Å windows as benchmarks to estimate the power-law, the uncertainty in the power-law inevitably becomes larger because these two benchmarks have relatively little wavelength separation. Using these two windows, the power-law is estimated based on only 1/3 of the entire wavelength range for a  $z = 1.512$  spectrum. Therefore, this dip may not reflect any true trends of spectral indices (Richards 2007, private communication). At redshifts lower than this local minimum, the spectral index tends to decrease with redshift. Spectral indices of less luminous quasars with  $40 \lesssim \log l_\lambda \lesssim 41$  change faster with redshift than more luminous quasars with  $41 \lesssim \log l_\lambda \lesssim 43$ . At low luminosity ( $\log l_\lambda \lesssim 40$ ), the number of quasar spectra used to generate the composite is small.

## 2.4.2 Properties of Emission Lines

### 2.4.2.1 Baldwin Effect (BEff)

Baldwin (1977) discovered that the rest equivalent width (EW) of C IV is anti-correlated with quasar luminosity density at 1450 Å,  $l_\lambda(1450 \text{ Å})$ . Later work found that the BEff exists for other strong UV/optical metal emission lines such as MgII  $\lambda 2798$  and C III]  $\lambda 1908$  (e.g., Kinney et al. 1990; Dietrich et al. 2002).

The BEff was initially thought to have the potential to predict quasar luminosity using only EW measurements. This idea was shown to be infeasible due to large scatter of this relationship (e.g., Baldwin et al. 1989; Kinney et al. 1990; Osmer & Shields 1999; Wu et al. 2009), especially in small samples or ones that do not cover a sufficiently large range in luminosity. The scatter is contributed by multiple sources, including the intrinsic BEff (e.g., Pogge & Peterson 1992), variability (Wu et al. 2009 estimated that  $\sim 60\%$  of the C IV  $\lambda 1549$  BEff can be attributed to variability), population effects (Baldwin 1977 noted that flat-radio-spectrum quasars showed the tightest C IV  $\lambda 1549$  BEff) and possibly evolutionary effects. Investigation to the evolutionary effects has been difficult for most flux-limited samples because it requires wide ranges of coverage in both redshift and luminosity. Our composite spectrum set, with its coverage of the redshift-luminosity panel (Figure 2.2) is ideal for the study of the behavior of BEff with redshift.

In Figure 2.13, we present seven composite spectra at the same redshift bin ( $z = 2.020$ ), but spanning more than an order of magnitude in luminosity. The spectra are normalized to unity at 2000 Å, and color-coded according to their luminosity. While the spectral shapes are similar, the strengths of several emission lines, such as Ly $\alpha$ , C IV  $\lambda 1549$ , C III]  $\lambda 1908$ , and Mg II  $\lambda 2798$ , exhibit evident BEff, i.e., more luminous quasars tend to have relatively weaker emission lines.

In Figure 2.14, we plot the behavior of the EW of eight strong emission lines with  $\log l_\lambda$ . The BEff exists for emission lines such as Ly $\alpha$ , SiIV  $\lambda 1396$ , CIV  $\lambda 1549$ , HeII  $\lambda 1640$ , and CIII]  $\lambda 1908$ . This result is consistent with previous works based on Bright Quasar Survey sample (Baskin & Laor 2004), SDSS sample (Richards et al. 2002a), 2dF+6dF sample (Croom et al. 2002) and samples consisting of type 1 AGNs (e.g., Wilkes et al. 1999; Dietrich et al. 2002). In Figure 2.15, we plot the distribution of uncertainties of log EW for all the data points in each panel of Figure 2.14. The uncertainties of log EW for most emission lines are less than 0.1. Larger uncertainties occur for composite spectra generated with small number ( $< 100$ ) of SDSS spectra .

We fit the BEff relationships of different emission lines at a given redshift as  $\log(\text{EW}/\text{EW}_0) = k \log(l_\lambda/l_\lambda^0) + b$ , in which  $\text{EW}_0 = 10 \text{ Å}$ , and  $l_\lambda^0 = 10^{42} \text{ erg s}^{-1}$ . The slopes and intercepts of three prominent emission lines (C IV  $\lambda 1549$ , Mg II  $\lambda 2798$  and SiIV  $\lambda 1396$ ) are tabulated in Table 2.6. The mean slope of the C IV  $\lambda 1549$ , Mg II  $\lambda 2798$  and SiIV  $\lambda 1396$  BEff is  $-0.276$ ,  $-0.207$  and  $-0.129$ , respectively. The uncertainty of BEff slope values depends on the number of data points in the linear fitting and should be  $\delta k \lesssim 0.1$ . The relatively large slope uncertainties arise because they are obtained by fitting only a few data points (typically  $< 10$ ). As a result, the slope can be sensitive to the EW value of a single or a few data points, which are affected by systematic errors transited from “line-free” window selection in the power-law estimation. The traditional ordinary least square fitting method is based on the Gaussian distributional

data points and is not suitable for small number of data points and may underestimate the uncertainties. Therefore, we perform a bias correction using the Jackknife statistics (Shao & Tu 1995).

For data points at a given redshift, we first perform a linear fit to all  $n$  data points and obtain the slope  $\hat{k}$  and intercept  $\hat{b}$  (note that the “intercept” is defined as  $l_\lambda^0 = 10^{42} \text{ erg s}^{-1}$ ). Then, we remove the  $i$ -th data point and fit the remaining ones to obtain a slope  $k_i$ . We repeat the second step for  $n$  times by looping  $i$  over each pair of ( $\log l_\lambda$ ,  $\log \text{EW}$ ). We can then estimate the Jackknife bias as  $k_b = (n-1)(\bar{k} - \hat{k})$ , which may be used as a correction term to produce re-estimate of  $\hat{k}$ :  $k_{\text{JK}} = \hat{k} - k_b = n\hat{k} - (n-1)\bar{k}$ , in which  $\bar{k} = \frac{1}{n} \sum k_i$ . The variance can be expressed as  $\delta k_{\text{JK}} = \sqrt{\frac{n-1}{n} \sum (k_i - \hat{k})^2}$ . In Table 2.6, we tabulate the slopes and intercepts of BEff relationships for C IV  $\lambda 1549$ , Mg II  $\lambda 2798$ , and Si IV  $\lambda 1396$  at each available redshift using the Jackknife method. In Figure 2.16, we plot the BEff slope as a function of redshift for the three emission lines based on the fitting results in Table 2.6. The typical Jackknife biases are  $\lesssim 0.01$  for both slopes and intercepts although it can be up to  $\sim 0.1$ .

The BEff slopes for Si IV  $\lambda 1396$  are consistent with that in Dietrich et al. (2002, here after D02), but our slopes for C IV  $\lambda 1549$  and Mg II  $\lambda 2798$  are  $\sim 0.15$  and  $\sim 0.1$  lower, respectively than those in D02. This discrepancy could be caused by the non-linear nature of the BEff, because the D02 sample contains not only quasars but also Seyfert 1 galaxies. In D02  $\log l_\lambda(1450 \text{ \AA})$  ranges from 38.55 to 44.55 (Figure 7 of Dietrich et al. 2002), which covers 6 orders of magnitude. It has been found that the BEff slope becomes shallower for low luminosity AGNs; for example, D02 perform linear fits to objects with  $\log l_\lambda \gtrsim 40.55$  for all emission lines, and found that the slopes are steeper than fitting all objects including lower luminosity. The lowest luminosities of our sample for these emission lines are 0–1.45 dex higher than 40.55. For C IV  $\lambda 1549$ , the data points start to depart from linearity below  $\log l_\lambda = 40.55$  so we obtained the same slope as D02. As a result, the luminosity range used to measure the BEff may have a considerable influence on the BEff slope and increase the scatter of this relationship.

The slopes of the BEff for all lines do not exhibit significant change over redshift. In Figure 2.16, the BEff slopes  $k$  for all of these three emission lines are consistent of being constant over redshift within the error bars. The error bars estimated by Jackknife method indicate that the variations are caused by systematic measurement uncertainties in spectral fitting process. The relatively large uncertainties are caused by performing linear fits with only a small number of data points. For example, each slope of the Mg II  $\lambda 2798$  BEff curve in Figure 2.16 is obtained by fitting only 7–9 data points; some slope values on the C IV  $\lambda 1549$  and Si IV  $\lambda 1396$  BEff curve are obtained by fitting 4–5 data points.

The scale factors ( $b$ , the intercept) of the BEff exhibit changes with redshift (Figure 2.14), namely, the BEff of quasars at higher redshift tend to have larger scale factors (larger EW at a given luminosity). For C IV  $\lambda 1549$  the increment is  $\sim 0.17$  dex from  $z = 1.512$  to  $z = 3.786$ ; for Mg II  $\lambda 2798$  the increment is  $\sim 0.25$  dex from  $z = 0.318$  to  $z = 2.020$ ; for Si IV  $\lambda 1396$ , however, the increment is only  $\sim 0.03$  dex from  $z = 1.754$  to  $z = 4.248$ . The changes of BEff scale factors over redshift contribute to scatter of the overall BEff correlation, and this contribution may be significant for a sample consisting

of objects covering a large range of redshift. The redshift dependence of BEff is an important factor that makes BEff impossible to be used for luminosity estimation.

Among the above three emission lines, the Mg II  $\lambda 2798$  BEff has the largest change of scale factor with redshift. Figure 2.20 shows the Mg II  $\lambda 2798$  EW as a function of redshift for a given luminosity. Quasars with  $41 \lesssim \log l_\lambda \lesssim 42$  exhibit a noticeable change from  $z \sim 1.5$  to  $z \sim 2.5$ . The decrease of EW at  $z = 1.512$  is caused by using different benchmark windows from  $z = 1.291$  (see Figure 2.12 and explanations in Section 2.4.1). The EW of quasars with  $42 \lesssim \log l_\lambda \lesssim 44$  do not exhibit a similar amount of change over redshift. Composites at the lowest and highest redshifts in Figure 2.20 have large uncertainties because they are generated from small number of quasar spectra.

In contrast to our results, D02 reported no detection of any significant evolutionary effect of EW for MgII and other emission lines. We suspect that this disagreement arises because their luminosity range for  $Mg\text{ II }43.0 \leq \log L_\lambda(1450\text{\AA}) \leq 44.0$ , corresponds to the high luminosity range of our composites. As seen from Figure 2.20, the EW change in that luminosity range is not as significant as lower luminosity ones.

In our emission line fitting process, we attempt the difficult separation of Ly $\alpha$  and Nv. We do not find Nv  $\lambda 1239$  exhibits a significant BEff. The mean slope of a linear fit to log EW and  $\log l_\lambda$  for Nv is  $-0.027 \pm 0.025$ ; this weak dependence of Nv EW with luminosity is in agreement with Osmer et al. (1994) and D02. However, similar to emission lines with BEff, we do see an increase of the average EW with redshift, which is about 0.3 dex from  $z = 2.020$  to  $z = 4.248$ .

The H $\beta$  EW in our sample also exhibits only a weak dependence on luminosity; an average BEff slope is  $-0.029 \pm 0.009$  (see Table 2.7). This measurement is consistent with results by Boroson & Green (1992) and Dietrich et al. (2002). However, using a sample of 22,000 quasars from 2dF+6dF, Croom et al. (2002) finds an *inverse* BEff for H $\beta$  and H $\gamma$  in which BEff slope  $k = 0.2$ . Greene & Ho (2005) found this slope is  $k = 0.1$  using 229  $z \leq 0.35$  AGNs (so that H $\alpha$  is in the bandpass) selected from SDSS DR3 (Abazajian et al. 2005). Croom et al. (2002) postulates that this result is caused by a luminosity change in the ratio of disk to non-disk continuum components. The weak or inverse BEff of H $\beta$  probably reflects that individual species have different responses to changes of ionizing flux. High ionization ions are more sensitive to the changes of luminosity than low ionization ions.

It has been reported that the BEff slope is anti-correlated with ionization potential of the ion producing an emission line (Espey & Andreadis 1999; Dietrich et al. 2002). The ionization potentials used in Vanden Berk et al. (2001) and Dietrich et al. (2002) have different definitions. The former uses the ionization potential to *ionize* a particular ion  $U_1$ , e.g., from MgII  $\rightarrow$  MgIII, while the latter uses the ionization potential  $U_0$  to *create* this ion, e.g., from MgI to MgII. To be consistent, we adopt the convention in Vanden Berk et al. (2001). In Figure 2.7, we plot the BEff slopes of prominent emission lines as a function of their ionization potential  $U_1$ . Comparing these results with Figure 9 of D02, the values of slopes are different for several lines such as C IV  $\lambda 1549$ , and Mg II  $\lambda 2798$ , which could be due to difference of luminosity range (see argument above), but the general trend remains, i.e., emission lines produced by ions with higher potentials tend to have steep BEff slopes. This trend may indicate a stratified structure of the broad line region (BLR). Under the locally optically emitting clouds (LOC) model (Baldwin et al.

1995), the each emission line is optimally produced at a certain radius with a narrow range of density. Higher ionization lines are produced closer to the central continuum source than lower ionization lines, and are more sensitive to luminosity changes than lower ionization lines, resulting in a steeper BEff slope. The N V line is an exception, probably because of its abnormal abundance pattern (Korista et al. 1998, , here after K98). K98 found in their calculations that while the EW of Ly $\alpha$ , C IV  $\lambda$ 1549 and O VI  $\lambda$ 1032 exhibit almost no dependence on metallicity, the EW of N V  $\lambda$ 1239 shows a strong dependence on gas metallicity (Figure 1 of K98). As a secondary element, nitrogen's rapidly increasing abundance with increasing metallicity ( $N/H \propto Z^2$ ) compensates for the losses in its EW emitted by gas illuminated by softer continua in higher-luminosity quasars. In addition, there is a growing evidence that the gas metallicity increases with increasing luminosity. Earlier work based on the N V/He II and N V/C IV line ratios (Hamann & Ferland 1992, 1993; Ferland et al. 1996) indicates that metallicities are typically solar or higher and nitrogen is selectively enhanced. There is also a significant trend in these line ratios, suggesting higher metallicities in more luminous quasars (e.g., Osmer et al. 1994; Korista et al. 1998). Figure 7 in Hamann & Ferland (1999) clearly shows that emission line ratio of N V/He II increases with luminosity. The combination of the facts that metallicity increases with luminosity and the strong dependence of nitrogen abundance on metallicity provide a plausible explanation that N V  $\lambda$ 1239 does not exhibit a BEff.

In early work by Cristiani & Vio (1990) and Francis & Koratkar (1995), the SiIV+OIV] blend shows little or no evidence of BEff. In our fitting process, we attempt to separate the blend and found an evident BEff for Si IV  $\lambda$ 1396 (Figure 2.14).

#### 2.4.2.2 Systematic Emission Line Shift

The expression “emission line shift” refers to the emission line center offset from the laboratory wavelength after transforming a spectrum from the observer’s frame to the quasar’s rest frame (Gaskell 1982). Because quasar redshifts are measured by fitting reference emission/absorption lines, it is usually difficult to define the *absolute* emission line shift for quasar spectra which do not cover proper reference lines. Redshift estimates based on the detection of photospheric absorption from stars in the spatially averaged spectrum of the quasar host galaxy might be expected to provide a close to “ideal” systematic redshift. The most reliable reference used to measure quasar redshifts are perhaps the Ca II  $\lambda\lambda$ 3934.8, 3969.6 doublet. However, given the nature of the SDSS spectra, coupled with the large luminosity of quasars at rest-frame compared to their host galaxies, it is impossible to detect the Ca II H&K feature for the majority of SDSS quasars. As a result, we must use other emission lines as surrogates to Ca II H&K. Thus, for most quasars we measure the *relative* emission line shift which is the offset of an emission line centroid after shifting the spectrum to the rest-frame of a reference emission line.

The narrow forbidden emission lines of [O III]  $\lambda\lambda$ 4960, 5008 are prominent in many low redshift ( $z < 0.8$ ) SDSS quasar spectra and are usually taken as a good reference line to redshift measurements (e.g., Hewett & Wild 2010). In Figure 2.18, we present distributions of relative emission line shifts for individual or composites. The emission line centroids are measured using the eigenspectrum technique described in Section 2.2.

The positions of distribution peaks thus represent the average line shift with respect to a given reference line. In the upper left panel of Figure 2.18, the distribution peak is well consistent with  $0 \text{ km s}^{-1}$  with a few tens of  $\text{km s}^{-1}$  redshift indicating that the redshift measured based on  $\text{H}\beta$  emission line is close to the redshift of the quasar itself (from [O III]). The upper right panel demonstrates that the  $\text{Mg II } \lambda 2798$  line has a systematic offset of  $\lesssim -100 \text{ km s}^{-1}$  in the rest-frame of [O III]. The distribution at the lower left indicates that this offset becomes larger ( $\lesssim 200 \text{ km s}^{-1}$ ) if the redshift is measured according to  $\text{H}\beta$ . The panel on the lower right indicates that the  $\text{C IV } \lambda 1549$  line is systematically shifted towards blue by  $\sim 400 \text{ km s}^{-1}$  in the rest-frame of  $\text{Mg II } \lambda 2798$ . Combining information from these distributions, it is clear that the  $\text{Mg II } \lambda 2798$  and  $\text{C IV } \lambda 1549$  are systematically blueshifted with respect to the quasar rest-frame of by  $\sim 150 \text{ km s}^{-1}$  and  $\sim 800 \text{ km s}^{-1}$ , respectively. The emission line shift values measured in spectra whose redshifts are determined by different emission lines can be different, so emission line shift correlations must be studied using a set of spectra whose redshifts are measured using the same reference line.

Line-shifts of different ions exhibit a general anti-correlation with ionization potentials of the ions contributing to emission lines (e.g., Tytler & Fan 1992; Vanden Berk et al. 2001; Richards et al. 2002a). Most emission lines are blue shifted with respect to the quasar host galaxy (e.g., see Table 4 in Vanden Berk et al. 2001), suggesting that emission line regions are moving towards the observer along the line-of-sight. This anti-correlation is also seen in our quasar sample. In Figure 2.19, we plot the mean emission line shifts which are averaged over a redshift range vs. ionization potentials for a number of emission lines. The four panels correspond to four ways to measure object redshifts. The mean value of emission line shift may vary in different panels because the redshifts are measured using different reference lines. Most of emission line shifts in Panel (c) are positive, indicating that the  $\text{C IV}$  line is a strongly blue shifted relative to the quasar redshift.

The emission line shifts of permitted and semi-forbidden lines are in general an order of magnitude larger than forbidden lines. This result is consistent with Figure 9 of Vanden Berk et al. (2001). Most forbidden lines are narrow lines and are produced in the NLR, which is distant from the central continuum source. The ionization radiation has much weaker influence to these line emitters than in the broad line region which is much closer to the central engine.

Our result is consistent with that by D02 in that the  $\text{C IV } \lambda 1549$  is systematically blueshifted by  $800 \text{ km s}^{-1}$  on average with respect to the  $\text{Mg II } \lambda 2798$ . In Panel (b) of Figure 2.19, the average separation between  $\text{C IV } \lambda 1549$  and  $\text{Mg II } \lambda 2798$  is  $\sim 600 \text{ km s}^{-1}$ . In Figure 2.21, we plot the emission line redshift as a function of luminosity, which shows that this shift increases with luminosity from  $\sim 100 \text{ km s}^{-1}$  for low luminosity ( $\log l_\lambda \sim 41.5$ ) to  $\sim 800 \text{ km s}^{-1}$  for high luminosity objects ( $\log l_\lambda \sim 43.0$ ).

The exact origin of line shift is not clear, but Richards et al. (2002a) suggested that the apparent blueshift is due to a lack of flux in the red wing of the CIV emission line profile. As a result, more blueshifted emission lines have stronger BEff. This relation does not apply for  $\text{N V } \lambda 1239$ , which exhibits a strong emission line shift (Panel 3 of Figure 2.19) but it does not exhibit strong BEff. Another example is  $\text{Si IV } \lambda 1396$ , which displays a strong BEff but has a small line shift with respect to  $\text{Mg II } \lambda 2798$ .

Some emission line shifts exhibit evident dependence with luminosity and/or redshift. Figure 2.21 presents the mean line shift over given redshift ranges for some emission lines as a function of luminosity. In Panel (a) of Figure 2.21, the H I  $\lambda 5877$  tends to have a smaller red emission line shift with respect to H $\beta$ . In Panel (b) of Figure 2.21, C IV emission line shift decreases with increasing luminosity. This result is consistent with Figure 9 in Hewett & Wild (2010). Combining this relation with the trend that more luminous quasars have smaller C IV EW, this fact indicates that these two phenomena are likely to be connected, as both of the EW and line shifts are correlated by the spectral energy distribution (SED). The strong anti-correlation between EW(C IV) and  $\alpha_{\text{ox}}$ <sup>3</sup> is presented in Figure 4 of Wu et al. (2009), indicating that the  $\alpha_{\text{ox}}$  is a fundamental driver of the change of EW(C IV). On the other hand, the C IV blueshift is also correlated with  $\alpha_{\text{ox}}$  (see Figure 9 of Richards et al. 2011), suggesting that objects with more negative values of  $\alpha_{\text{ox}}$  (softer SED) have more strongly blueshifted C IV. A more straightforward view of the relationship between C IV EW and its blueshift can be seen in Figure 10 of Richards et al. (2011). The strong correlation between these two quantities suggests that these two phenomena are not independent and both of them are possibly controlled by  $\alpha_{\text{ox}}$ , which approximately characterizes the SED between UV and X-ray bands. If this is true, both the C IV BEff and the luminosity dependence of line shift are both secondary effects, considering the UV luminosity is strongly correlated with  $\alpha_{\text{ox}}$  (e.g., Vignali et al. 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007). Richards et al. (2011) considers these two phenomena within the context of an accretion disk wind model in which the wind filters the continuum seen by the disk component. A detailed discussion of this model is beyond the scope of this work, but the C IV EW and blueshifts are capturing an important trade-off between “disk” and “wind” components to UV emission lines that ultimately sensitive to the shape of the ionizing continuum.

In Panel (c) of Figure 2.21, the strong increase of positive emission line shift of C III]  $\lambda 1908$  and Mg II  $\lambda 2798$  with luminosity is a result of strong blue shift of C IV, because these mean line shifts are with respect to C IV, considering that Mg II does not exhibit a strong systematic line shift with respect to H $\beta$  (upper left panel of Figure 2.19) and that C III]  $\lambda 1908$  does not exhibit a strong systematic line shift with Mg II (upper right panel of Figure 2.19).

In Figure 2.22, we see a decrease of line-shift of HeI  $\lambda 5877$  with redshift, but we do not see similar trends for other emission lines.

#### 2.4.2.3 Full Width at Half Maximum (FWHM) of Emission Lines

We also found changes of FWHMs for several emission lines as functions of redshift and luminosity. Figure 2.23 presents the plots of FWHM against luminosity for Mg II  $\lambda 2798$  and C IV  $\lambda 1549$  at different redshifts. For Mg II and C IV, we see a trend of increasing line width with luminosity. The average slope of this correlation is  $0.084 \pm 0.022$

---

<sup>3</sup>Defined as  $\alpha_{\text{ox}} = \log [l_{\nu}(2 \text{ keV})/l_{\nu}(2500 \text{ \AA})] / \log [\nu(2 \text{ keV})/\nu(2500 \text{ \AA})] = 0.3838 \log [l_{\nu}(2 \text{ keV})/l_{\nu}(2500 \text{ \AA})]$ .  $\alpha_{\text{ox}}$  is used to characterize the spectral hardness in the UV to X-ray band (e.g., Avni & Tananbaum 1982, 1986; Anderson & Margon 1987; Wilkes et al. 1994; Vignali, Brandt & Schneider 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007).

for C IV and  $0.051 \pm 0.008$  for Mg II, indicating that the FWHM has a weak positive correlation with luminosity.

In Figure 2.23, we qualitatively see a trend of the dependence of FWHM-luminosity relation on redshift. In Figure 2.24 we plot the FWHM against redshift, and find that the FWHM exhibits a mild evolutionary trend with redshift. For a given luminosity, higher redshift quasars tend to have smaller FWHM. This trend is seen in all luminosity ranges of C IV, but only low luminosity quasars for Mg II. For C IV, the change of FWHM can be up to  $2000 \text{ km s}^{-1}$  from  $z = 1.754$  to  $z = 4.248$ . The change of Mg II FWHM is only  $\approx 400 \text{ km s}^{-1}$  from  $z = 0.905$  to  $z = 1.512$ .

## 2.5 Discussion

The most intriguing conclusion we draw using this composite spectrum set is that the BEff evolves with redshift. This evolutionary trend was not seen in previous published work because the SDSS quasar catalog provides a dataset with unprecedented combination of wide ranges of redshift, luminosity and number of quasars. The origin of the BEff has long been a subject of investigation and a number of possible explanations have been proposed. For example, one promising explanation is that the UV/optical BEff is due to the softening of the SED at high luminosity, which lowers the ion populations having higher ionization potentials (e.g., Netzer et al. 1992; Zheng & Malkan 1993; Green 1996, 1998; Korista et al. 1998; Gibson et al. 2008; Wu et al. 2009). Other proposed BEff divers include the Eddington ratio (e.g., Baskin & Laor 2004; Bachev et al. 2004; Warner et al. 2004; Zhou & Wang 2005), and black hole mass (e.g., Netzer et al. 1992; Wandel et al. 1999; Shields 2007).

The quasar SEDs, Eddington ratios and black hole masses are all found to be correlated. For example, Shemmer et al. (2008) found a mild correlation (Spearman rank coefficient  $r = -0.33$ , confident level  $p < 1.3 \times 10^{-3}$ ) between  $\alpha_{\text{ox}}$  and  $L_{\text{bol}}/L_{\text{Edd}}$  using 81 sources from Steffen et al. (2006) sample with available FWHM(H $\beta$ ) measurements ( $z \lesssim 0.75$ ; Figure 6 in their paper); Grupe et al. (2010) claims a strong correlation (Spearman rank coefficient  $r = 0.55$ ) between  $\alpha_{\text{ox}}$  and  $L_{\text{bol}}/L_{\text{Edd}}$  (Figure 13 of their paper) using  $\sim 100$  soft X-ray select AGNs; a similar correlation is also found in 150 X-ray selected Type 1 AGNs in Lusso et al. (2010). The strong positive correlation between AGN black hole mass and luminosity has been established and confirmed by reverberation mapping results (e.g., Peterson et al. 2004) and AGN UV luminosity is found to be strongly correlated to  $\alpha_{\text{ox}}$  (e.g., Vignali et al. 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007). Combining all of these findings, BEff is transitively dependent on all of these fundamental parameters used to characterize quasars. Therefore, the evolution of BEff may be related to any of these parameters which exhibits evolution over redshift. Particularly, the black hole growth over redshift may affect the accretion rate which leads to a change of the SED shape. This can further cause a change of the ionization structure of the BLR and eventually result in a change of line emission, which leads to a systematic change of EW over redshift of a given luminosity.

The past decade has been great progress in measuring the black hole masses in both nearby normal galaxies and AGNs (e.g., Kaspi et al. 2000; Peterson et al. 2004). There has been growing number of observational evidence and constraints on the AGN

black hole growth across the cosmic time (e.g., Yu & Tremaine 2002; Cao 2010; Kelly et al. 2010), although whether the black hole growth is down or up sizing is still in debate. The black hole growth changes the accretion states which is also evolving (Yu et al. 2005). A possible scenario is that as the black hole mass grows up, the accretion becomes relatively inefficient, which lowers the accretion disk temperature. This makes the incident flux from the background continuum source softer which decreases the yield of the C<sup>3+</sup> ion (see Wu et al. 2009 for a positive correlation between EW(C IV) and  $\alpha_{\text{ox}}$ ). A further investigation of this problem may require a detailed model of quasar activity combined with observational constraints of black hole growth (e.g., Small & Blandford 1992; Marconi et al. 2004), which is beyond the scope of this work.

## 2.6 Conclusion

We have compiled a set of 186 composite quasar spectra from SDSS DR3 quasar catalog using the spectral binning, averaging and weighting techniques in Vanden Berk et al. (2001). The advantage of the composite spectra is that this technique averages the differences in individual spectra so as to highlight the common spectral properties. In addition, the composite spectrum has a significantly higher  $S/N$  than an individual SDSS spectrum so that even weak spectral features can be detected. This set of composite spectra allows one to investigate spectral properties of quasars as a function of luminosity and redshift. We developed a software package *QSFIT* to fit both the underlying continua and emission lines of all the composite spectra. Studying the continuum and emission line properties reveals a number of trends.

1. The UV spectral indices of all quasars have a mean value of  $\alpha_\lambda \approx -1.6$  (or equivalently,  $\alpha_\nu = -0.4$ ). Low luminosity quasars at small redshifts tend to be redder than high luminosity quasars at higher redshift. This trend applies to quasars with  $\log l_\lambda \lesssim 42$  and  $z \lesssim 2$ .
2. The Baldwin Effect exists not only for C IV  $\lambda 1549$  but also for other broad emission lines such as Ly $\alpha$ , Si IV  $\lambda 1396$ , He II  $\lambda 1640$ , and Mg II  $\lambda 2798$ , but this effect is not seen for N V  $\lambda 1239$  and H $\beta$ . While the slopes of the individual line BEffs do not exhibit significant dependence over redshift, the scale factor of BEff exhibits significant change with redshift. Typical changes are  $\sim 0.2$  dex for C IV  $\lambda 1549$  from  $z = 1.5$  to  $z = 3.8$ , and  $\sim 0.25$  dex for Mg II  $\lambda 2798$  from  $z = 0.3$  to  $z = 2.0$ . The scale factor change for Si IV  $\lambda 1396$  is only  $\sim 0.03$  dex from  $z = 1.7$  to  $z = 4.3$ . The change of scale factor is an important contributor to the dispersion in the BEff for a sample of quasars covering a wide range of redshifts.
3. We confirmed that in general the mean emission line shift decreases with the BEff slope and ionization potential, i.e., emission lines produced by highly ionized atoms are more blue-shifted, suggesting that the BLR region has a stratified structure. The emission line shift of He I  $\lambda 5877$  with respect to H $\beta$  decreases with luminosity; the emission line shifts of C IV  $\lambda 1549$  with respect to Mg II  $\lambda 2798$  becomes more negative with increasing luminosity reaching  $\sim 800 \text{ km s}^{-1}$ ; in contrast, Al III  $\lambda 1857$  and C III]  $\lambda 1908$  emission lines exhibit a positive velocity shift with respect to Mg II  $\lambda 2798$ , and the amount of line shift increases with luminosity.

4. The Mg II  $\lambda 2798$  and C IV  $\lambda 1549$  emission lines tend to be broader for high luminosity quasars; The average slope is  $\alpha = 0.084 \pm 0.022$  for C IV and  $\alpha = 0.051 \pm 0.008$  for Mg II ( $\text{FWHM} \propto l_\lambda^\alpha$ ). The amplitudes of this correlation also exhibit changes with redshift. At a given luminosity, the FWHMs of C IV  $\lambda 1549$  or Mg II  $\lambda 2798$  tend to be lower at higher redshift.
5. The BEff, and the emission line shift redshift dependences are likely to be connected. The strong BEff and blueshift of C IV  $\lambda 1549$  with respect to Mg II suggest that the background continuum may be shielded by the outflow from the accretion disk, which supports the accretion disk wind model proposed by Richards et al. (2002a). However, it remains unclear why some emission lines exhibit strong BEff but small amount of emission line shift, e.g., Mg II  $\lambda 2798$  and Si IV  $\lambda 1396$ .

The most intriguing conclusion we draw using this composite spectrum set is that the BEff evolves with redshift. This evolutionary trend was not seen in previous published work because the SDSS quasar catalog provides an dataset with unprecedented combination of wide ranges of redshift, luminosity and number of quasars. The origin of the BEff has long been a subject of investigation and a number of possible explanations have been proposed. For example, one promising explanation is that the UV/optical BEff is due to the softening of the SED at high luminosity, which lowers the ion populations having higher ionization potentials (e.g., Netzer et al. 1992; Zheng & Malkan 1993; Green 1996, 1998; Korista et al. 1998; Gibson et al. 2008; Wu et al. 2009). Other proposed BEff divers include the Eddington ratio (e.g., Baskin & Laor 2004; Bachev et al. 2004; Warner et al. 2004; Zhou & Wang 2005), and black hole mass (e.g., Netzer et al. 1992; Wandel et al. 1999; Shields 2007).

The quasar SEDs, Eddington ratios and black hole masses are all found to be correlated. For example, Shemmer et al. (2008) found a mild correlation (Spearman rank coefficient  $r = -0.33$ , confident level  $p < 1.3 \times 10^{-3}$ ) between  $\alpha_{\text{ox}}$  and  $L_{\text{bol}}/L_{\text{Edd}}$  using 81 sources from Steffen et al. (2006) sample with available FWHM(H $\beta$ ) measurements ( $z \lesssim 0.75$ ; Figure 6 in their paper); Grupe et al. (2010) claims a strong correlation (Spearman rank coefficient  $r = 0.55$ ) between  $\alpha_{\text{ox}}$  and  $L_{\text{bol}}/L_{\text{Edd}}$  (Figure 13 of their paper) using  $\sim 100$  soft X-ray select AGNs; a similar correlation is also found in 150 X-ray selected Type 1 AGNs in Lusso et al. (2010). The strong positive correlation between AGN black hole mass and luminosity has been established and confirmed by reverberation mapping results (e.g., Peterson et al. 2004) and AGN UV luminosity is found to be strongly correlated to  $\alpha_{\text{ox}}$  (e.g., Vignali et al. 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007). Combining all of these findings, BEff is transitively dependent on all of these fundamental parameters used to characterize quasars. Therefore, the evolution of BEff may be related to any of these parameters which exhibits evolution over redshift. Particularly, the black hole growth over redshift may affect the accretion rate which leads to a change of the SED shape. This can further cause a change of the ionization structure of the BLR and eventually result in a change of line emission, which leads to a systematic change of EW over redshift of a given luminosity. A further investigation of this problem may require a detailed model of quasar activity combined with observational constraints of black hole growth (e.g., Small & Blandford 1992; Marconi et al. 2004), which is beyond the scope of this work.

Because of limited redshift coverage, most previous work focus on quasar emission line and continuum properties at a particular redshift or within a narrow redshift range. Our work confirms and refines many existing relationships such as the emission line shift vs. ionization and emission line shift vs. BEff slope relations, over a wider range of redshift. Both of these new and refined correlations reflect *average* quasar properties, but not a small set of selected quasars. Because of this, the conclusions we obtained in this work are more *general*. These properties will provide strong constraints to current and future quasar models, and must be re-produced, as a function of redshift and luminosity, by successful emission models.

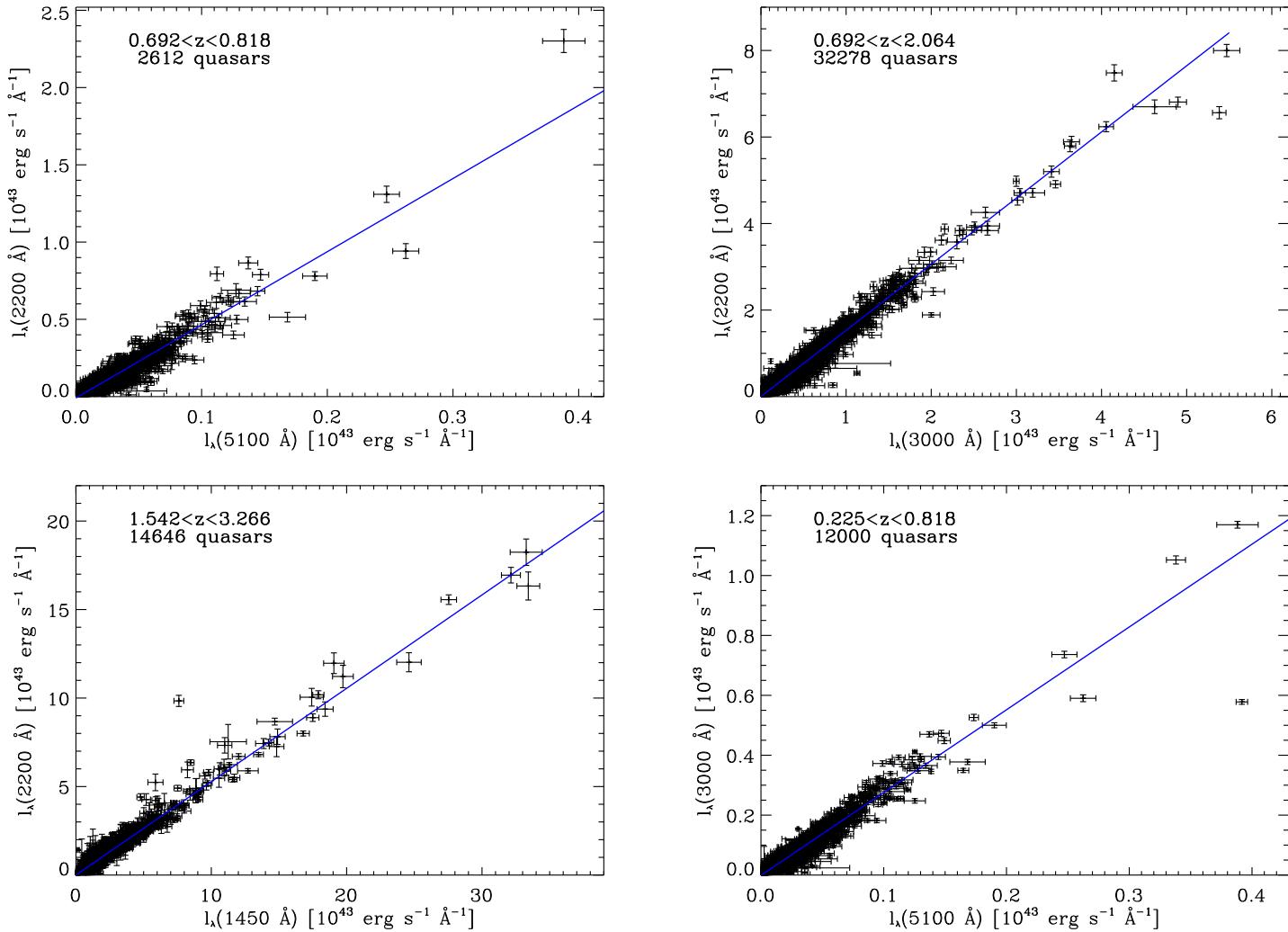


Fig. 2.1 Calibrated correlation of  $l_{\lambda}(2200 \text{ \AA})$  with monochromatic luminosities at 5100  $\text{\AA}$  (upper left), 3000  $\text{\AA}$  (upper right), and 1450  $\text{\AA}$  (bottom left). The panel at bottom right is a calibration of  $l_{\lambda}(5100 \text{ \AA})$  and  $l_{\lambda}(3000 \text{ \AA})$ . Luminosity density values are sampled from SDSS DR5 quasar spectra from DR3 quasar catalog. In each panel, the solid line is the linear fit of all data points. The remarkable linear correlations indicate that these monochromatic luminosities in UV and optical bands are well correlated. It is reasonable to use  $l_{\lambda}(2200 \text{ \AA})$  as a luminosity indicator.

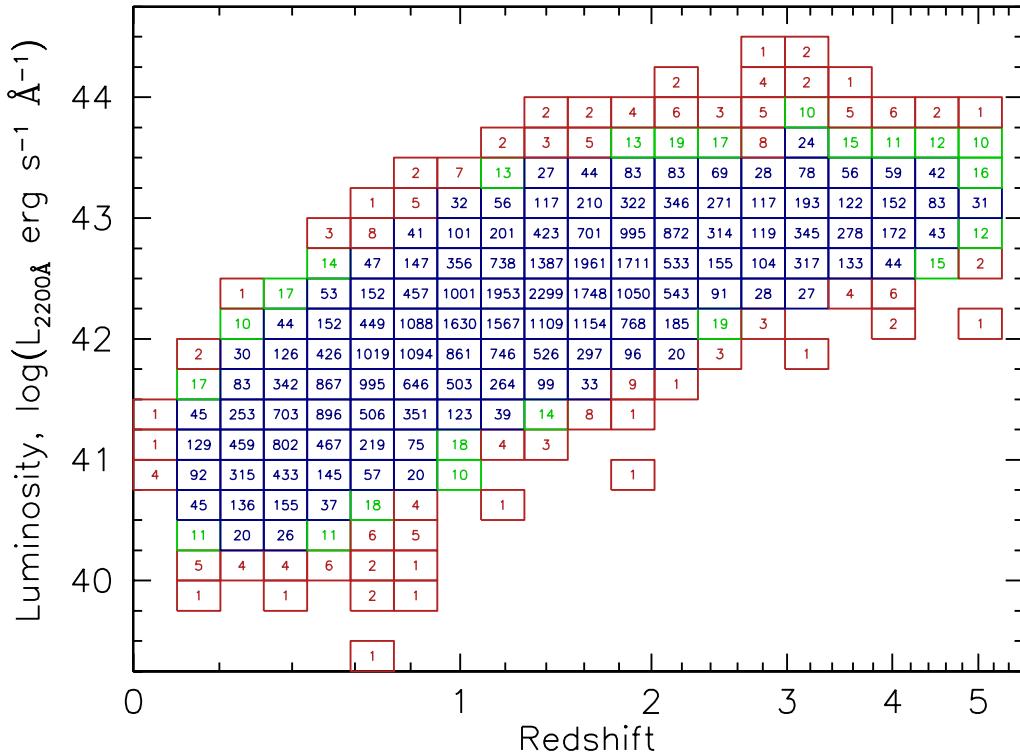


Fig. 2.2 The number of quasars in each redshift-luminosity density bin that represent the composite spectra. The composites are generated from 44,969 quasars from SDSS DR3 quasar catalog. Bins with twenty or more quasars are shown in blue, bins with ten to 19 quasars are shown in green, and those with less than ten are shown in red. Redshift bin limits are separated by a constant value of  $\Delta \log(1 + z) = 0.04$ ; luminosity density bin limits are separated by a logarithmic spacing of 0.25.

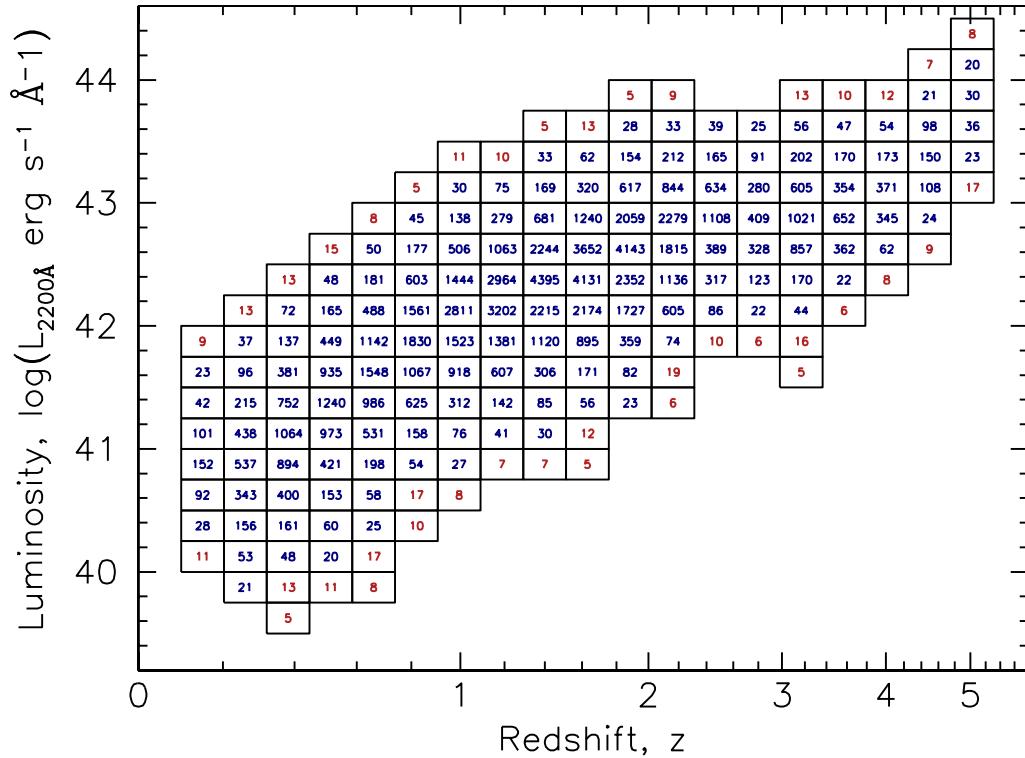


Fig. 2.3 Similar diagram as Figure 2.2, but based on SDSS DR7 quasar catalog . This set of composite spectra is generated based on 92,586 quasar spectra. Bins with numbers in red represent composite spectra generated from less than 20 real quasar spectra.

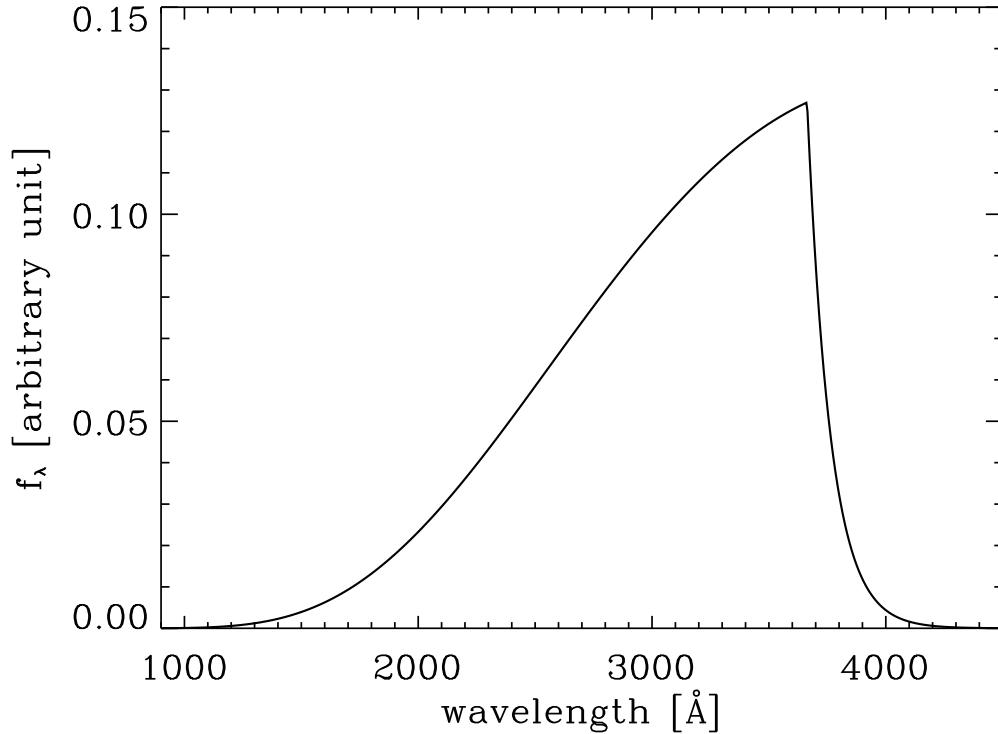


Fig. 2.4 The shape of the 3000 Å bump for  $T_B = 10^4$  K, and  $\tau_B = 1$ , which contains a modulated black body spectrum at  $\lambda \leq \lambda_B = 3663$  Å, and an exponentially decay at  $\lambda > \lambda_B$ . The flux density is given in linear arbitrary units.

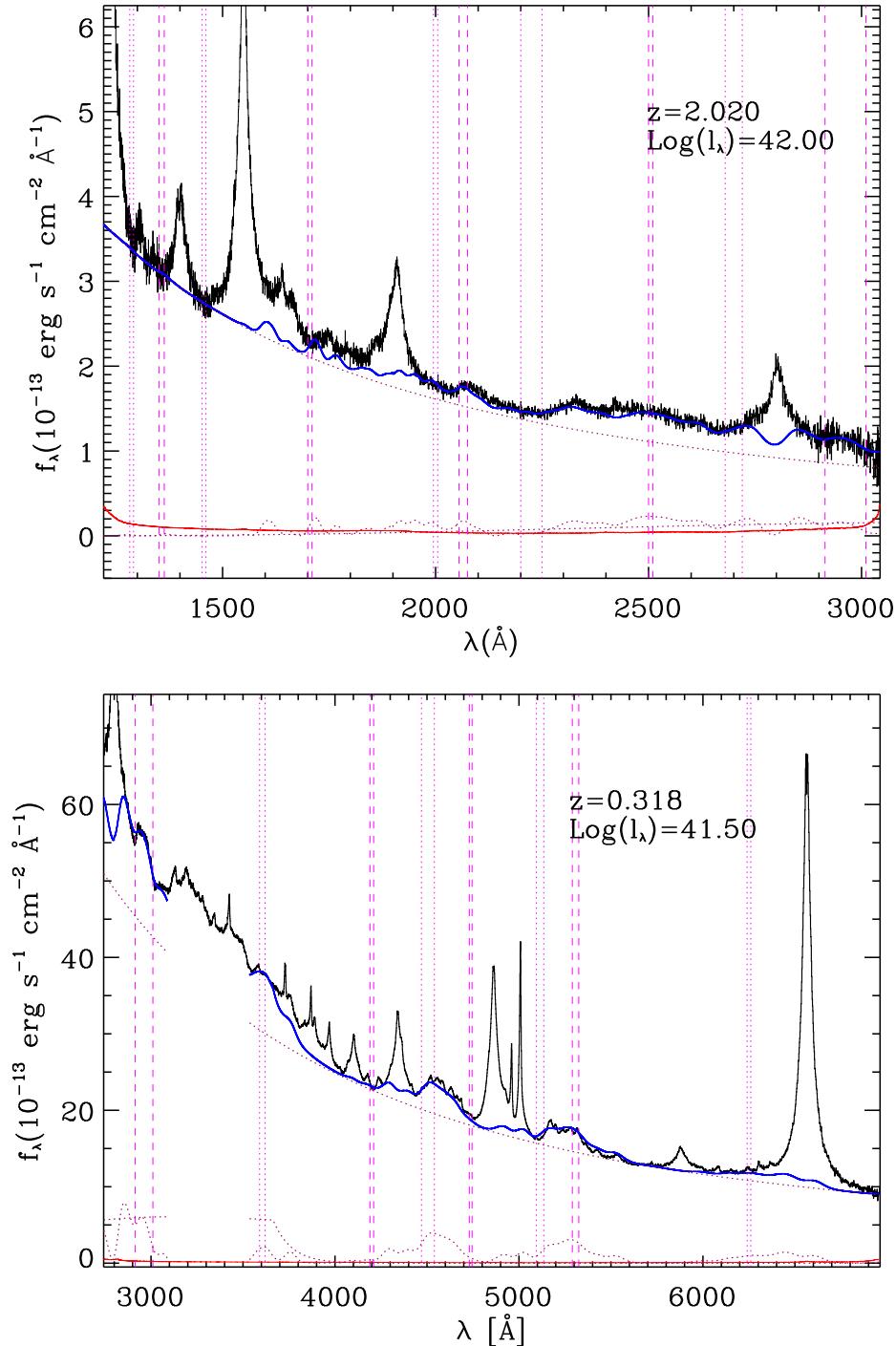


Fig. 2.5 *Upper panel*: An example continuum fitting to the composite spectrum at the bin  $(z, \log l_\lambda) = (2.020, 42.00)$ , generated from 185 quasar spectra. *Lower panel*: An example continuum fitting to the composite spectrum at the bin  $(z, \log l_\lambda) = (0.318, 41.50)$ , generated from 342 quasar spectra. The composite spectra have an effective spectral resolution of  $R \sim 1800$ . In each panel, the black histogram is the original composite spectrum; the solid blue line is the overall fit of the underlying continuum; the dotted purple lines are different continuum components. The magenta dotted and dashes lines mark the “line-free” windows used to evaluate the goodness-of-fit. There is a “gap” in the spectral fitting results at the lower panel at  $\sim 3500$  Å because of the lack of the iron template in this region.

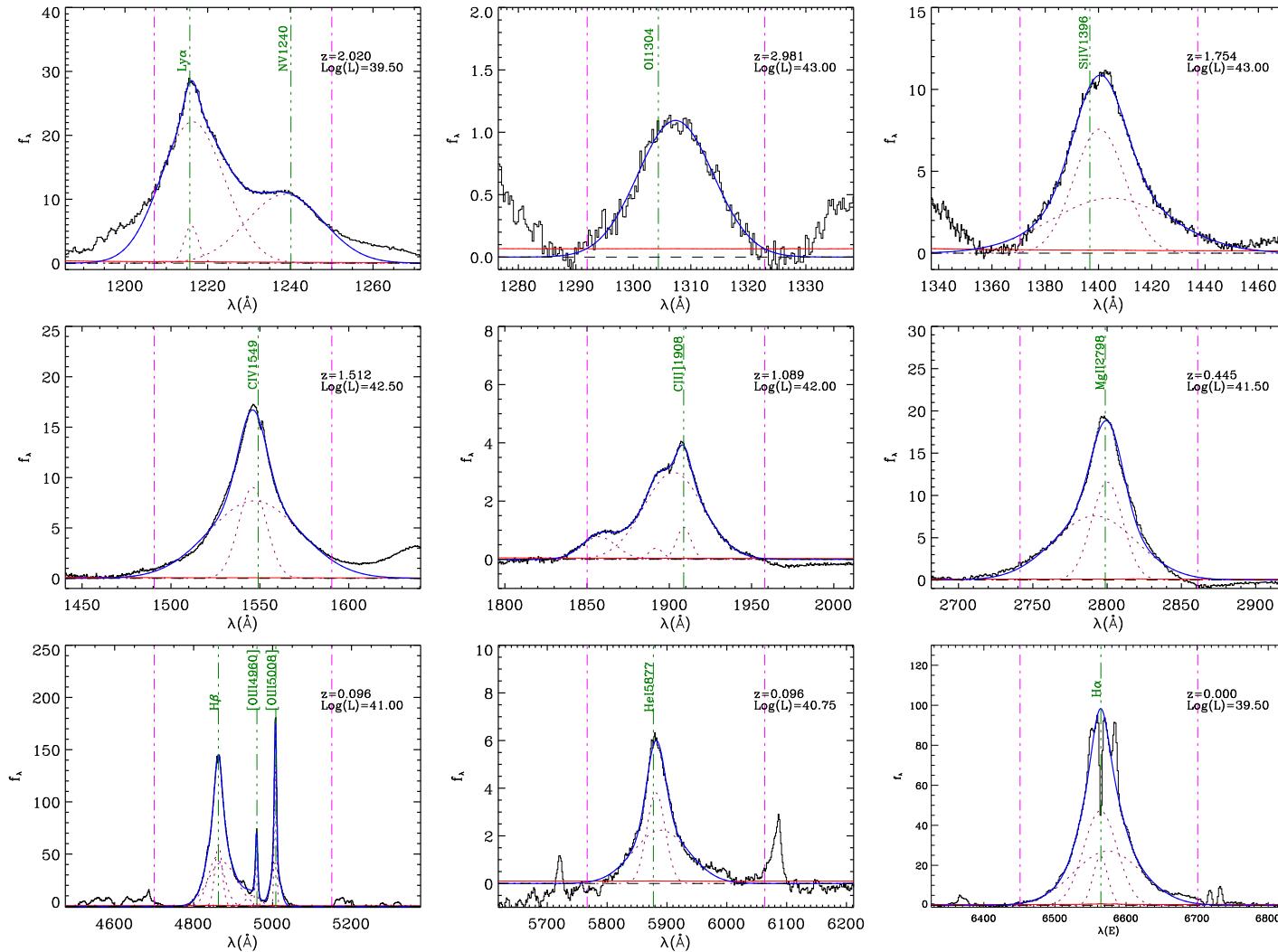


Fig. 2.6 Examples of emission line fitting to the composite spectra. The black histograms are the continuum-subtracted spectrum; the magenta dotted curves are Gaussian components; the blue solid curves are the synthesized emission line profiles; the magenta dash-dotted lines bound the regions used to fit the emission lines. In the He I  $\lambda 5877$  panel the [Fe VII] emission line around  $6100 \text{\AA}$  is fit separately. In the H $\alpha$  panel, the two weak sulphur lines beyond the red wing of H $\alpha$  and the two [N II] lines on the each side of H $\alpha$  peak are fit separately (see emission line fitting description in Section 2.3.2). The narrow absorption features at the peak of H $\alpha$  are due to over-subtraction of host galaxy component and are masked in the fitting process. The emission line identifications are printed in green.

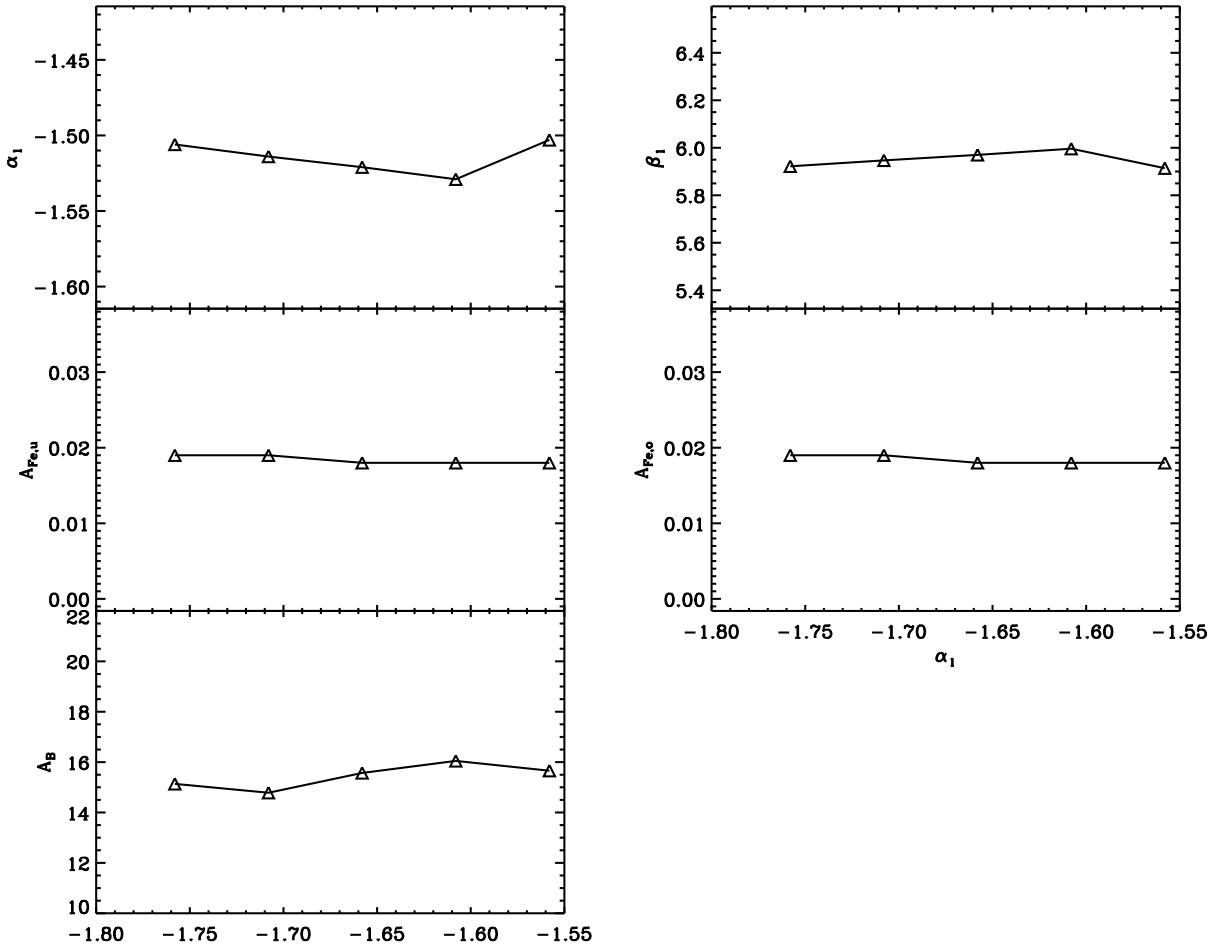


Fig. 2.7 Continuum parameter fitting results by varying UV spectral index  $\alpha_1$ . Refer to Table 2.3 for meanings of parameter symbols and Table 2.5 for varying ranges.

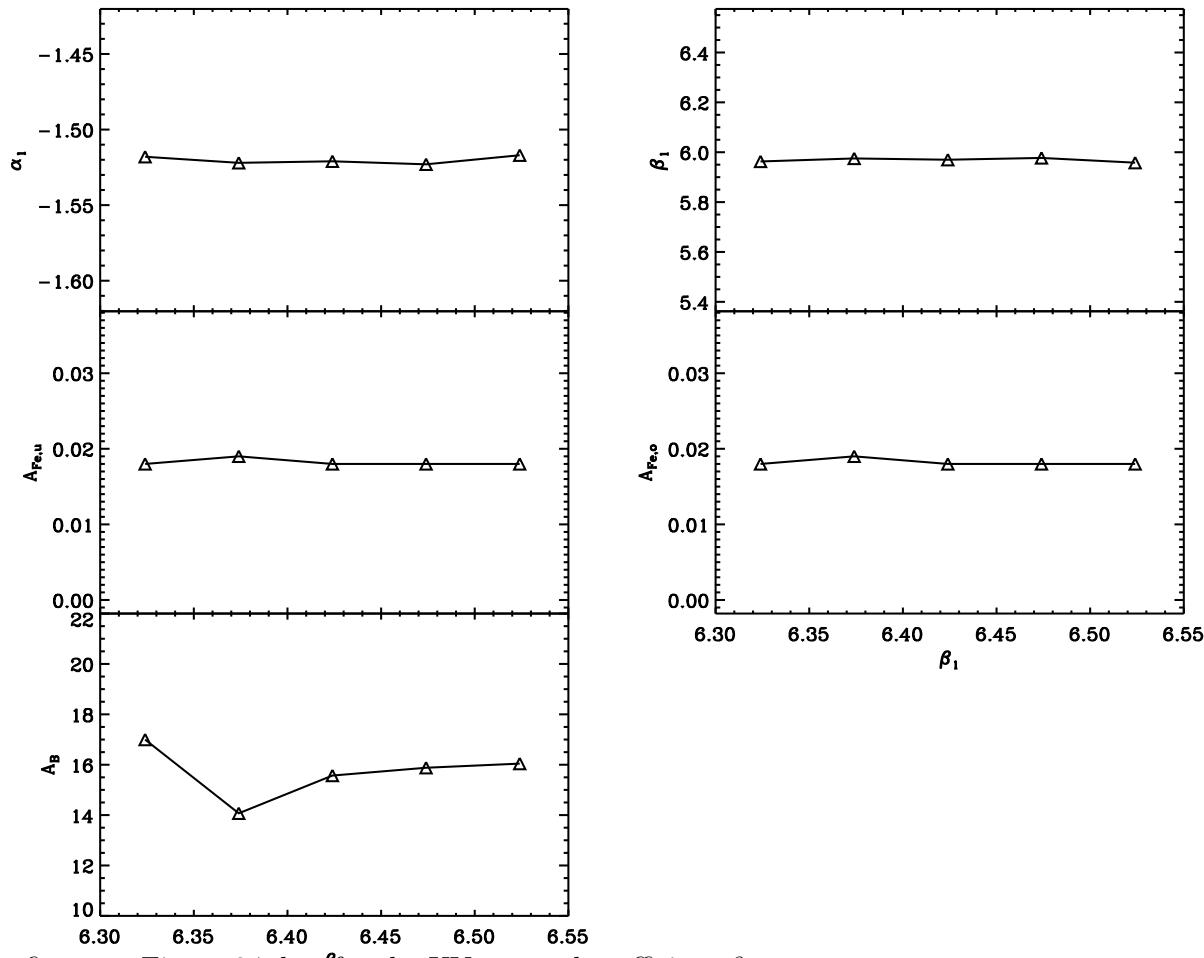


Fig. 2.8 The same figure as Figure 2.7 but for the UV spectral coefficient  $\beta_1$ .

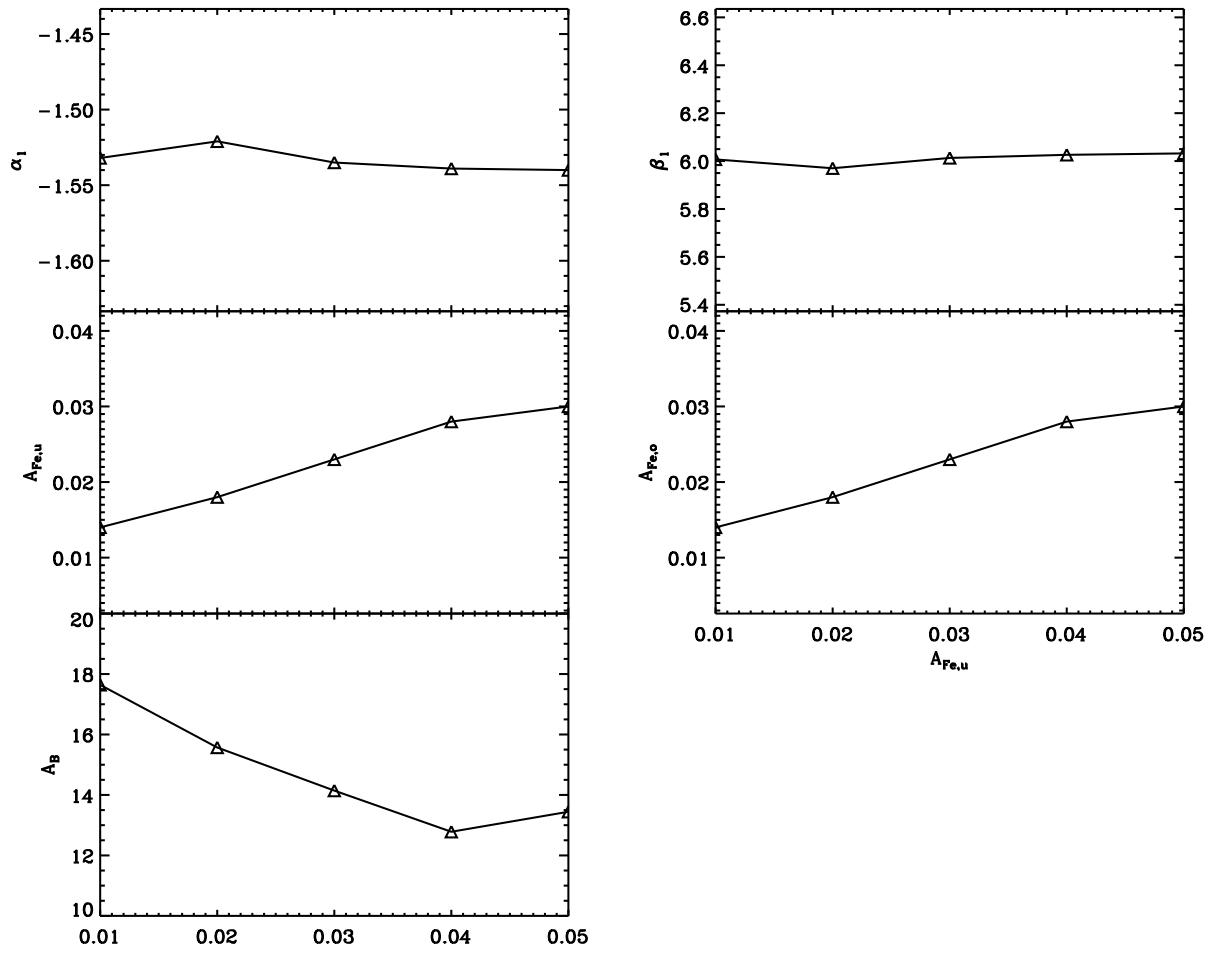


Fig. 2.9 The same figure as Figure 2.7 but for the UV spectral coefficient  $A_B$ .

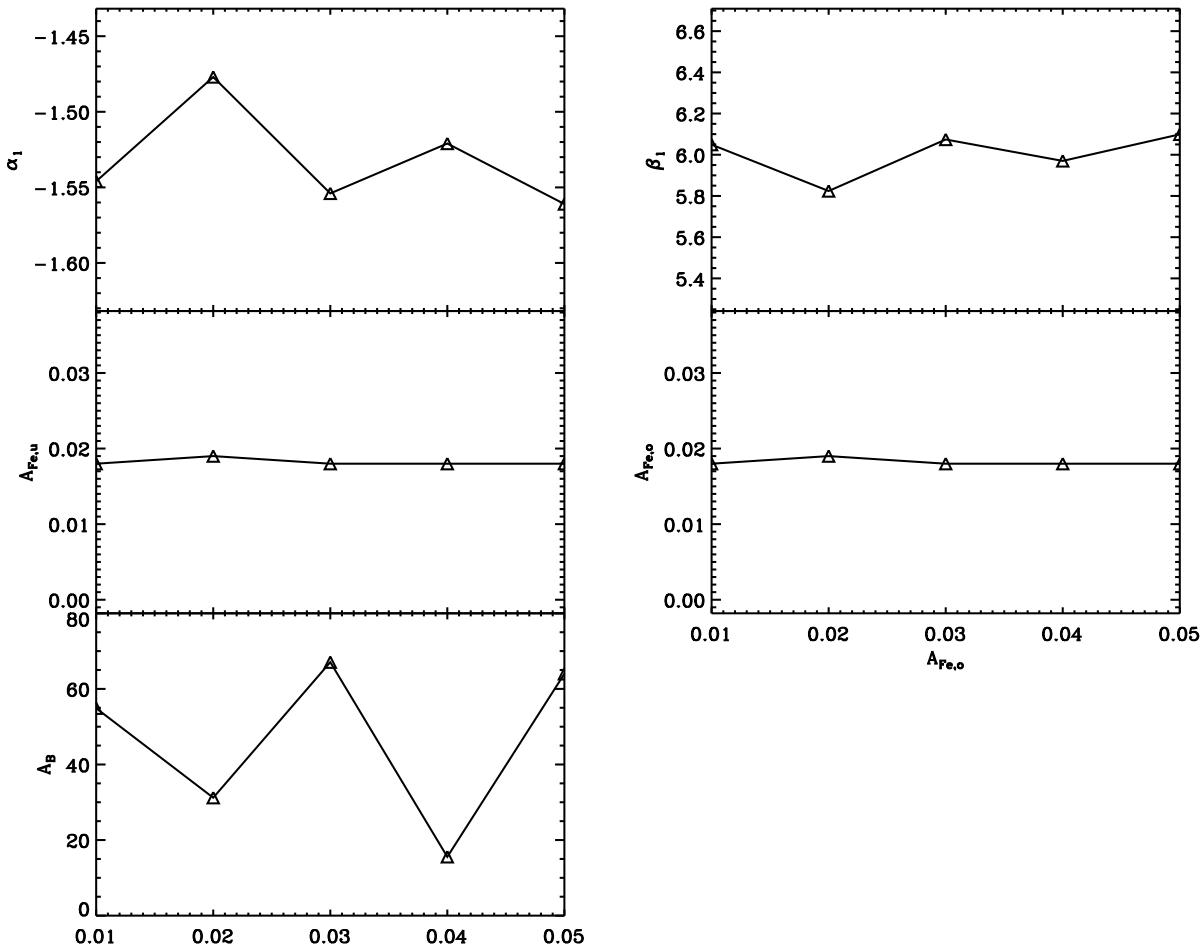


Fig. 2.10 The same figure as Figure 2.7 but for the UV spectral coefficient  $A_{Fe,u}$ .

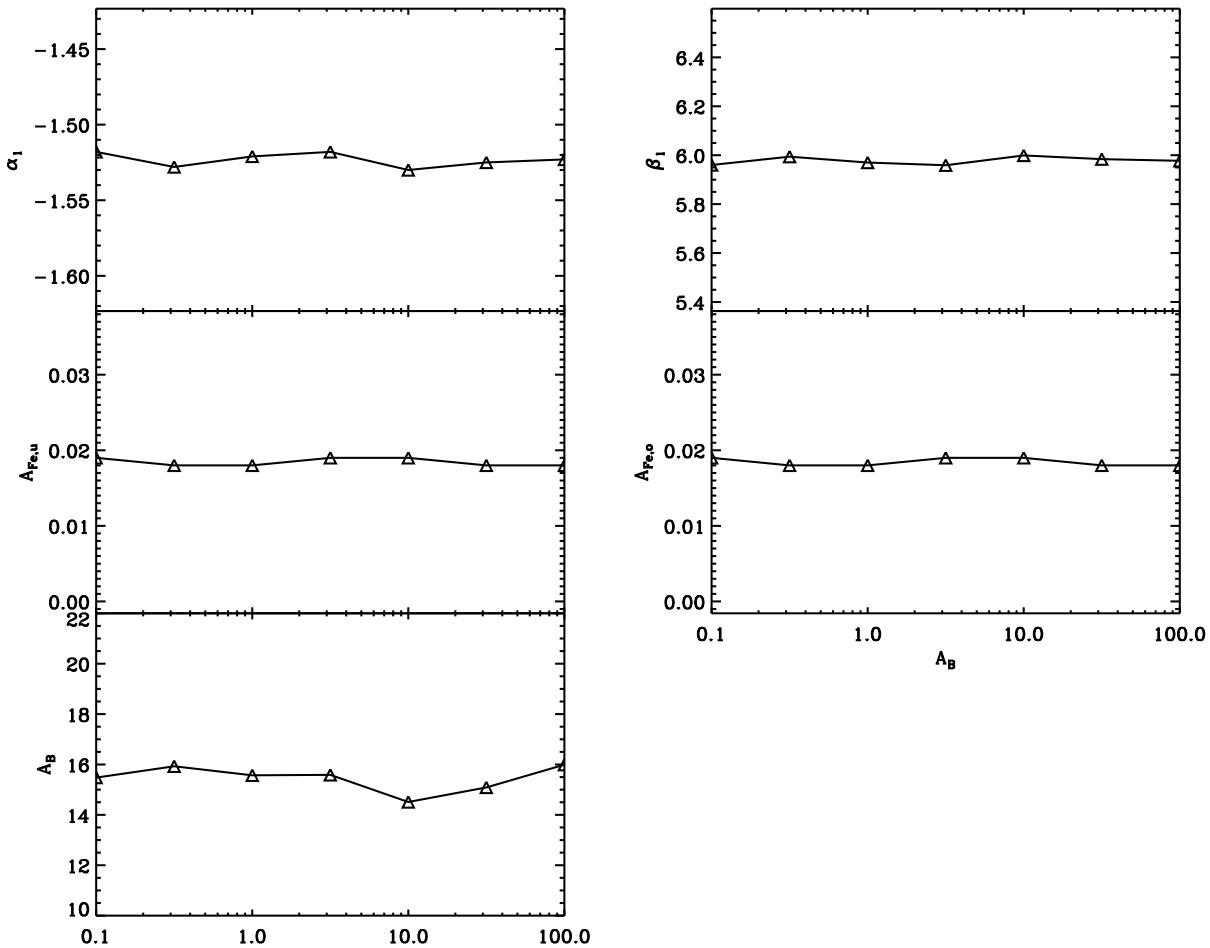


Fig. 2.11 The same figure as Figure 2.7 but for the UV spectral coefficient  $A_{Fe,o}$ .

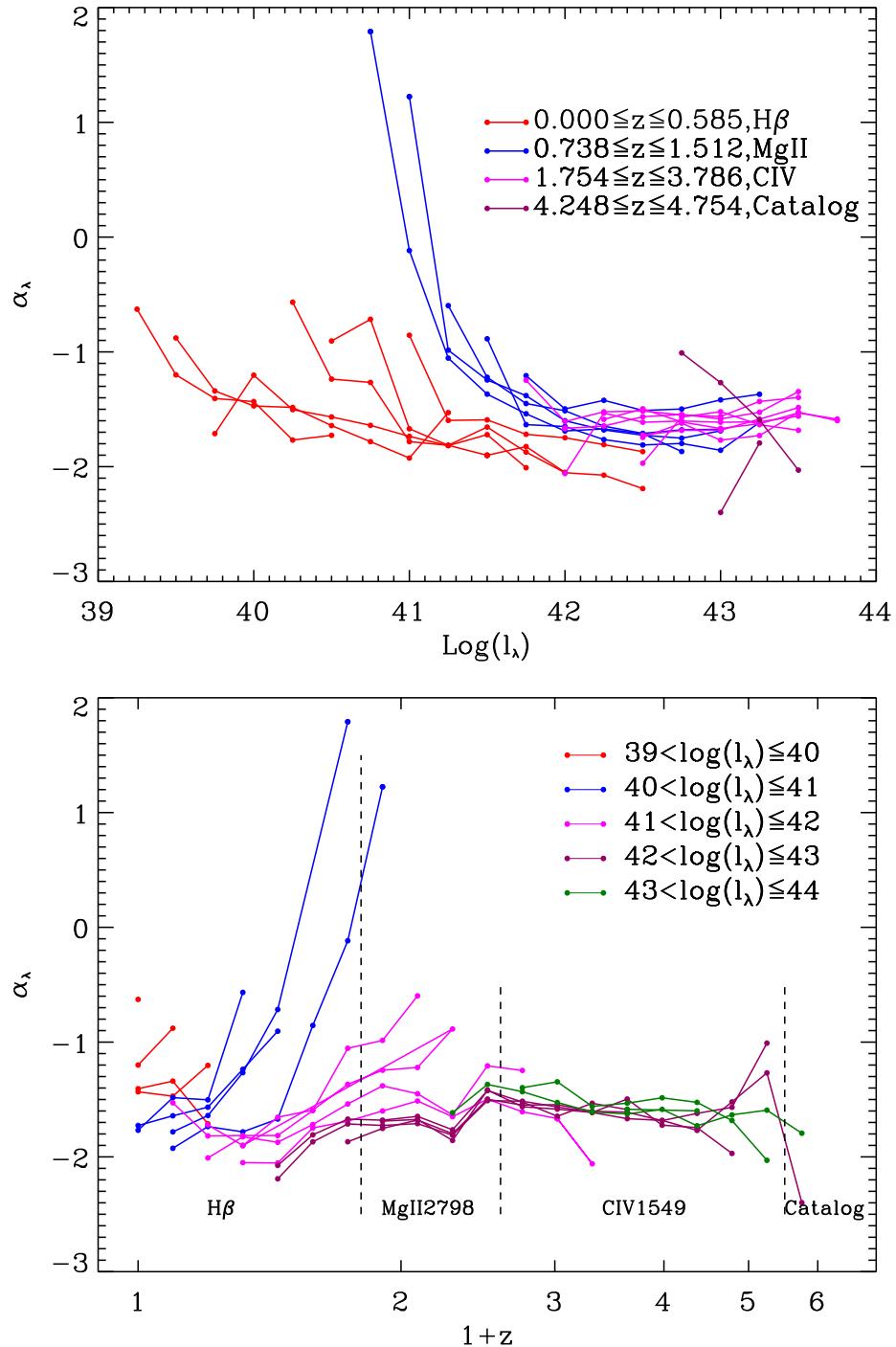


Fig. 2.12 *Upper panel*: The UV spectral indices as a function of luminosity. The redshifts of red dots are identified by H $\beta$ ; blue dots by Mg II  $\lambda 2798$ ; dots by C IV  $\lambda 1549$ ; and purple ones from the SDSS quasar catalog. Dots connected by straight lines have the same luminosity. *Lower panel*: The UV/optical spectral indices as a function of redshift. Dots connected by straight lines have the same luminosity ranges. The luminosity ranges represented by different colors are shown at the upper right corner of the plot. The black dashed lines divide data points whose redshift values are measured from different reference lines (H $\beta$ , Mg II  $\lambda 2798$ , C IV  $\lambda 1549$ ) or from SDSS quasar catalog.

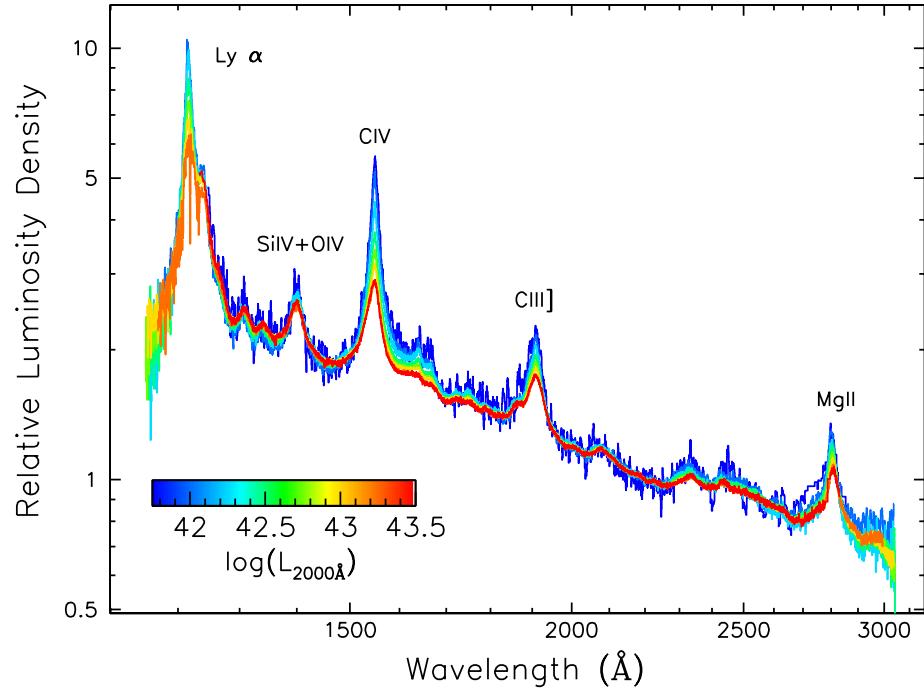


Fig. 2.13 A set of composite quasar spectra at  $z = 2.020$ . Spectra are color-coded according to their specific luminosity at  $2000 \text{ \AA}$ , shown as the color-bar in the legend. The flux densities have been normalized at  $2000 \text{ \AA}$ . The BEff is clearly present for a number of emission lines such as Ly $\alpha$ , C IV  $\lambda 1549$ , C III  $\lambda 1908$ , and Mg II  $\lambda 2798$ .

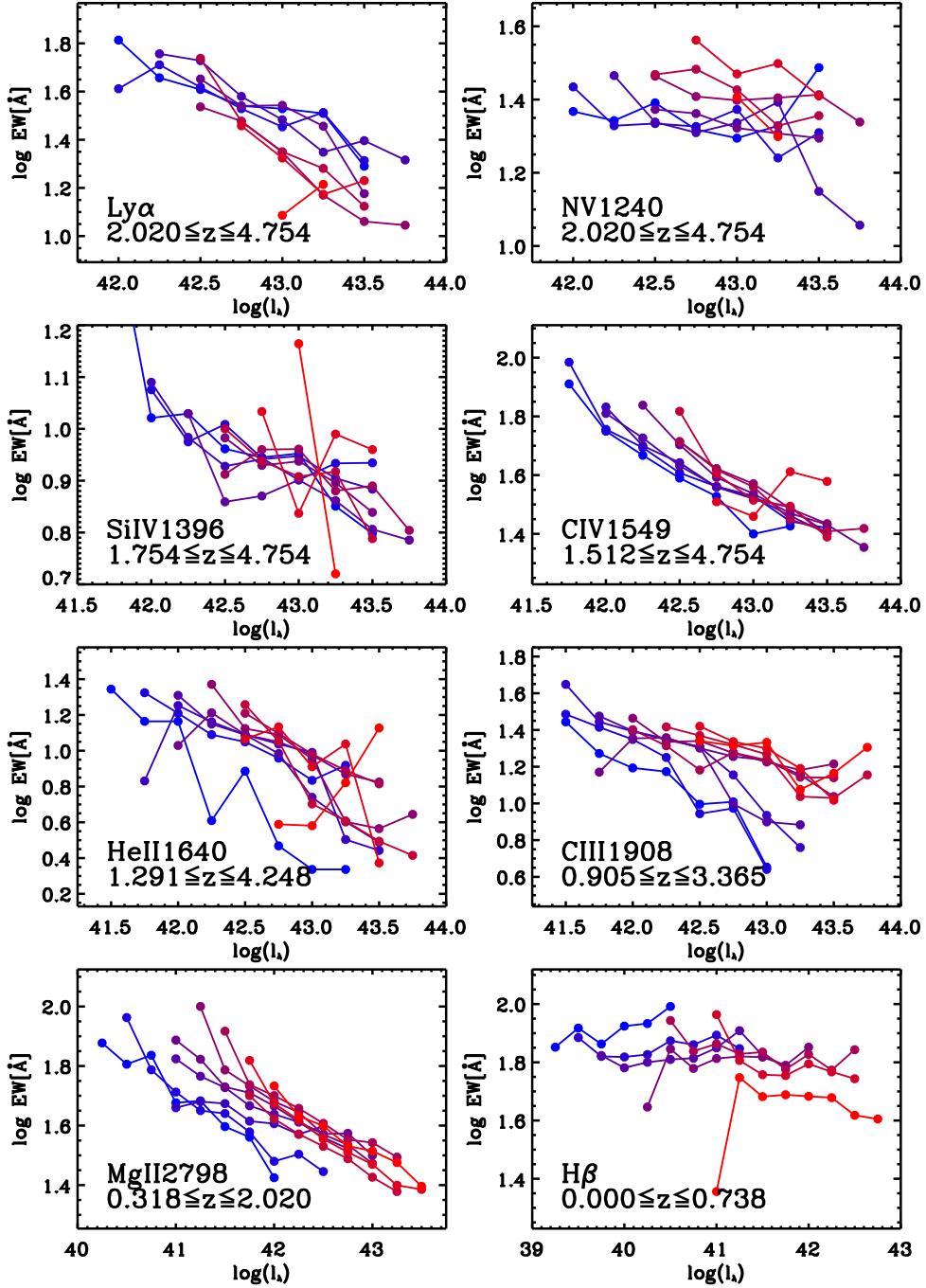


Fig. 2.14 Plots of  $\log EW$  (rest-frame) against  $\log l_{\lambda}$  for eight emission lines. The redshift changes from low to high as the color changes from blue to purple to red. Note that different emission lines have different redshift ranges. The BEff is clearly seen in all emission lines shown except  $H\beta$  and  $N\gamma \lambda 1239$ . The scale factors of the BEff also change with redshift for emission lines such as  $C\ IV\ \lambda 1549$  and  $Mg\ II\ \lambda 2798$ .

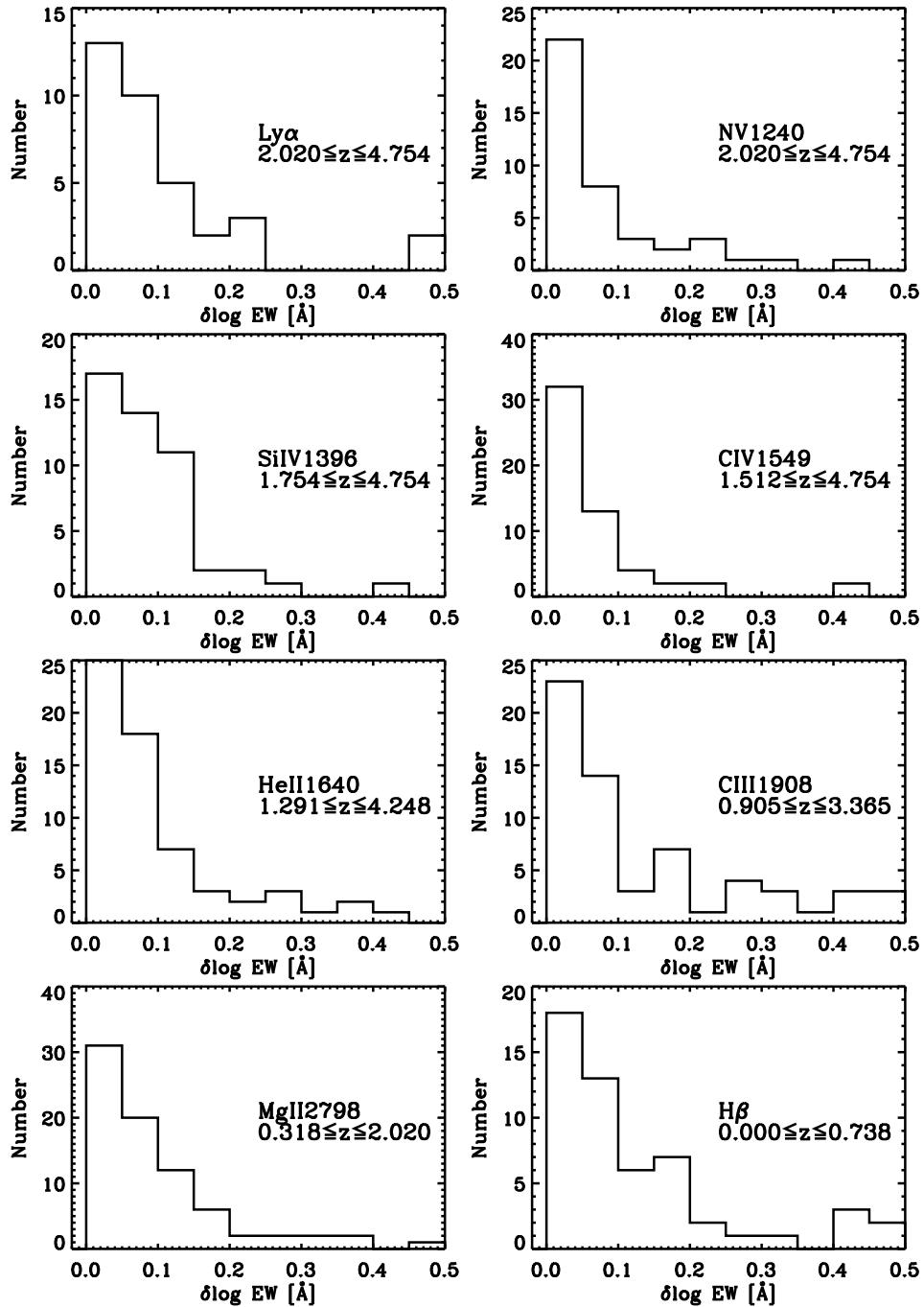


Fig. 2.15 Distributions of EW uncertainties for multiple emission lines in Figure 2.14. In each panel, the errors are evaluated to all data points in the corresponding panel of Figure 2.14. The typical error bar for  $\delta \log E\lambda$  is  $\sim 0.1$ .

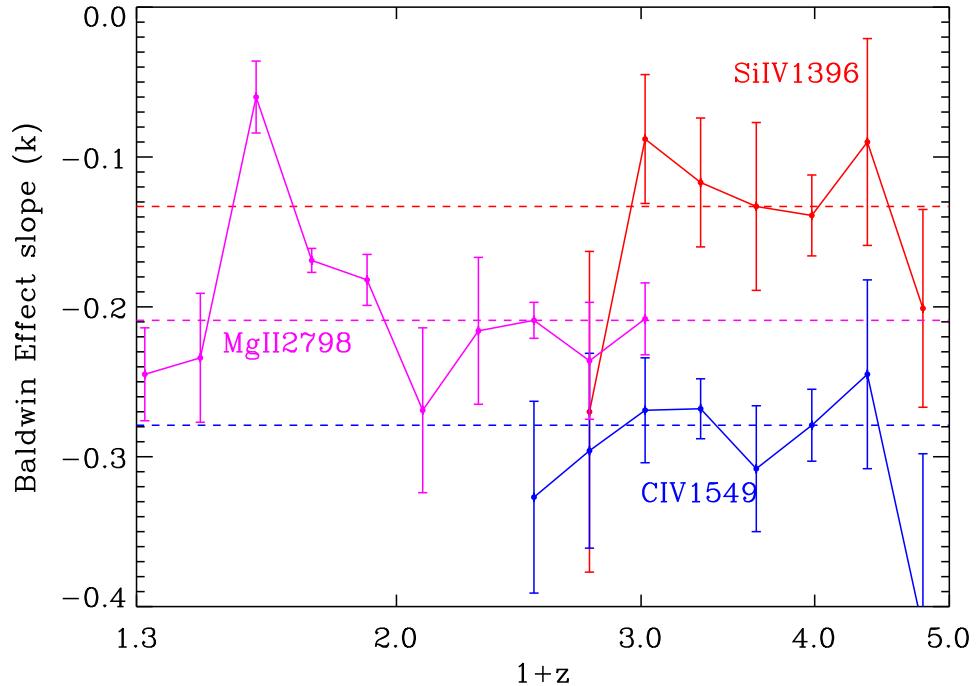


Fig. 2.16 Evolution of the BEff slopes  $k$  as a function of redshift for Mg II  $\lambda 2798$ , C IV  $\lambda 1549$ , and SiIV  $\lambda 1396$  represented by filled triangles, squares and circles, respectively. The slope values are obtained by fitting BEff curves for these three emission lines presented in Figure 2.14 in form of  $\log(EW/EW_0) = k \log(l_\lambda/l_\lambda^0) + b$ , in which  $EW_0 = 10 \text{ \AA}$  and  $l_\lambda^0 = 10^{42} \text{ erg s}^{-1}$ . Dashed lines are median values of BEff slopes for these three emission lines. There appears to be no evidence for any redshift evolution. The positive BEff slope for C IV at  $z = 4.248$  is not plotted because it is based upon a fit of only four data points.

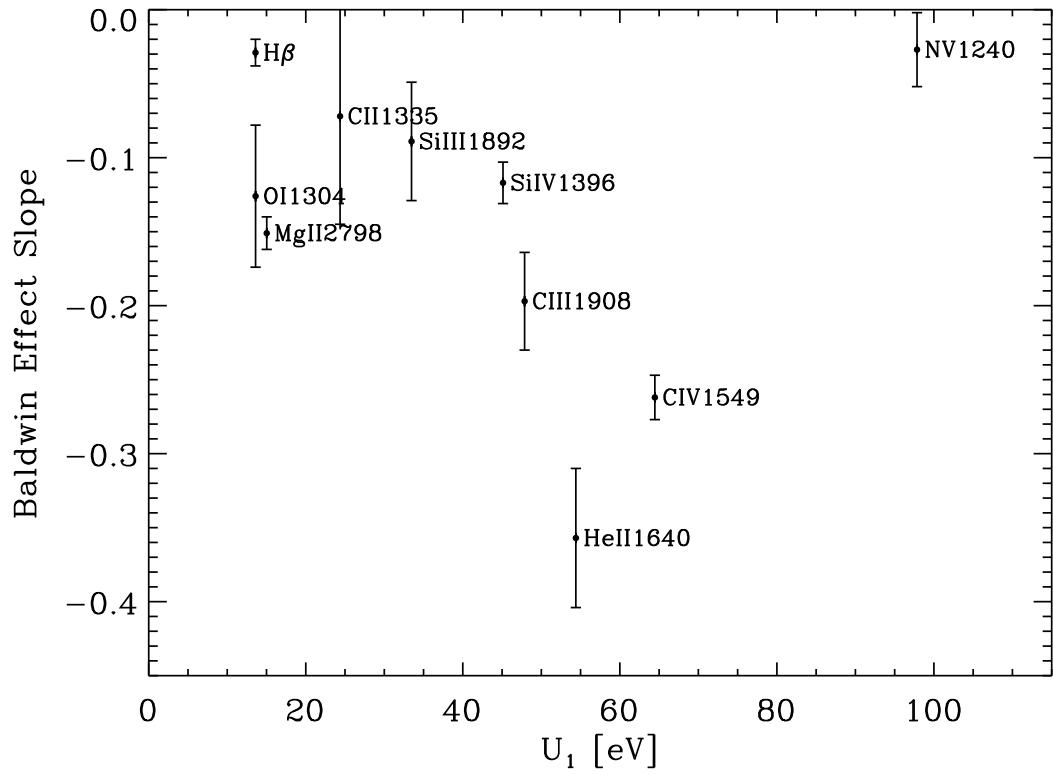


Fig. 2.17 BEff slopes vs. ionization potentials  $U_1$  for a number of prominent broad emission lines. The BEff slopes are *average* slopes by fitting all data points ( $\log EW$  vs.  $\log l_\lambda$ ) for a given emission line. We do not plot Ly $\alpha$  because of poor statistics (small luminosity range).

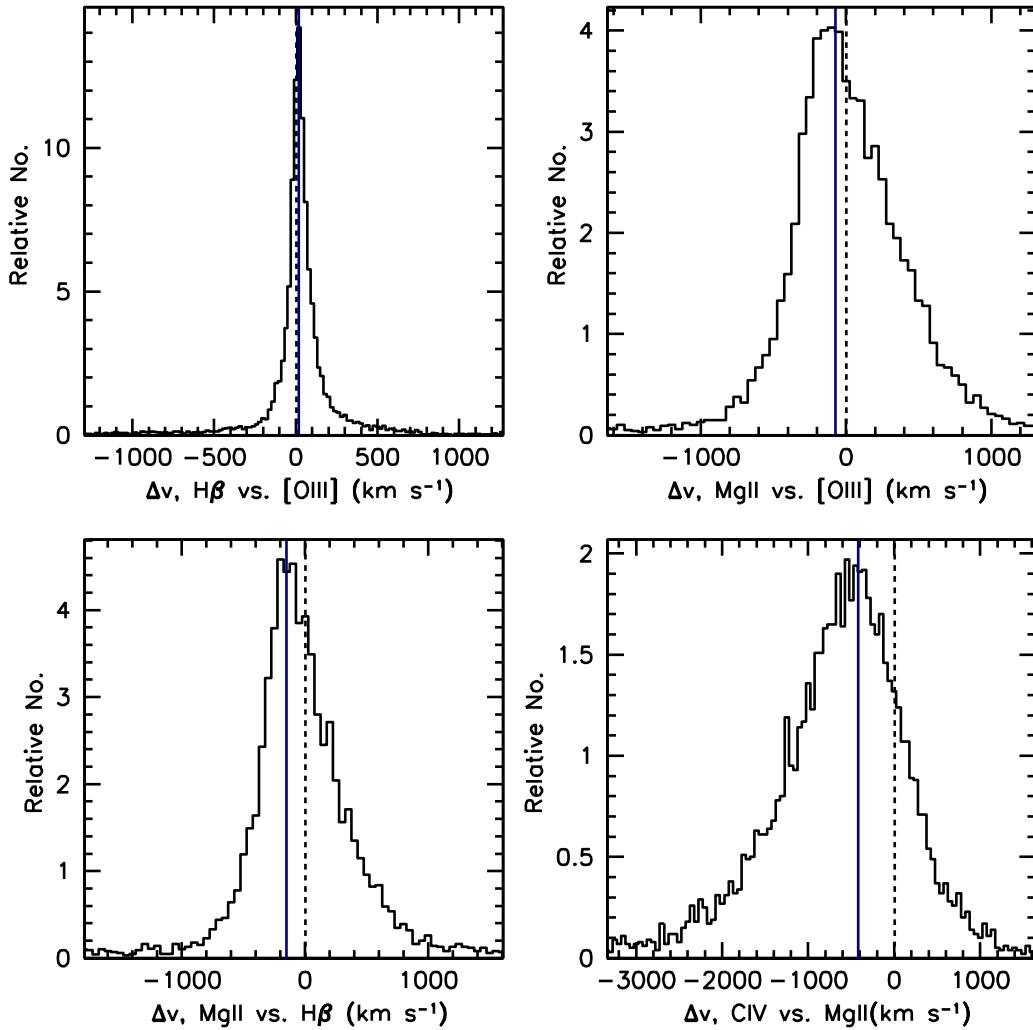


Fig. 2.18 Distributions of emission line shifts with respect to [O III]  $\lambda 5007$ , H $\beta$  and Mg II  $\lambda 2798$  emission lines. Emission line centers are measured using eigenspectrum decomposition technique to all quasar spectra covering both emission lines displayed in each panel. The  $x$ -axis is the relative emission line shift in km s $^{-1}$ ; the  $y$ -axis is the renormalized number. The dotted line in each panel is the position of zero emission shift; the solid blue line in each panel is the peak position for each distribution histogram.

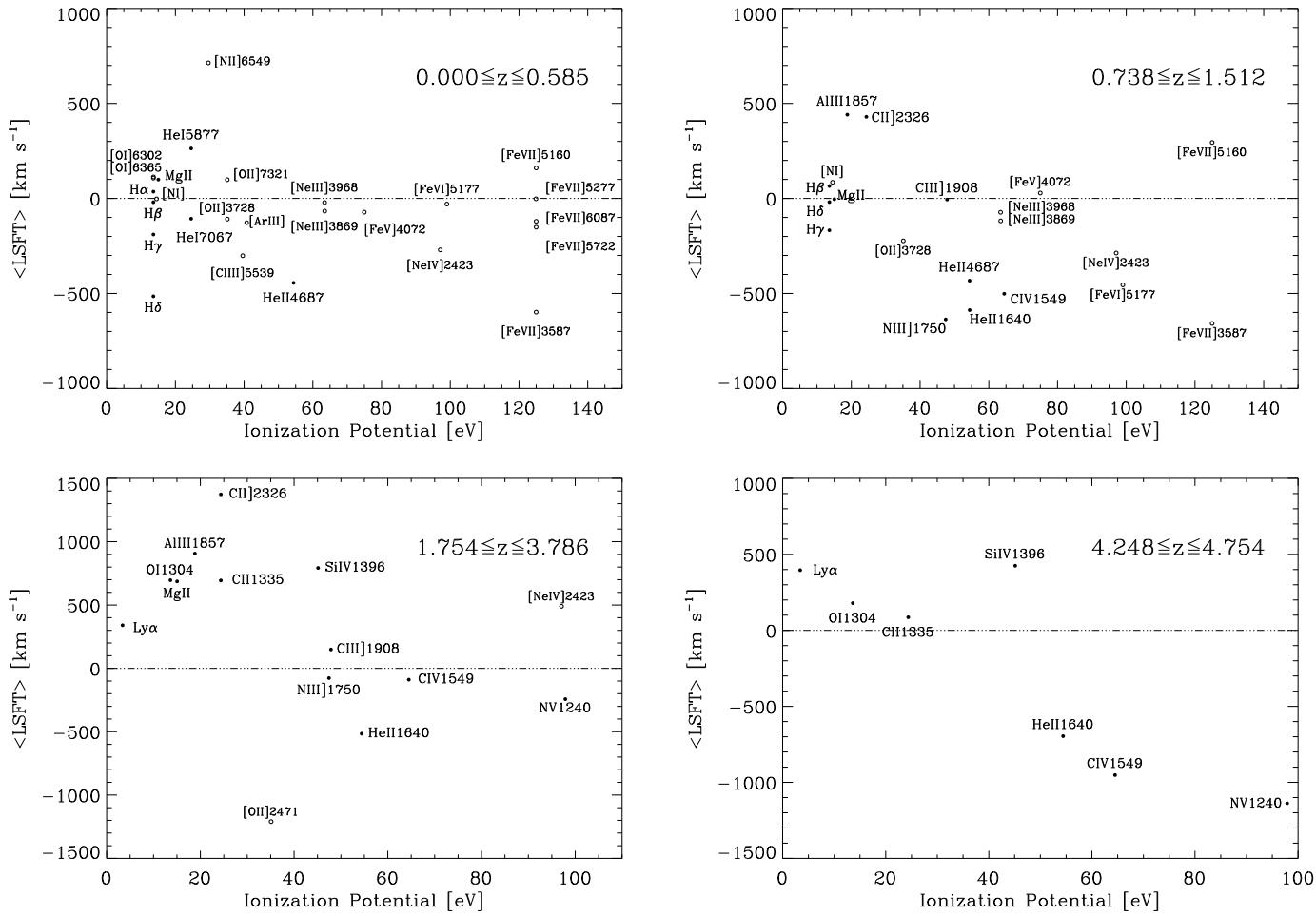


Fig. 2.19 Plots of mean line-shift ( $\langle \text{LSFT} \rangle$ ) against ionization potential for emission lines available in the spectra within particular redshift ranges. The mean value for a particular emission line is taken by averaging all the line-shifts within a redshift range shown in each panel. Some emission line notations are abbreviated to save space: [ArIII] for  $[\text{ArIII}] \lambda 7137$ , [Ni] for  $[\text{Ni}] \lambda 5200$  and Mg II for  $\text{Mg II} \lambda 2798$ . Filled circles are permitted and semi-forbidden lines; open circles are forbidden lines.

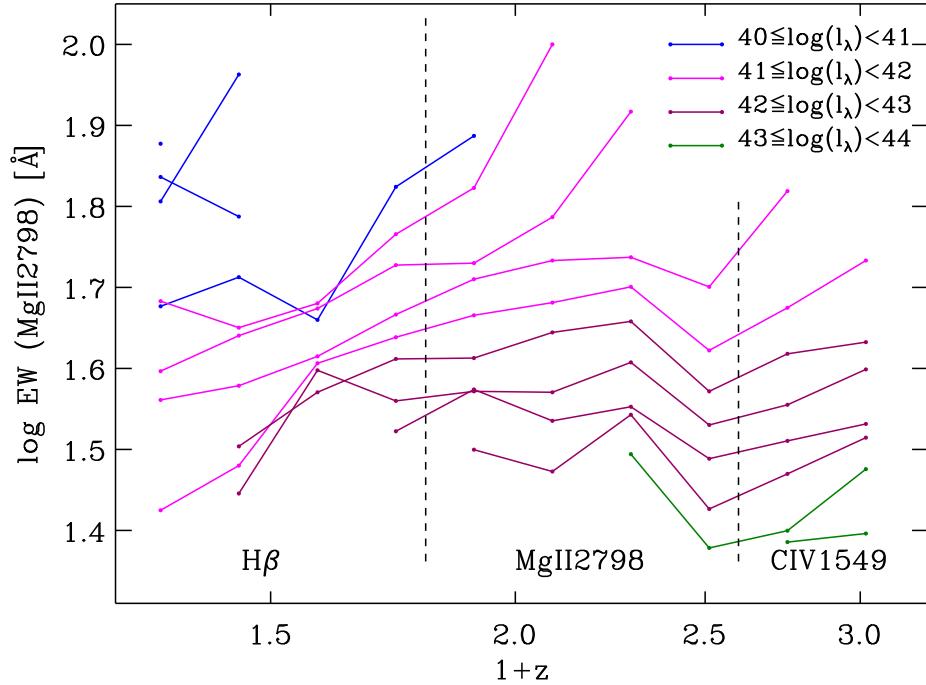


Fig. 2.20 The rest-frame EW for Mg II  $\lambda 2798$  as a function of redshift for different luminosities. The points are color-coded by the luminosity ranges indicated at the upper right corner. We divide the redshift ranges based on the reference lines used to measure redshifts.

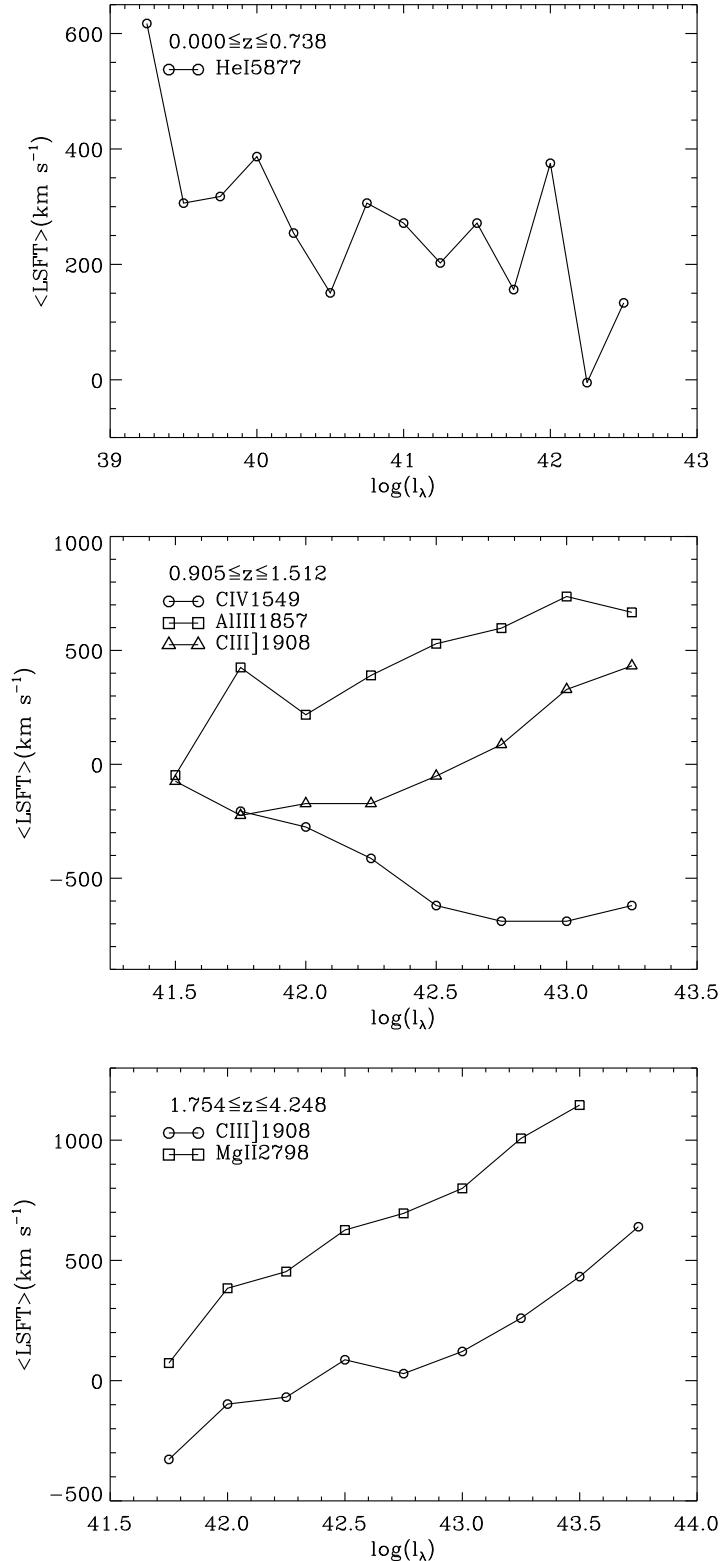


Fig. 2.21 Plots of mean line-shift ( $\langle \text{LSFT} \rangle$ ) over a restricted redshift ranges against luminosity. Emission lines are symbol coded as indicated in the legends. The three panels correspond to the three redshift bands in which different emission lines are used as redshift references. The lines not presented do not exhibit significant line-shift changes with luminosity.

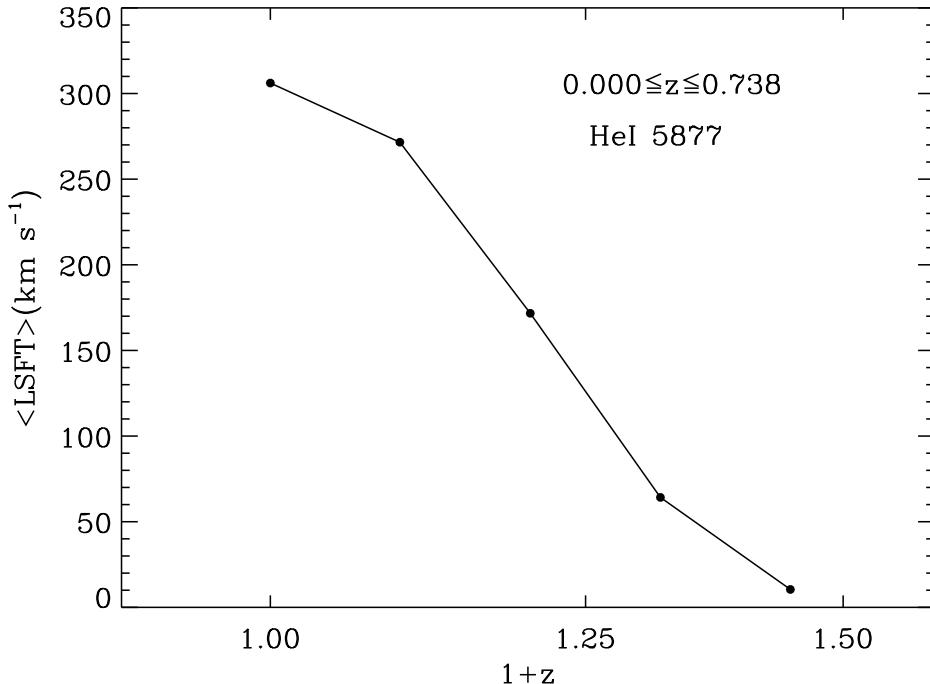


Fig. 2.22 The plot of mean line-shift ( $\langle \text{LSFT} \rangle$ ) with respect to H $\beta$  over the available luminosities in a redshift group as a function of redshift for the He I  $\lambda 5877$  emission line. This is the only emission feature that displays a  $\langle \text{LSFT} \rangle$ –redshift dependence.

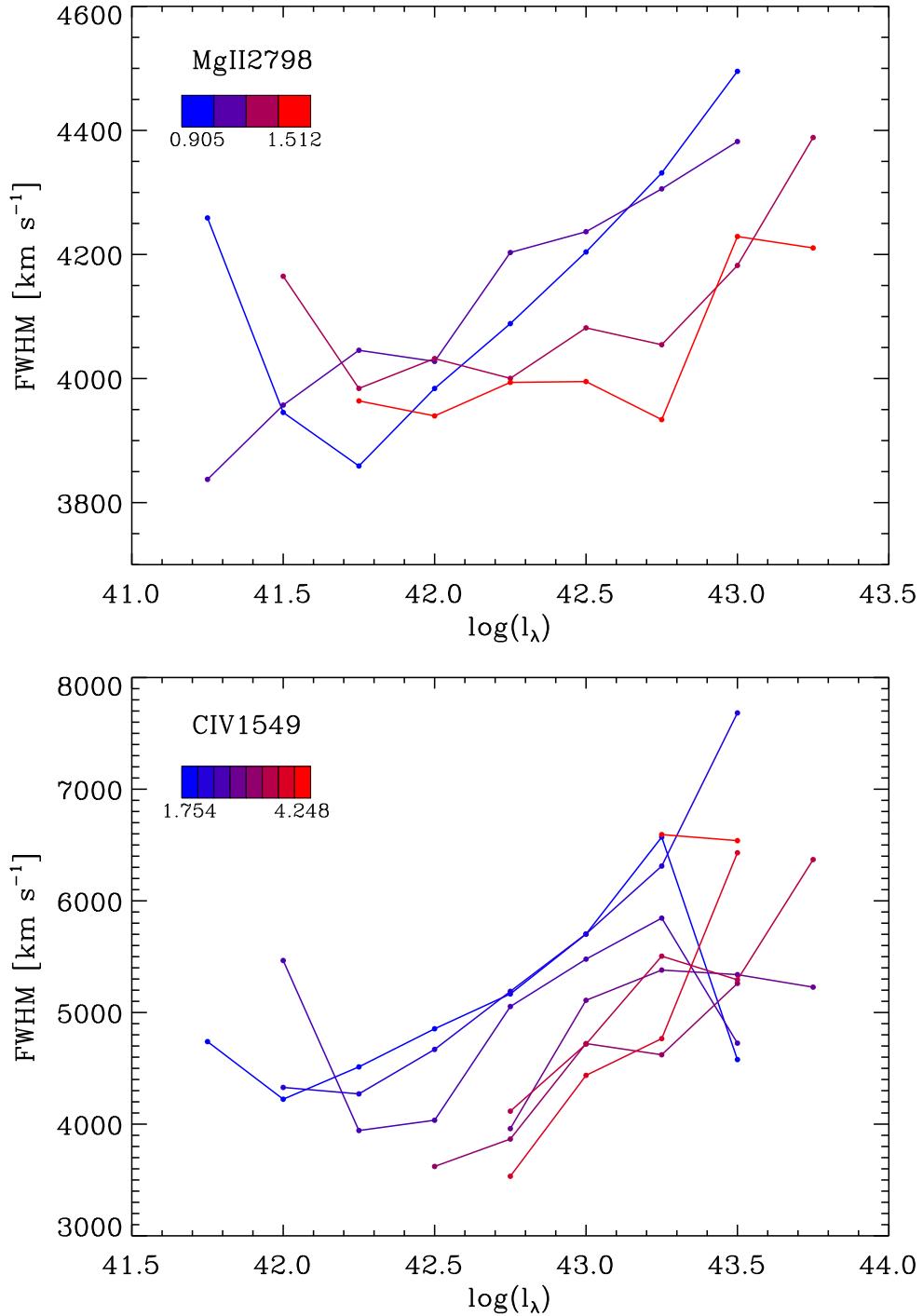


Fig. 2.23 Plots of FWHM as a function of luminosity for Mg II  $\lambda 2798$  and C IV  $\lambda 1549$ . The data points and connecting lines are color-coded. The color changes from blue to purple to red as redshift increases. The lowest and highest redshifts are labeled with color bars in each panel. Reference lines used to measure redshifts are indicated in the legends.

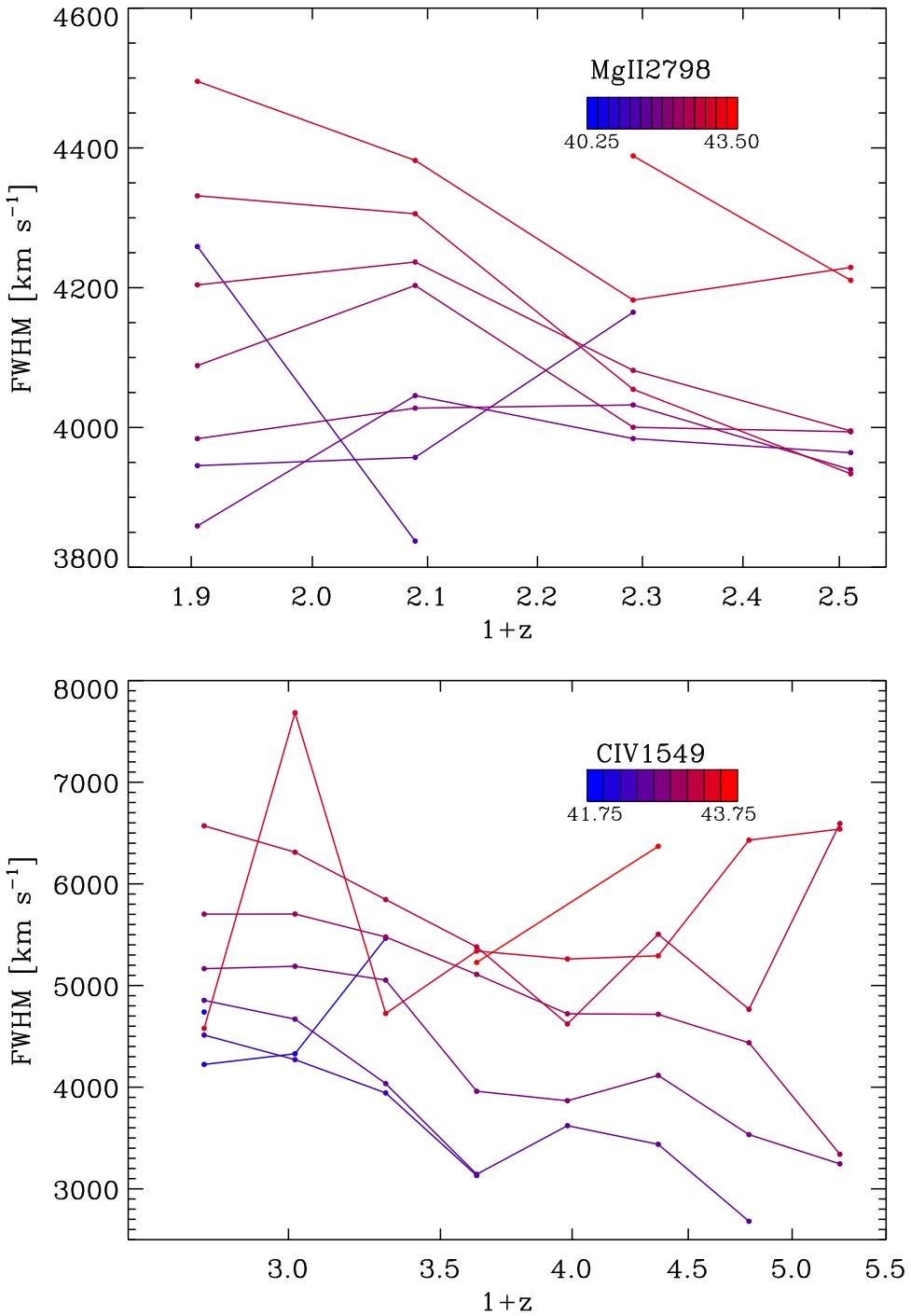


Fig. 2.24 Plots of FWHM as a function of redshift ( $1+z$ ) for MgII  $\lambda 2798$  and CIV  $\lambda 1549$ . The data points are color-coded. The color changes from blue to purple to red as luminosity increases. The lowest and higher luminosities are labeled with color bars in each panel.

Table 2.1. Empirical line-free windows used to estimate the power-law component.

Priority	Wavelength Windows (Å)
2	1282–1291
2	1350–1365
1	1445–1455
2	1452–1461
2	1695–1705
2	2048–2058
1	2196–2202
2	3801–3818
1	4198–4218
2	4209–4219
2	5095–5105
1	5630–5640
2	6145–6155
2	6895–6905
1	6935–6945
2	7595–7605

Table 2.2. Benchmark wavelengths used to evaluate the goodness of fit.

Waveband	Benchmark Wavelengths
UV	1327.0, 1363.0, 1444.8, 1469.2, 1693.0, 1729.6, 1772.1, 1989.4 2016.4, 2026.0, 2036.8, 2062.1, 2142.7, 2173.1, 2197.5, 2202.1 2222.4, 2242.7, 2274.6, 2528.2, 2538.6, 2612.5, 2621.3, 2714.4 2876.8, 2912.6, 2947.3, 3025.7, 3605.7, 3621.5, 3623.1, 3747.8 4037.8, 4200.0, 4480.3, 4520.3, 5090.0, 5150.3, 5304.6, 5419.1 5460.0
Optical	5614.0, 5648.4, 5790.0, 6071.7, 7546.2

Table 2.3. Underlying continuum parameters in joint fitting phase.

Parameter	Initial Values	Allowed Range
$\alpha_1$	power-law estimation $\alpha_{1,0}$	$[\alpha_{1,0} - 1.0, \alpha_{1,0} + 1.0]$
$\beta_1$	power-law estimation $\beta_{1,0}$	$[\beta_{1,0} - 1.0, \beta_{1,0} + 1.0]$
$\alpha_2$	power-law estimation $\alpha_{2,0}$	$[\alpha_{2,0} - 1.0, \alpha_{2,0} + 1.0]$
$\beta_2$	power-law estimation $\beta_{2,0}$	$[\beta_{2,0} - 1.0, \beta_{2,0} + 1.0]$
$A_B$	1.0	$\geq 0.0$
$A_{\text{Fe},u}$	0.02	$> 0.0$
$A_{\text{Fe},o}$	0.04	$\geq 0.0$
$v_{\text{Fe},u}$	3000	$1250\text{--}10,000 \text{ km s}^{-1}$
$v_{\text{Fe},o}$	3000	$1250\text{--}10,000 \text{ km s}^{-1}$
$\delta z_u$	0.0	$\geq 0$
$\delta z_o$	0.0	$\geq 0$

Table 2.4. Quasar emission line list.

Group Name <sup>1</sup>	Emission Line	$\lambda_{\text{lab}}(\text{\AA})$	Fitting Region (\text{\AA})	$N_G$ <sup>2</sup>	Continuum Type <sup>3</sup>
Ly $\alpha$	Ly $\alpha$	1215.67	1207.05–1254.03	2	G
Ly $\alpha$	N V	1240.14	1207.05–1254.03	1	G
	O I	1304.35	1291.96–1322.85	1	G
	C II	1335.30	1320.35–1360.76	1	L
	Si IV	1396.76	1370.56–1437.23	2	G
	C IV	1549.06	1490.56–1590.48	2	G
He II	He II	1640.42	1610.56–1690.48	1	G
He II	O III]	1663.48	1610.56–1690.48	1	G
	N III]	1750.26	1710.90–1780.21	1	L
Al III	Al III	1857.40	1850.03–1958.03	1	G
Al III	Si III]	1892.03	1850.03–1958.03	1	G
Al III	C III]	1908.73	1850.03–1958.03	2	G
	C II]	2326.44	2300.11–2370.68	1	L
[Ne IV]		2423.83	2380.85–2470.07	2	L
[O II]		2471.03	2450.85–2500.07	1	L
[Al II]		2669.95	2645.23–2710.83	1	L
Mg II		2798.75	2741.42–2860.96	2	G
[Fe VII]		3587.34	3550.08–3615.22	1	L
[O II]		3728.48	3700.22–3750.02	1	L
[Fe VII]		3759.99	3732.02–3788.68	1	L
[Ne III]		3869.85	3850.94–3890.96	2	L
He I		3889.74	3870.94–3909.96	2	L
[Ne III]		3968.58	3920.60–4020.71	2	L
H $\delta$	[Fe V]	4072.39	4041.51–4170.53	1	L
H $\delta$	H $\delta$	4102.89	4041.51–4170.53	2	L
H $\gamma$	H $\gamma$	4341.68	4256.24–4500.57	2	L
H $\gamma$	[O III]	4364.44	4256.24–4500.57	2	L
	He II	4687.02	4672.54–4715.40	2	L
H $\beta$	H $\beta$	4862.68	4700.05–5150.11	3	L
H $\beta$	[O III]	4960.30	4700.05–5150.11	2	L
H $\beta$	[O III]	5008.24	4700.05–5150.11	3	L
[Fe VI]	[Fe VII]	5160.33	5100.82–5230.24	1	L
[Fe VI]	[Fe VI]	5177.48	5100.82–5230.24	1	L
[Fe VI]	[N I]	5200.53	5100.82–5230.24	1	L
	[Fe VII]	5277.85	5260.29–5295.40	2	L
	[Fe XIV]	5304.34	5295.29–5335.40	1	L
	[Cl III]	5539.43	5490.64–5580.28	1	L
	[Fe VII]	5722.30	5680.09–5751.62	1	L
	He I	5877.29	5766.58–6063.46	2	L
	[Fe VII]	6087.98	6050.46–6138.82	2	L
	[O I]	6302.05	6279.82–6335.20	1	L
	[O I]	6365.54	6340.22–6400.80	2	L
H $\alpha$	H $\alpha$	6564.61	6450.92–6700.80	3	G
	He I	7067.20	7000.88–7130.38	2	L
	[Ar III]	7137.80	7100.88–7170.38	1	L
	[O II]	7321.48	7261.38–7388.93	1	L

<sup>1</sup>We use a prominent emission line to name a line group. Emission lines in the same group are fit simultaneously.<sup>2</sup>The number of Gaussian profiles.<sup>3</sup>Continuum type: “G” = global continuum and “L” = local continuum.

Table 2.5. Varying initial values of continuum fitting parameters.

Parameter	Initial values
$\alpha_1$	$\alpha_{1,0} + [-0.1, -0.05, 0, 0.05, 0.1]^1$
$\beta_1$	$\beta_{1,0} + [-0.1, -0.05, 0, 0.05, 0.1]^1$
$\log A_B$	$[-0.1, -0.5, 0, 0.5, 1.0, 1.5, 2.0]$
$A_{\text{Fe,u}}$	$[0.01, 0.02, 0.03, 0.04, 0.05]$
$A_{\text{Fe,o}}$	$[0.01, 0.02, 0.03, 0.04, 0.05]$

<sup>1</sup> $\alpha_{1,0}$  and  $\beta_{1,0}$  come from power-law estimation (see Table 2.3).

Table 2.6. BEff slopes and intercepts for three emission lines.

Emission line	Redshift	Slope ( $k$ )	Intercept ( $b$ )
C IV $\lambda 1549$	1.512	$-0.327 \pm 0.064$	$0.781 \pm 0.036$
	1.754	$-0.296 \pm 0.065$	$0.817 \pm 0.062$
	2.020	$-0.269 \pm 0.035$	$0.798 \pm 0.037$
	2.311	$-0.268 \pm 0.020$	$0.793 \pm 0.020$
	2.631	$-0.308 \pm 0.042$	$0.875 \pm 0.053$
	2.981	$-0.279 \pm 0.024$	$0.848 \pm 0.026$
	3.365	$-0.245 \pm 0.063$	$0.812 \pm 0.062$
	3.786	$-0.415 \pm 0.117$	$0.994 \pm 0.144$
	4.248	$0.078 \pm 0.105$	$0.455 \pm 0.131$
Mg II $\lambda 2798$	0.318	$-0.245 \pm 0.031$	$0.465 \pm 0.034$
	0.445	$-0.234 \pm 0.043$	$0.528 \pm 0.020$
	0.585	$-0.060 \pm 0.024$	$0.614 \pm 0.012$
	0.738	$-0.169 \pm 0.008$	$0.645 \pm 0.004$
	0.905	$-0.182 \pm 0.017$	$0.677 \pm 0.009$
	1.089	$-0.269 \pm 0.055$	$0.718 \pm 0.029$
	1.291	$-0.216 \pm 0.049$	$0.738 \pm 0.036$
	1.512	$-0.209 \pm 0.012$	$0.637 \pm 0.011$
	1.754	$-0.236 \pm 0.039$	$0.707 \pm 0.035$
	2.020	$-0.208 \pm 0.024$	$0.712 \pm 0.025$
Si IV $\lambda 1396$	1.754	$-0.270 \pm 0.107$	$0.173 \pm 0.109$
	2.020	$-0.088 \pm 0.043$	$0.039 \pm 0.040$
	2.311	$-0.117 \pm 0.043$	$0.047 \pm 0.048$
	2.631	$-0.133 \pm 0.056$	$0.010 \pm 0.078$
	2.981	$-0.139 \pm 0.027$	$0.053 \pm 0.025$
	3.365	$-0.090 \pm 0.069$	$-0.006 \pm 0.086$
	3.786	$-0.201 \pm 0.066$	$0.106 \pm 0.048$
	4.248	$-0.118 \pm 0.250$	$0.117 \pm 0.336$

Note. — At each redshift, the BEff is fit linearly as  $\log(EW) - 1 = k(\log l_\lambda - 42) + b$  using Jackknife method (see Section 2.4.2.1).

Table 2.7. Baldwin Effect slope  $k$  for different emission lines as a function of ionization potentials.

$U_1$	Emission Line	$\langle k \rangle^1$
13.60	Ly $\alpha$	$-0.395 \pm 0.038$
13.60	H $\beta$	$-0.029 \pm 0.009$
15.03	Mg II $\lambda 2798$	$-0.151 \pm 0.011$
47.87	C III $\lambda 1908$	$-0.197 \pm 0.033$
45.12	Si IV $\lambda 1396$	$-0.117 \pm 0.014$
54.40	He II $\lambda 1640$	$-0.357 \pm 0.047$
64.46	C IV $\lambda 1549$	$-0.262 \pm 0.015$
97.85	N V $\lambda 1239$	$-0.227 \pm 0.025$

<sup>1</sup>Average BEff slope.

## Chapter 3

# Probing the Origins of the C IV and Fe K $\alpha$ Baldwin Effects

### 3.1 Introduction

One of the important properties of AGNs is the relation between the emission-line strength, characterized by the equivalent width (EW), and continuum luminosity, because it reveals that the regions emitting these two spectral components are associated. Baldwin (1977) found that the EW of C IV  $\lambda 1549$  (C IV) is inversely correlated with the quasar monochromatic luminosity at 1450 Å,  $l_\lambda(1450 \text{ \AA})$ , namely,  $\log \text{EW}(\text{C IV}) = k \log l_\nu(1450 \text{ \AA}) + b$ . Carswell & Smith (1978) referred to this trend as the “Baldwin Effect” (BEff), a designation now widely used to describe line strength-luminosity relations. Baldwin (1977) identified this relation using only 20 quasi-stellar objects with  $29.8 \lesssim \log l_\nu(1450 \text{ \AA}) \lesssim 32.0$  and  $1.24 < z < 3.53$ . Subsequent UV/optical surveys have enabled investigation of this relation with wider luminosity and redshift ranges (e.g., Kinney, Rivolo & Koratkar 1990; Zamorani et al. 1992). It has been found that the BEff exists for not only C IV but many other broad emission lines such as Ly $\alpha$ , C III]  $\lambda 1908$ , Si IV, Mg II (Dietrich et al. 2002; Vanden Berk 2011), UV iron emission lines (Green et al. 2001), and even forbidden lines such as [O II] and [Ne V]  $\lambda 3426$  (Croom et al. 2002). Applying a spectral-composite technique (Vanden Berk et al. 2001) to the Sloan Digital Sky Survey (SDSS; York et al. 2000) Data Release Three (DR3) quasar catalog (Schneider et al. 2005), Vanden Berk (2011) found that the BEff evolves with redshift, which is a source of scatter in this relation for a sample with a wide range of redshift.

The X-ray BEff (or Fe K $\alpha$  BEff), in which the EW of the narrow Fe K $\alpha$  line at 6.4 keV (hereafter abbreviated to Fe K $\alpha$ ) is anti-correlated with X-ray luminosity,  $l_\nu(2 \text{ keV})$ , was discovered in the early 1990s from observations by the X-ray observatory *Ginga* (Iwasawa & Taniguchi 1993). This relation has been subsequently confirmed using data from *ASCA* (Tanaka et al. 1994; Nandra et al. 1997) and from *Chandra* and *XMM-Newton* (Page et al. 2004a; Zhou & Wang 2005; Jiang et al. 2006; Bianchi et al. 2007). Possible sites of origin for narrow Fe K $\alpha$  emission include the broad emission-line region (BELR), the outskirts of the accretion disk, and the molecular torus (e.g., Weaver et al. 1992, Antonucci 1993; Krolik, Madau & Zycki 1994).

Although it is well accepted that the BEff exists for many UV/optical emission lines (e.g., Osmer & Shields 1999; Shields 2007), there is currently no theoretical model that provides a compelling and complete explanation of this well-known phenomenon. Several physical explanations have been proposed to account for the UV/optical BEff.

One promising explanation is that the continuum shape may be luminosity dependent. In this model, the UV/optical BEff is due to the softening of the spectral energy

distribution (SED) at high luminosity, which lowers the ion populations having high ionization potentials (Netzer et al. 1992; Korista et al. 1998). It has been found, using *Einstein Observatory* data, that the quasar SED, parameterized by  $\alpha_{\text{ox}}$ <sup>1</sup> (Tananbaum et al. 1979), depends on UV luminosity (e.g., Zamorani et al. 1981). Later research using radio-quiet (RQ) optically selected quasar samples from the SDSS Early Data Release (Stoughton et al. 2002) confirmed and extended this result (e.g., Vignali, Brandt & Schneider 2003). Using optically selected AGNs Strateva et al. (2005) and Steffen et al. (2006) firmly established the correlation of  $\alpha_{\text{ox}}$  with UV luminosity for these sources. The idea that the UV/optical BEff is attributable to SED-driven ionization effects is supported both observationally (Zheng & Malkan 1993; Green 1996, 1998) and theoretically (Netzer et al. 1992). Recent work on a sample of non-Broad Absorption Line (BAL), radio-quiet, optically selected quasars indicates that the EW of C IV depends both on UV and X-ray luminosity. The physics of the C IV BEff is apparently associated with both UV and X-ray emission (e.g., Gibson, Brandt & Schneider, 2008, hereafter GBS08, and references therein).

Other proposed BEff drivers include the Eddington ratio,  $L/L_{\text{Edd}}$  (Baskin & Laor 2004; Bachev et al. 2004; Warner et al. 2004; Zhou & Wang 2005), the black-hole mass (Netzer et al. 1992; Wandel et al. 1999; Shields 2007), and the luminosity dependence of metallicity (Warner et al. 2004).

In this paper we investigate the origin of the BEff for the C IV emission line in a sample of 272 Type 1 AGNs and quasars. Although C IV is not the only UV/optical broad emission line that exhibits a BEff, we selected it to study this phenomenon not only because it is a representative and well-accepted BEff emission line, but also because C IV resides in a relatively clean spectral region where the local continuum can be well approximated as a single power-law, with few blends with other emission lines (in particular the iron emission forest) and limited contamination from the AGN host galaxy. These properties make it relatively straightforward to perform spectral fitting and obtain accurate emission-line parameters for C IV. We also use partial-correlation analysis (PCA) and linear-regression regression analysis to investigate the correlations between EW, monochromatic luminosity, and  $\alpha_{\text{ox}}$  for C IV and narrow Fe K $\alpha$  emission lines.

Over the past three decades, there have been a large number of studies of the BEff. Our work on the C IV/Fe K $\alpha$  lines combines the following important features (1) a wide range in redshift ( $0.009 \lesssim z \lesssim 4.720$ ) and luminosity ( $27.81 \lesssim \log l_{\nu}(2500 \text{ \AA}) \lesssim 33.04$ ) that allows one to disentangle evolutionary vs. luminosity, so that we are not narrowing our study for quasars with a particular luminosity or at a certain redshift; (2) a relatively large sample size (272 objects); (3) the use of partial correlation analysis; and (4) a high X-ray detection rate ( $\sim 94\%$ ); and (5) estimates of the effects of observational errors and object variability.

---

<sup>1</sup>Defined as  $\alpha_{\text{ox}} = \log [l_{\nu}(2 \text{ keV})/l_{\nu}(2500 \text{ \AA})] / \log [\nu(2 \text{ keV})/\nu(2500 \text{ \AA})] = 0.3838 \log [l_{\nu}(2 \text{ keV})/l_{\nu}(2500 \text{ \AA})]$ .  $\alpha_{\text{ox}}$  is used to characterize the spectral hardness in the UV to X-ray band (e.g., Avni & Tananbaum 1982, 1986; Anderson & Margon 1987; Wilkes et al. 1994; Vignali, Brandt & Schneider 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007).

We describe the sample selection in § 5.2 and the methods used to process the data in § 3.3. In § 3.4, we perform partial-correlation and linear-regression analyses to investigate the roles of  $\alpha_{\text{ox}}$  in the C IV and Fe K $\alpha$  BEffs. In § 3.5, we probe the connections between C IV and Fe K $\alpha$  relationships in EWs, fluxes and luminosities. We present our conclusions in § 3.6. Throughout this work, we adopt the following cosmology:  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 3.2 Sample Construction

The quasar sample in our study is drawn from three sources: 50 objects from Jiang et al. (2006), which will be referred to as “Sample A”; 98 objects from GBS08, which will be referred to as “Sample B”; and 124 objects from Just et al. (2007), which is referred to as “Sample C”. We define our “combined sample” as the combination of these three data sets.

### 3.2.1 Sample A

Jiang et al. (2006) compiled a dataset of 101 Type 1 AGNs with Fe K $\alpha$  observations from both the *Chandra* and *XMM-Newton* archives. The detection fraction of the Fe K $\alpha$  line in their combined sample is around 55%. The redshifts of the AGNs range from 0.003 to 3.366, but most of the objects (87%) are low-redshift ( $z \lesssim 0.4$ ) AGNs. The monochromatic luminosities,  $l_\nu(2500 \text{ \AA})$ , range from  $10^{26.0}$  to  $10^{31.5} \text{ erg s}^{-1} \text{ Hz}^{-1}$ . We chose this dataset because it is the most complete Fe K $\alpha$  BEff sample with high-quality data obtained from the most sensitive X-ray missions. However, because the number of AGNs observed in the X-ray band is much smaller than the number observed optically and not all X-ray observed AGNs present Fe K $\alpha$  emission lines, the Fe K $\alpha$  sample is significantly limited in size.

We searched for UV/optical spectra of all 101 objects from the archival databases for *HST* and *IUE*. We found 82 spectra covering the wavelength region around the C IV emission line (containing at least the 1500–1600 Å band). If observations are available from both *HST* and *IUE*, we selected the *HST* observations due to their generally higher signal-to-noise (S/N) ratio. For spectra with multiple observations using the same instrument, we preferentially use spectra with higher S/N.

Next, we excluded all the radio-loud (RL) objects from the core sample, because additional X-ray emission is produced by the radio jet (e.g., Brinkmann et al. 2000) and changes the value of  $\alpha_{\text{ox}}$  as well as the slope of the Fe K $\alpha$  BEff.

We further excluded 3 objects for the following reasons:

- *MCG-06-30-15*: The Fe K $\alpha$  profile of MCG-06-30-15 is well fit using a broad disk line model (e.g., Tanaka et al. 1995). The narrow component is not well resolved or very weak. In this work, we only study the narrow component of Fe K $\alpha$ , so this object is excluded.
- *IC 4329a*: The spectrum has low S/N, and the C IV emission line is cannot be accurately measured (e.g., Crenshaw & Kraemer 2001).

- *PG 1407+265*: This object was termed as an “unusual” quasar (McDowell et al. 1995) because it contains extremely weak Ly $\alpha$  and C IV lines. Because of its peculiarity, we exclude it from our sample.

We removed 5 objects with strong associated absorption lines of C IV  $\lambda$ 1549: PG 1411+442 (Wise et al. 2004), NGC 4151 (Crenshaw et al. 1999), Ark 564 (Crenshaw et al. 1999), NGC 4051 (Collinge et al. 2001), and PG 1114+445 (Shang et al. 2007). These features prohibit reconstruction of the unabsorbed C IV  $\lambda$ 1549 emission-line profile. The X-ray absorption associated with the UV line absorption (Brandt et al. 2000) might also lead to an under-prediction of continuum flux at 2 keV and, therefore, an incorrect estimation of the intrinsic value of  $\alpha_{\text{ox}}$ .

Finally, we exclude 8 objects that are classified as Seyfert 1.5 (Sy 1.5) and Sy 1.9<sup>2</sup> (Osterbrock 1981; Osterbrock & Ferland 1989). These objects are intermediate between Sy 1 and Sy 2 galaxies and are often subjected to obscuration along the line of sight. Thus their UV, X-ray luminosities and  $\alpha_{\text{ox}}$  values are also potentially affected.

The final version of Sample A consists of 50 AGNs. Among these objects, 34 are found in the *HST* archive (FOS<sup>3</sup> or STIS<sup>4</sup>), and 16 are found in the *IUE* archive (SWP<sup>5</sup> or LWP<sup>6</sup>). The Fe K $\alpha$  detection rate is 55%, the same as in the entire Jiang et al. (2006) study.

### 3.2.2 Sample B

Objects in Sample B were selected from 536 SDSS DR5 (Adelman-McCarthy et al. 2007) quasars in GBS08. These quasars, taken from the DR5 Quasar Catalog (Schneider et al. 2007), are at redshift  $1.7 \leq z \leq 2.7$  and have been observed by *Chandra* or *XMM-Newton*. The lower redshift limit ensures that all SDSS spectra in this sample cover the C IV region; the upper redshift limit ensures that the rest-frame flux at 2500 Å is covered so that  $\alpha_{\text{ox}}$  can be measured accurately.

We excluded the BAL quasars in Table 1 of GBS08. BAL quasars can have strong absorption features in the C IV spectral region that prohibit accurate fitting of the continuum and emission-line profiles. In addition, BALs are usually associated with relatively strong X-ray absorption (e.g., Brandt, Laor & Wills 2000; Gallagher et al. 2006). We only retain objects with *Chandra* observations with angular offsets  $< 10'$  to avoid large X-ray flux uncertainties caused by variations of the point spread function. This restriction reduces our sample size to 149. In addition, we excluded RL objects and strong associated absorption-line (AAL) objects. The final version of Sample B consists of 98 objects with  $l_{\nu}(2500 \text{ \AA})$  between  $10^{30.53}$  and  $10^{31.67} \text{ erg s}^{-1} \text{ Hz}^{-1}$ . This is a relatively narrow luminosity range; however, it does provide a significant extension of Sample A since the latter is mostly composed of low-luminosity and low-redshift AGNs.

---

<sup>2</sup>Based on NASA/IPAC Extragalactic Database: <http://nedwww.ipac.caltech.edu/>

<sup>3</sup>Faint Object Spectrograph

<sup>4</sup>Space Telescope Imaging Spectrograph

<sup>5</sup>Short Wavelength Prime

<sup>6</sup>Long Wavelength Prime

### 3.2.3 Sample C

To examine the relations between  $\alpha_{\text{ox}}$ ,  $l_\nu(2500 \text{ \AA})$ , and  $l_\nu(2 \text{ keV})$ , Just et al. (2007) compiled a sample of 372 objects, including 26 from their core sample, 332 from Steffen et al. (2006), and 14 from Shemmer et al. (2006b). BAL quasars, RL quasars, and gravitationally lensed objects have already been excluded from this sample. The gravitationally lensed quasars are removed because their fluxes are strongly amplified and thus their luminosities are uncertain.

Among the 372 objects, 38 are already in Sample A. For the rest of the AGNs, we searched for existing spectra with C IV coverage preferentially from SDSS, then the *HST* and *IUE* archives. Finally, we removed 5 AGNs whose spectra contain strong AALs. These restrictions leave 124 objects in Sample C, in which 91 objects are from the SDSS DR5 quasar catalog, 13 from the *HST* archive, and 20 from the *IUE* archive. The redshift of this sample ranges from 0.015 to 4.720 and  $l_\nu(2500 \text{ \AA})$  ranges from  $10^{28.12}$  to  $10^{32.32} \text{ erg s}^{-1} \text{ Hz}^{-1}$ .

### 3.2.4 Combined Sample

The combined sample (Table 3.1) consists of a total of 272 objects: 189 (69.5%) have spectra from SDSS, 47 (17.3%) from *HST*, and 36 (13.2%) from *IUE*. The redshifts range from 0.009 to 4.720 (Fig. 3.1). The gap between  $z \sim 0.5$  and  $z \sim 1.5$  is caused by instrumentation limitations. Because of the wavelength coverage of the SDSS spectrographs, the redshifts of SDSS quasars having C IV coverage must be greater than 1.5. Most intermediate-redshift AGNs ( $0.5 \lesssim z \lesssim 1.5$ ) are too faint for their UV/optical spectra to be taken by *IUE* and *HST*. Our sample exhibits a strong redshift-luminosity correlation (Fig. 3.1); we discuss this issue further in our analyses below.

The  $l_\nu(2500 \text{ \AA})$  of this combined sample ranges between  $10^{26.53}$  and  $10^{33.04} \text{ erg s}^{-1} \text{ Hz}^{-1}$ , including Seyfert galaxies to the most-luminous quasars in the Universe. The X-ray detection rate is 94.9%. The UV properties of the combined sample as well as the UV and X-ray properties of Sample A are tabulated in Table 3.2 and Table 3.3. The combined sample will be used to investigate the C IV BEff, while only Sample A will be used to study the Fe K $\alpha$  BEff.

## 3.3 Data Processing

### 3.3.1 Bad Pixel Removal and Reddening Correction

To ensure that we use high-quality data to perform the continuum and emission-line fitting, we remove bad pixels in the SDSS spectra based on the mask column contained in the SDSS quasar spectral files. We removed all the bad pixels in the spectra of SDSS objects in Sample C.<sup>7</sup> The excluded pixels cover less than 10% of the total pixels for over 98% of SDSS objects, and the maximum fraction of removed pixels for a single object is 15%.

---

<sup>7</sup>For data processing of Sample B, refer to GBS08.

We perform Galactic reddening corrections to all the spectra using the  $E(B - V)$  dependent extinction curve of Fitzpatrick (1999). Values of  $E(B - V)$  are calculated following Schlegel et al. (1998).

### 3.3.2 Spectral Fitting

#### 3.3.2.1 Sample A

For each UV/optical spectrum from *HST* or *IUE*, we fit the local continuum in the vicinity of C IV (typically 1300–1700 Å) using a single power-law. We do not expect our measurements to be significantly affected by host-galaxy components because 1) our sample contains only Type I AGNs and quasars in which emission from the nucleus dominates the light from host galaxies in the UV band; and 2) C IV is in the UV region in which the non-nuclear emission primarily arises from massive stars such as O and B stars. Examination of the spectra does not reveal any stellar absorption lines, indicating that the contribution from starlight is negligible. We do not subtract the iron emission forest as this component is usually not strong around the C IV emission line (e.g., Shen et al. 2008), and the wavelength coverage of the *HST* and *IUE* spectra is frequently too narrow (only a few hundred Angstroms) to fit this component. The Balmer continuum “small blue bump” only appears in the wavelength range between 2000–4000 Å, so its contribution is negligible around the C IV region.

The emission-line spectrum is obtained after subtracting the power-law continuum. We fit the C IV emission lines using two Gaussian profiles. The model always produces visually acceptable fits. We mask narrow absorption-line features appearing near the emission-lines so as not to under-predict the emission line flux. The equivalent widths, the emission-line luminosities under the assumption of isotropy, and the continuum monochromatic luminosities at 2500 Å are calculated under our adopted cosmology and are tabulated in Table 3.2.

To first order, we use 2–10 keV luminosities tabulated in Table 1 of Jiang et al. (2006) to estimate  $l_\nu(2 \text{ keV})$  under the assumption that the X-ray continuum is a single power-law with photon index  $\Gamma = 2$  from 2 keV to 10 keV (e.g., Page et al. 2005; Shemmer et al. 2005; Vignali et al. 2005), so that

$$l_\nu(2 \text{ keV}) = \frac{L(2\text{--}10 \text{ keV})}{\nu_2 \ln 5} \quad (3.1)$$

where  $h\nu_2 = 2 \text{ keV}$  and  $h$  is Planck’s constant. To obtain the Fe K $\alpha$  emission-line flux, we further assume that the Fe K $\alpha$  emission line resembles a single Gaussian profile:  $f_l(\nu) = A \cdot e^{-(\nu - \nu_0)^2/2\sigma^2}$ . This allows one to express the line flux  $F_l$  in terms of  $l_\nu(2 \text{ keV})$  or  $F(2\text{--}10 \text{ keV})$ , the Fe K $\alpha$  equivalent width EW(Fe K $\alpha$ ), and the central energy ( $\epsilon_0 = 6.4 \text{ keV}$ ) of the emission line. Assuming the continuum flux  $f_c = C \cdot \nu^{-1}$ , we can derive

$$\text{EW}(\text{Fe K}\alpha) = \int_0^{+\infty} \frac{f_l(\nu)}{f_c(\nu)} d\nu = \frac{A}{C} \left[ \sigma^2 e^{-\nu_0^2/2\sigma^2} + \nu_0 \left( \sqrt{2\pi}\sigma - \int_{\nu_0}^{+\infty} e^{-x^2/2\sigma^2} dx \right) \right] \approx \frac{A}{C} \sqrt{2\pi} \nu_0 \sigma. \quad (3.2)$$

The Fe K $\alpha$  emission-line flux is

$$F_l = \int_0^{+\infty} f_l(\nu) d\nu = A \left( \sqrt{2\pi}\sigma - \int_{\nu_0}^{+\infty} e^{-x^2/2\sigma^2} dx \right) \approx A\sqrt{2\pi}\sigma. \quad (3.3)$$

The approximations are valid because generally  $\nu_0 \gg \sigma$  so  $\nu_0^2/2\sigma^2 \gg 1$ . For instance,  $\epsilon_0(\text{Fe K}\alpha) = 6.4$  keV while the width,  $\sigma(\text{Fe K}\alpha)$ , is usually  $\lesssim 0.1$  keV; as a result,  $e^{-\nu_0^2/2\sigma^2} \approx 0$ . We can therefore calculate the Fe K $\alpha$  line flux by

$$F_l = \frac{\text{EW}(\text{Fe K}\alpha)}{\epsilon_0} \frac{F(2-10 \text{ keV})}{\ln 5}. \quad (3.4)$$

The Fe K $\alpha$  EWs, emission-line luminosities, and 2 keV monochromatic luminosities of Sample A are tabulated in Table 3.3.

### 3.3.2.2 Sample B

For objects in Sample B, we directly adopt the fitting results from GBS08. In their paper, the SDSS spectral continua were fit with polynomials, and the C IV emission lines were fit with Voigt profiles. The different model in GBS08 from our work used to fit the C IV spectral region will not cause significant differences; because the continuum around C IV is not contaminated with other emission/absorption lines, the polynomial fit will produce almost the same result as the simple power-law fit. In addition, since both multiple Gaussian and Voigt profiles produce acceptable fits to the emission line, they will give nearly the same line flux. The X-ray spectral continua were fit using a broken power-law with the power-law break fixed at 2 keV in the rest-frame in order to obtain  $l_\nu(2 \text{ keV})$ .

### 3.3.2.3 Sample C

We fit the UV/optical spectra from the SDSS using a routine, described by Vanden Berk (2011), developed for SDSS quasar spectra. This routine fits the “underlying continuum” using three components simultaneously: a power-law, the iron emission forest, and the small blue bump. We adopt the UV iron emission template (1075–3090 Å) from Vestergaard & Wilkes (2001) and the optical template (3535–7534 Å) from Véron-Cetty et al. (2004). We first make a preliminary estimate of the power-law component by connecting two “line-free” points in the spectra. This provides initial estimates of the power-law parameters for the subsequent comprehensive processing in which the spectra are fit by considering all three components mentioned above. We evaluate the fitting quality by calculating  $\chi^2$  values within some “line-free” windows. Finally, the continuum fit is subtracted from the original spectrum and the residuals are used to conduct emission-line fitting. For Sample C, the C IV emission-lines are fit by superpositions of two Gaussian profiles, which always yields acceptable fits for the data. We then calculate the C IV EW, emission-line luminosity, and monochromatic luminosity at 2500 Å (Table 3.2). We adopt the values of  $\alpha_{\text{ox}}$  and 2 keV monochromatic luminosity from Just et al. (2007).

Examples of continuum and emission-line fits of the *HST*, *IUE* and SDSS spectra are presented in Fig. 3.2.

### 3.3.3 Error Analysis

In order to estimate the uncertainties in the measured quantities for objects in Samples A and C, we ran Monte Carlo simulations assuming a model with a perfect correlation between  $\text{EW}(\text{C IV})$ ,  $f(\text{C IV})$ , and  $f_\nu(2500 \text{ \AA})$ . For each spectrum, we add random noise to the original best fit to produce artificial spectra. The random noise follows a Gaussian distribution, and its amplitude is determined in one of two ways. For a spectrum from the *HST* or *IUE* database, we apply a low band pass filter, filtering out low-frequency signals via fast Fourier transformation.<sup>8</sup> The residual signal is mostly noise. We then calculate the root-mean-square (RMS) of the noise and take this value as the random noise amplitude to be added onto the model spectrum. For a spectrum from the SDSS database, we simply use the uncertainty level associated with each pixel as the noise amplitude. For each spectrum, we produce 100 artificial spectra and fit them in exactly the same way as the observed spectrum. The error bar of a spectral parameter is then calculated as the RMS of 100 fitting results. The typical error bar shown in each plot (e.g., Fig. 3.3) is the median of all the error bars of points in that plot.

When evaluating the uncertainties of monochromatic luminosities, e.g.,  $l_\nu(2500 \text{ \AA})$ , we must consider the contribution from the error in the luminosity distances (important for the low-redshift objects). To validate this, we calculate the ratio of luminosity uncertainties without considering distance errors ( $\delta_f$ ; the error in the flux measurement) to luminosity uncertainties after considering distance errors ( $\delta_{f,d}$ ) for objects in Sample A, and denote it as  $\delta_f/\delta_{f,d}$ , in which  $f$  stands for flux and  $d$  stands for distance. Assuming that the square of total uncertainty can be expressed as the quadratic summation of the uncertainties of flux and distance individually, the square of the ratio,  $(\delta_f/\delta_{f,d})^2$ , is more relevant than the ratio itself. We find that there are 5 out of 50 objects (in Sample A) whose ratios are above 0.1; three of them are even greater than 0.5. These objects have very low redshifts but large redshift uncertainties ( $\delta_z > 0.001$ ). For consistency, it is necessary to include the distance into the gross uncertainty calculation for all the objects in our samples. These uncertainties are also estimated using a Monte Carlo method. We adopt the most accurate values of redshift and their uncertainties from the NASA/IPAC Extragalactic Database.<sup>9</sup> These uncertainties are treated as noise amplitudes to be added to the redshift values. The luminosity uncertainties are calculated using the following error-propagation equation:

$$\delta L = 4\pi d_L \sqrt{(d_L \delta f)^2 + 4f^2 (\delta d_L)^2}$$

We adopt a 20% uncertainty for each X-ray continuum luminosity, e.g.,  $l_\nu(2 \text{ keV})$ ; their measurement errors are not available in the literature. The relative uncertainty varies considerably depending on the number of X-ray counts. For Sample A, when

---

<sup>8</sup><http://www.msi.umn.edu/software/idl/tutorial/idl-signal.html>

<sup>9</sup><http://nedwww.ipac.caltech.edu/>

the Fe K $\alpha$  emission lines are detected, the total number of counts is at least  $\sim 500$ . The Fe K $\alpha$  emission-line luminosity and flux errors are calculated using the maximum error estimates. For instance, the upper bound of  $L(\text{Fe K}\alpha)$  is calculated using the upper bounds of both EW(Fe K $\alpha$ ) and  $L(2\text{-}10 \text{ keV})$ ; the upper bound of  $f(\text{Fe K}\alpha)$  is calculated using the upper bound of  $L(\text{Fe K}\alpha)$  and the lower bound of luminosity distance  $d_L - \delta d_L$ . This uncertainty ignores any systematic uncertainty produced by errors in the cosmological model.

For Sample B, because the parameter uncertainties are not given in GBS08, we simply apply a 20% uncertainty for all luminosity values as a first-order approximation.

### 3.4 Drivers of the C IV Baldwin Effect

#### 3.4.1 Comparison with Previous Work

We will first examine some important relations to determine if our measurements are consistent with previous work. It has been argued that  $\alpha_{\text{ox}}$  has no detectable redshift dependence (e.g., Strateva et al. 2005; Steffen et al. 2006; but see Kelly et al. 2007), so in this paper we neglect any redshift evolution of  $\alpha_{\text{ox}}$ .

Fig. 3.3 displays the plot of EW(C IV) against  $l_\nu(2500 \text{ \AA})$  for the combined sample, distinguished by luminosity. We use the monochromatic luminosity at 2500  $\text{\AA}$  rather than the traditional BEff wavelength (1450  $\text{\AA}$ ) because our choice is more convenient to compare the BEff with the correlation between C IV and  $\alpha_{\text{ox}}$ . The luminosities at these two wavelengths are well correlated. The quantity  $f_{1400}/f_{2500}$  is Gaussian distributed with a dispersion of  $\sigma \sim 0.15$  (Fig. 3 in Gibson et al. 2009), so using  $l_\nu(2500 \text{ \AA})$  instead of  $l_\nu(1450 \text{ \AA})$  should only add a small dispersion to the data points but will not significantly affect the slope of the BEff. We fit the data points linearly in logarithmic space using the EM (Expectation-Maximization) method (Dempster, Laird & Rubin 1977; Table 3.4). It is clear that EW(C IV) decreases with  $l_\nu(2500 \text{ \AA})$  (Fig. 3.3).

It has been reported that the slope of the BEff becomes steeper for high-luminosity quasars. For example, Dietrich et al. (2002) obtained a C IV BEff slope ( $-0.14 \pm 0.02$ ), which was shallower than the value reported in previous studies ( $-0.22 \pm 0.05$ , Green 1996; also see Osmer, Porter & Green 1994; Laor et al. 1995). Using only the EW(C IV) measurements for their high-luminosity subsample with  $\lambda L_\lambda(1450 \text{ \AA}) \gtrsim 10^{44} \text{ erg s}^{-1}$ , Dietrich et al. (2002) obtained a steeper slope of the BEff of  $-0.20 \pm 0.03$  for C IV. To investigate the slope change with luminosity, we fit our data points with  $l_\nu(2500 \text{ \AA}) < 30.5$ . We find that the slope of the low-luminosity sample is consistent with the slope of the entire sample within  $1 \sigma$  (Table 3.4). Therefore, although our dataset exhibits a suggestive *trend* which disagrees with Dietrich et al. (2002), we do not find significant changes of slope over luminosity.

The large scatter in the BEff could have several causes, including observational error, intrinsic variation of the BEff (Osmer & Shields 1999; Shields 2007 and references therein), luminosity dependence of the BEff slope, and the redshift dependence of the BEff. Vanden Berk et al. (2009, in preparation) found that the slope of the BEff does not change across redshift, but its scaling factor (or equivalently, the EW of a broad emission line at a fixed monochromatic luminosity) exhibits a significant change Our

BEff slope is an overall average across the redshift and luminosity range we covered; remember that there is a strong redshift-luminosity correlation in our sample.

We examined the  $l_\nu(2500 \text{ \AA})$ - $l_\nu(2 \text{ keV})$  and  $l_\nu(2500 \text{ \AA})$ - $\alpha_{\text{ox}}$  relations for both Sample A and the combined sample using survival analysis (ASURV, Lavalle, Isobe & Feigelson, 1992) if censored data are involved, and find that all the results are consistent with previous work. When we calculate the slope of the Fe K $\alpha$  BEff, we apply the Buckley-James method (Table 3.4, Buckley & James 1979; Lavalle, Isobe & Feigelson 1992) included in ASURV to perform linear regression because the EW(Fe K $\alpha$ ) contains censored values. The slope of the Fe K $\alpha$  BEff of Sample A ( $-0.115 \pm 0.062$ ) is consistent with the result from the RQ sample ( $-0.102 \pm 0.052$ ) in Jiang et al. (2006).

### 3.4.2 The Effects of $\alpha_{\text{ox}}$ on the C IV and Fe K $\alpha$ BEffs

Fig. 3.4 shows the plot of EW(C IV) against  $\alpha_{\text{ox}}$  for the combined sample; the regression result from the EM algorithm for the linear relation is

$$\log \text{EW}(\text{C IV}) = (1.035 \pm 0.075)\alpha_{\text{ox}} + (3.301 \pm 0.119) \quad (3.5)$$

and the Spearman correlation coefficient is 0.607 ( $P_0 < 0.001$ ).<sup>10</sup> Because  $\alpha_{\text{ox}}$  is an indicator of the hardness of the SED which controls the ionization level of C IV surrounding the central engine, Fig. 3.4 demonstrates that as the ionizing flux becomes harder ( $\alpha_{\text{ox}}$  increases), the C IV emission has a strong positive response to  $\alpha_{\text{ox}}$ .

Both  $\alpha_{\text{ox}}$  and  $l_\nu(2500 \text{ \AA})$  are correlated with EW(C IV); which is a more fundamental driver? To investigate this issue, we applied PCA to EW(C IV),  $l_\nu(2500 \text{ \AA})$ , and  $\alpha_{\text{ox}}$  using the combined sample, Sample A, and a reduced sample. Table 3.5 presents Pearson, Spearman, and Kendall's correlation and partial-correlation coefficients (if available), along with significance levels for these three samples. We present the statistical results of Sample A for comparing the correlation results of C IV and Fe K $\alpha$ . Because the combined sample contains censored data for the  $\alpha_{\text{ox}}$  values, we must use survival analysis to calculate the correlation coefficients. However, algorithms are not available for calculating *all* the correlation and partial-correlation coefficients for censored data. For example, the empty entries in Table 3.5 are due to the unavailability of corresponding algorithms. In order to compare the changes of correlation strength when the third parameter is controlled, we construct a reduced sample with the censored data removed, considering that this only excludes a small fraction ( $\sim 6\%$ ) of the entire data set and should not affect the statistical properties of the sample. We can see that the C IV BEff is significantly weakened when  $\alpha_{\text{ox}}$  is held fixed; the correlation coefficient drops from  $-0.580$  to  $-0.224$  (Spearman). On the other hand, the correlation coefficient between EW(C IV) and  $\alpha_{\text{ox}}$  also drops significantly when  $l_\nu(2500 \text{ \AA})$  is held fixed, from  $0.615$  to  $0.332$ . This implies that both  $\alpha_{\text{ox}}$  and  $l_\nu(2500 \text{ \AA})$  are driving the change of EW(C IV).

The Fe K $\alpha$  BEff plot (not shown) of our sample (Sample A) is very similar to the correlation shown of Fig. 4 in Jiang et al. (2006) except that our sample size is smaller. Fig. 3.5 shows EW(Fe K $\alpha$ ) plotted against  $\alpha_{\text{ox}}$ . The Spearman test gives a

---

<sup>10</sup> $P_0$  is the confidence level of the null hypothesis. Therefore, the smaller  $P_0$  is the more likely the correlation exists.

much weaker correlation coefficient ( $-0.230$  with  $P_0 = 0.111$ ) than that for C IV ( $-0.304$  with  $P_0 < 0.001$ ). In addition, simple  $\chi^2$  fitting produces a slope of  $0.046 \pm 0.154$ , consistent with zero. The consistency between the correlation analysis and regression result provides strong evidence that EW(Fe K $\alpha$ ) is not correlated with  $\alpha_{\text{ox}}$ .

### 3.4.3 Effects of AGN Variability on BEff Relation Scatter

Because of the ubiquity of AGN variability, combined with the different times of the optical and X-ray observations, our values of  $\alpha_{\text{ox}}$  do not reflect the spectral hardness at a specific time but are randomly distributed around their mean values. The deviation of  $\alpha_{\text{ox}}$  from its mean value would be  $\sim 0.083$ , assuming that the variation amplitudes are 30% for  $l_\nu(2500 \text{ \AA})$  and 40% for  $l_\nu(2 \text{ keV})$  (e.g., Strateva et al. 2005, GBS08).

To check how much of the scatter of our correlations could be attributed to variability, we performed two simple tests on our combined sample. We follow the method used in § 3.1 of GBS08 and introduce  $\Delta \log \text{EW}$ , which is the difference between the observed EW(C IV) and the EW calculated from linear regression (Eq. 3.7), i.e.,  $\Delta \log \text{EW} = \log \text{EW} - \log \text{EW}(l_\nu(2500 \text{ \AA}))$ . We define  $\mu$  and  $\sigma$  as the mean and dispersion of the distribution of  $\Delta \log \text{EW}$ . To calculate  $\mu$  and  $\sigma$ , we maximize the likelihood function (Maccacaro et al. 1988):

$$L = \prod_i \frac{1}{\sqrt{2\pi(\sigma_i^2 + \sigma^2)}} \exp \left[ -(\Delta \log \text{EW}_i - \mu)^2 / 2(\sigma_i^2 + \sigma^2) \right] \quad (3.6)$$

in which the subscript  $i$  represents each object and  $\sigma_i$  is the uncertainty of  $\Delta \log \text{EW}$  associated with each  $\Delta \log \text{EW}_i$ . The maximization of  $L$  requires  $\mu = 0.01$  and  $\sigma = 0.23$ .

Next, we estimate the potential scatter of  $\Delta \log \text{EW}$  due to variability. We need to consider two terms:  $\log \text{EW}$  and  $\log \text{EW}(l_\nu(2500 \text{ \AA}))$ . To first approximation,  $\text{EW} \sim f_{\text{line}}/f_{\text{cont}}$  in which  $f_{\text{line}}$  is the emission-line flux and the  $f_{\text{cont}}$  is the continuum flux. The emission-line variability of six luminous quasars at  $z = 2.2\text{--}3.2$  was recently reported by Kaspi et al. (2007). The mean fractional variation  $F_{\text{var}}^{11}$  is  $\approx 0.096$  by averaging the fractional variation of C IV  $\lambda 1549$  of all six quasars. The C IV emission-line variability of a number of Seyfert galaxies has been studied, including Fairall 9 (Rodriguez-Pascual et al. 1997), NGC 5548 (Clavel et al. 1991), NGC 7469 (Wanders et al. 1997), NGC 3783 (Reichert et al. 1994), and 3C 390.3 (O'Brien et al. 1998). By averaging the fractional variations of Seyferts and quasars above, we obtain an *average* emission-line variation  $\langle F_{\text{var}} \rangle = 0.130$ . The typical variation of  $l_\nu(2500 \text{ \AA})$  is  $\sim 30\%$  (e.g., Strateva et al. 2005). Therefore, the scattering of  $\log \text{EW}$  contributed from variability is *estimated* (assuming all independent variables are Gaussian distributed) as

$$\sigma(\Delta \log \text{EW}) = \frac{1}{\ln 10} \sqrt{\left( \frac{\delta f_{\text{line}}}{f_{\text{line}}} \right)^2 + \left( \frac{\delta f_{\text{cont}}}{f_{\text{cont}}} \right)^2 + a^2 \left[ \frac{\delta l_\nu(2500 \text{ \AA})}{l_\nu(2500 \text{ \AA})} \right]^2} \approx 0.144.$$

---

<sup>11</sup> $F_{\text{var}}$  is defined as the RMS of the intrinsic variability relative to the mean flux (Rodriguez-Pascual et al. 1997).

In the calculation above,  $a = 0.198$  (Eq. (3.7)). This exercise indicates that at least 60% of the scatter around the BEff in our sample can be attributed to AGN variability.

We performed a similar test for the  $\alpha_{\text{ox}}$ -EW(C IV) relation. Because the set of  $\alpha_{\text{ox}}$  contains censored data, we can only use the reduced data set (258 objects). The maximization of  $L$  (Eq. 3.6) yields  $\mu = 0.01$  and  $\sigma = 0.22$ . The potential dispersion of this relation assuming all scatter comes from variability is *estimated* (assuming Gaussian distributions) as

$$\sigma(\Delta \log \text{EW}) = \frac{1}{\ln 10} \sqrt{\left( \frac{\delta f_{\text{line}}}{f_{\text{line}}} \right)^2 + \left( \frac{\delta f_{\text{cont}}}{f_{\text{cont}}} \right)^2 + (0.3838a)^2 \left[ \left( \frac{\delta l_{\nu}(2500 \text{ \AA})}{l_{\nu}(2500 \text{ \AA})} \right)^2 + \left( \frac{\delta l_{\nu}(2 \text{ keV})}{l_{\nu}(2 \text{ keV})} \right)^2 \right]} \approx 0.16$$

In the calculation above,  $a = 1.035$  (Eq. 3.5). This indicates that variability produces at least 75% of the scatter around the  $\alpha_{\text{ox}}$ -EW(C IV) relationship.

In summary, the above two tests demonstrate that a substantial fraction, if not the majority, of the scatter in the correlations above can be attributed to X-ray and UV/optical variability. It should be possible to make the correlations tighter if the UV/optical and X-ray data are observed simultaneously.

### 3.4.4 Regressions of EW and Luminosity

The BEff provides a potential avenue to infer the luminosity of a quasar from emission-line observations. Type Ia supernovae (SNe) are treated as classical standard candles (e.g., Phillips 1993; Burrows 2000), but only a few are observed beyond  $z \sim 1.5$ . If quasars, which are much easier to detect than SNe and can be observed to much a higher redshift, can be used as standard candles, they would be an important tool for cosmological studies. Soon after the discovery of the BEff, many investigations considered the possibility of treating emission-line EW as a luminosity indicator (e.g., Baldwin 1977; Wampler 1980). Unfortunately, the C IV BEff usually has a large scatter (Osmer & Shields 1999; Shields 2007); given the small slope in the log EW-log  $l_{\nu}$  plot (on the order of  $-0.2$ ), the predicted luminosities are very inaccurate. It has also been shown that the C IV BEff is redshift dependent (Francis & Koratkar 1995; Vanden Berk et al. 2009, in preparation), making it a less valuable probe of cosmology.

Because we are focusing on the influence of  $\alpha_{\text{ox}}$  at the moment, and will put the redshift factor aside, we will concentrate on the issue of whether the scatter can be reduced if we regress EW(C IV) with  $l_{\nu}(2500 \text{ \AA})$  and/or  $\alpha_{\text{ox}}$ , and if the prediction of luminosity can be made more accurate with this approach. The linear-regression results of EW(C IV) with  $\alpha_{\text{ox}}$  are already shown in Eq.(3.5). Similar regressions from  $l_{\nu}(2500 \text{ \AA})$  and both of  $l_{\nu}(2500 \text{ \AA})$  and  $\alpha_{\text{ox}}$ , using the fully parametric EM algorithm, are

$$\log \text{EW(C IV)} = (-0.198 \pm 0.015) \log l_{\nu}(2500 \text{ \AA}) + (7.764 \pm 0.461) \quad (3.7)$$

$$\log \text{EW(C IV)} = (-0.107 \pm 0.021) \log l_{\nu}(2500 \text{ \AA}) + (0.615 \pm 0.106) \alpha_{\text{ox}} + (5.944 \pm 0.536). \quad (3.8)$$

in which  $\text{EW}(\text{C IV})$  is in Å and  $l_\nu(2500 \text{ Å})$  is in  $\text{erg s}^{-1} \text{ Hz}^{-1}$ . The last regression was performed on the combined sample without the censored data (258 objects) because the EM algorithm in ASURV does not allow both independent variables to contain censored data points.

To evaluate the scatter, we subtract the predicted EW values calculated using the above equations from the observed values and compute the RMS values of the residuals. We see a slight improvement of the RMS values, using  $\alpha_{\text{ox}}$  and both  $l_\nu(2500 \text{ Å}) + \alpha_{\text{ox}}$  (Table 3.6). The last regression result (Eq.(3.8)) is consistent with Eq.(11) in GBS08, indicating that at least part of the scatter of the BEff is due to  $\alpha_{\text{ox}}$ .

Next, we regress luminosity against  $\text{EW}(\text{C IV})$  and/or  $\alpha_{\text{ox}}$ , using the combined sample. The results are

$$\log l_\nu(2500 \text{ Å}) = (-1.980 \pm 0.154) \log \text{EW}(\text{C IV}) + (34.076 \pm 0.256) \quad (3.9)$$

$$\log l_\nu(2500 \text{ Å}) = (-0.852 \pm 0.168) \log \text{EW}(\text{C IV}) - (2.864 \pm 0.262) \alpha_{\text{ox}} + (27.668 \pm 0.635). \quad (3.10)$$

The last regression was performed on the censored-data-excluded sample. We then calculate the RMS values of the residuals after subtracting the predictions from the equations above (Table 3.6). The RMS value shrinks by 18% using  $\text{EW}(\text{C IV}) + \alpha_{\text{ox}}$ , compared to using  $\text{EW}(\text{C IV})$  alone. We use the standard *F*-test to check if the two sets of residuals have consistent variance. The testing gives an *F*-statistic of 1.48 with a significance 0.002, indicating that these two sets of residuals have significantly different variances and 18% is a statistically significant improvement. To use quasars as standard candles via the BEff, we should at least confine the luminosity within an uncertainty of 30%, or, equivalently,  $\text{rms} < 0.1$ . This cannot be achieved using our current dataset and controlled parameters.

### 3.5 Relation Between Fe K $\alpha$ and C IV

Fe K $\alpha$  is important in AGN studies because it is the strongest emission line appearing in the X-ray band. However, the strength of this emission line varies significantly from object to object, and the line is not detected in most X-ray observations of quasars. Given this lack of direct observational measurement, it would be useful to develop a way to predict the expected strength of the line empirically.

The  $\text{EW}(\text{C IV})$  and  $\text{EW}(\text{Fe K}\alpha)$  do not exhibit a significant correlation (Fig. 3.6) in Sample A; the correlation has a low Spearman rank-correlation coefficient (0.319 with  $P_0 = 0.027$ ). Although the EWs of the two lines are not well correlated, their luminosities and fluxes are strongly correlated (Fig. 3.7), with Spearman correlation coefficients 0.529 ( $P_0 < 0.001$ ) and 0.551 ( $P_0 < 0.001$ ), respectively (Table 3.7). The linear correlations regressed in Fig. 3.7 are

$$\log L(\text{Fe K}\alpha) = (0.588 \pm 0.079) \log L(\text{C IV}) + (16.164 \pm 3.416) \quad (3.11)$$

$$\log f(\text{Fe K}\alpha) = (0.978 \pm 0.188) \log f(\text{C IV}) - (2.082 \pm 2.228). \quad (3.12)$$

One must always question the significance of relations such as (3.11) because even if there is no correlation between the observed fluxes of the lines, the fact that the line luminosities for a given object contain the same distance factor will introduce an apparent correlation in the luminosities. To investigate whether this effect is important for our study, we perform a test in which we conduct correlation and regression analysis for a sub-sample of Sample A. Objects in this sub-sample have similar redshifts, and thus they have approximately the same distances. First, we use objects with  $0.06 < z < 0.09$  because this redshift bin contains a large number of objects. This sub-sample contains 10 objects. The correlation coefficient of  $f(\text{Fe K}\alpha)$ - $f(\text{C IV})$  is 0.624 ( $P_0 = 0.061$ ) and of  $L(\text{Fe K}\alpha)$ - $L(\text{C IV})$  is 0.709 ( $P_0 = 0.033$ ). The regression results are

$$\log L(\text{Fe K}\alpha) = (0.748 \pm 0.200) \log L(\text{C IV}) + (9.330 \pm 8.702) \quad (3.13)$$

$$\log f(\text{Fe K}\alpha) = (0.771 \pm 0.241) \log f(\text{C IV}) - (4.280 \pm 2.819). \quad (3.14)$$

Both the correlation and regression results of this sub-sample are consistent with the results of the entire sample within uncertainties, suggesting that the luminosity correlation of the C IV and Fe K $\alpha$  lines may be real and is not a consequence of multiplying the two fluxes of a given object by the same large distance factor. The luminosity of C IV emission line increases faster than the luminosity of Fe K $\alpha$ .

One might be concerned that the correlation between  $L(\text{Fe K}\alpha)$  and  $L(\text{C IV})$  is *artificial* because the calculation of the Fe K $\alpha$  measurements involves  $L(2\text{--}10 \text{ keV})$  (Eq. (3.4)), which is proportional to  $l_\nu(2 \text{ keV})$  (Eq. (3.1)), and  $l_\nu(2 \text{ keV})$  is correlated with  $l_\nu(2500 \text{ \AA})$ , which is correlated with EW(C IV), i.e., the C IV BEff. EW(C IV) is calculated from continuum and emission line luminosity, so apparently, the  $L(\text{Fe K}\alpha)$  and  $L(\text{C IV})$  are not independent before we perform the correlation. However, our calculation of Fe K $\alpha$  is simply reversing the  $F(2\text{--}10 \text{ keV})/\text{EW}$  calculation of Jiang et al. (2006), so  $F(\text{Fe K}\alpha)$  is not actually dependent of  $F(2\text{--}10 \text{ keV})$  and hence  $l_\nu(2500 \text{ \AA})$ . In essence, we have the values of  $F(\text{Fe K}\alpha)$  independent of  $F(\text{C IV})$ . Therefore, our  $L(\text{Fe K}\alpha)$  and  $L(\text{C IV})$  correlation, which is *expected* from existing relations, is not an artifact correlation.

That the EWs of C IV and Fe K $\alpha$  are uncorrelated is consistent with the result by Page et al. (2004a) and further demonstrates that the C IV and Fe K $\alpha$  emission lines are unlikely to have the same origin. This result is not surprising because Fe K $\alpha$  and C IV are produced in different processes. The correlation between their luminosities is probably a combination of effects between their EWs (uncorrelated) and continuum (strongly correlated). The flux correlation, although empirical and not very tight, is a useful first order estimation of the Fe K $\alpha$  line flux given the UV spectra in the rest-frame of an AGN.

### 3.6 Discussions and Conclusions

We have compiled a sample of 272 Type 1 AGNs and quasars that have UV and X-ray measurements, among which Fe K $\alpha$  emission lines are detected in 50 objects. The sample covers a wide range of redshift ( $0.009 \lesssim z \lesssim 4.720$ ), and a wide range of luminosity from Seyfert galaxies to the most-luminous quasars ( $27.81 \lesssim \log l_\nu(2500 \text{ \AA}) \lesssim 33.04$ ).

These properties allow us to study the overall properties of AGNs rather than focusing on a particular redshift or luminosity. It also has a high X-ray detection rate ( $\sim 96\%$ ), which lets us obtain robust statistics. We have performed correlation and regression analyses using this sample and draw the following conclusions:

1. The C IV BEff is driven by both  $\alpha_{\text{ox}}$  and  $l_{\nu}(2500 \text{ \AA})$ , or equivalently, by  $l_{\nu}(2 \text{ keV})$  and  $l_{\nu}(2500 \text{ \AA})$ . This implies that changes in the ionizing flux induce changes in the ionization state of the BELR, producing more C IV ions when the SED becomes harder and vice versa. This is supported both by correlation and regression analyses:
  - The partial correlation between EW(C IV) and  $l_{\nu}(2500 \text{ \AA})$  when  $\alpha_{\text{ox}}$  is controlled is weaker than the regular correlation between EW(C IV) and  $l_{\nu}(2500 \text{ \AA})$ .
  - The scatter in the linear regression decreases when we regress EW with  $\alpha_{\text{ox}} + l_{\nu}(2500 \text{ \AA})$  compared with  $l_{\nu}(2500 \text{ \AA})$  alone.

Although the reduction of the scatter due to adding another regression parameter is not sufficiently large to treat quasars as standard candles, it demonstrates that a significant fraction of the scatter attributes to  $\alpha_{\text{ox}}$ , and can be reduced by including it in regression analysis.

2. EW(Fe K $\alpha$ ) exhibits no strong correlation with either  $\alpha_{\text{ox}}$  or EW(C IV). This implies that Fe K $\alpha$  is not likely to have the same origin as C IV.
3. There may be a correlation between the luminosities of Fe K $\alpha$  and C IV with a logarithmic slope of  $0.588 \pm 0.079$ . This correlation is possibly because both of these two quantities involve a factor related to the scale of the line emitting regions and the slope indicates that the C IV emission line luminosity increases faster than the Fe K $\alpha$ .

Although  $\alpha_{\text{ox}}$  is a fundamental influence on EW(C IV), there is still a significant scatter in the EW(C IV)- $\alpha_{\text{ox}}$  diagram. As we have demonstrated, most of the scatter is contributed by variability, but another likely contribution source is the nature of  $\alpha_{\text{ox}}$  which only connects the flux points at 2500  $\text{\AA}$  and 2 keV but misses the Big Blue Bump, which is expected to play an important role in the photoionization process. The shape of an AGN SED can be very different depending on Eddington ratio ( $L_{\text{bol}}/L_{\text{Edd}}$ ) but still have a fairly constant  $\alpha_{\text{ox}}$  (Vasudevan & Fabian 2007). It is perhaps more appropriate to use a point near  $\sim 250 \text{ \AA}$  instead of 2500  $\text{\AA}$  to calculate a revised  $\alpha_{\text{ox}}$  (Shemmer et al. 2008). This new  $\alpha_{\text{ox}}$  might be more strongly correlated with EW(C IV). However, this requires challenging observations that cannot be achieved at present for most AGNs.

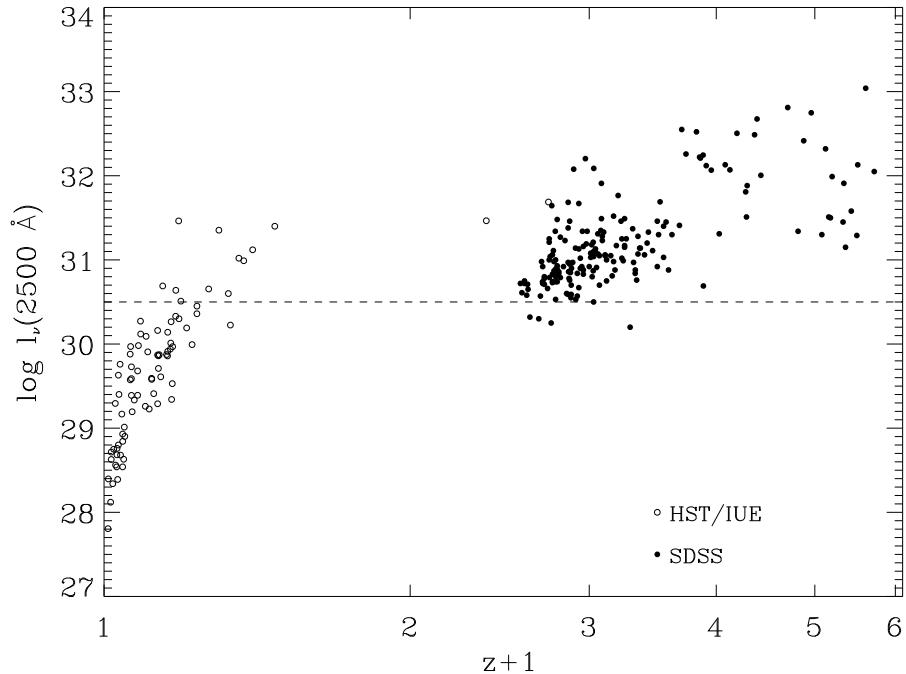


Fig. 3.1 The luminosity-redshift diagram of the combined sample containing 272 objects distinguished by observational facilities. The gap for  $0.5 \lesssim z \lesssim 1.5$  arises because of the detection limit of *HST/IUE* and wavelength coverage of the SDSS. The dashed line marks the position where  $l_\nu(2500 \text{ \AA}) = 30.5$ .

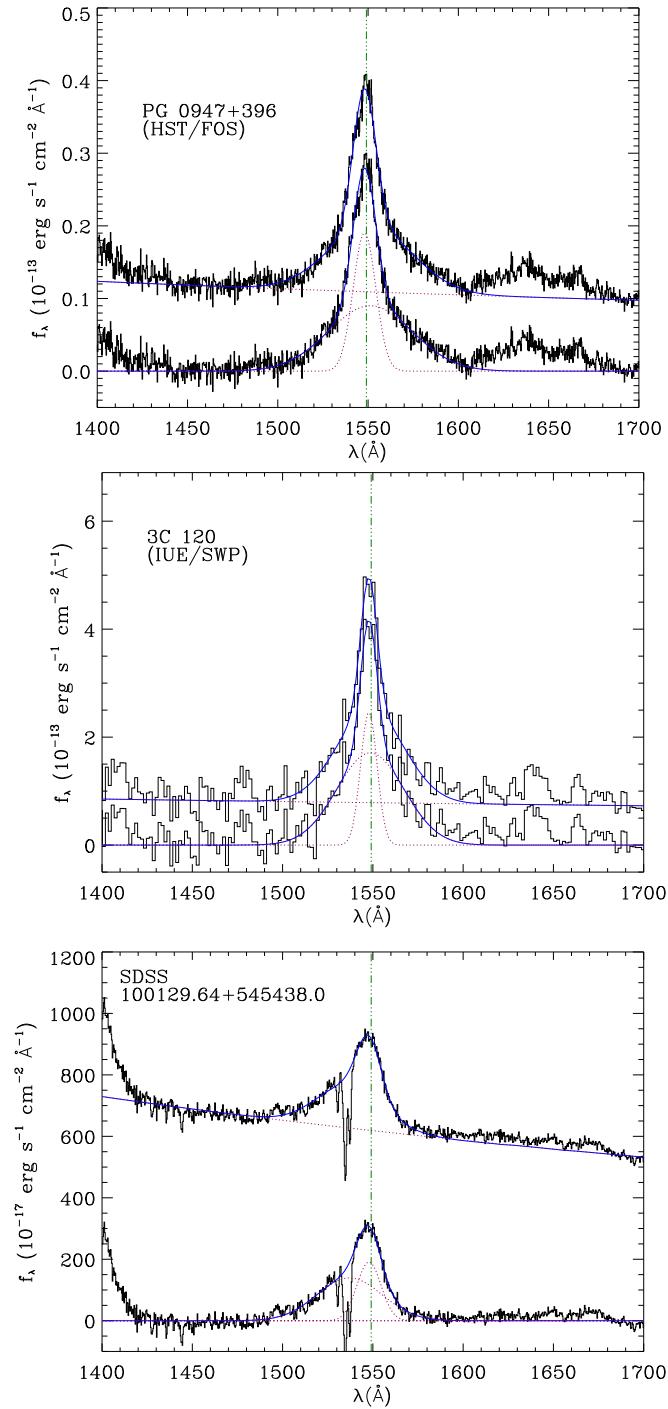


Fig. 3.2 Continuum and C IV emission-line fit examples. In each panel, the upper spectrum is the original and the lower spectrum is continuum subtracted; blue solid curves are the fits to the spectra. The cyan dotted curves are emission-line components. For PG 0947+396 and 3C 120, these components include the power-law continua and two Gaussian profiles for C IV. For SDSS J100129.64+545438.0, we also plot the iron emission forest and small Balmer bump components, although they are so weak that they are almost invisible. The rest-frame spectral resolutions of these spectra are, from top to bottom,  $\sim 1.6 \text{ Å}$ ,  $\sim 5 \text{ Å}$ ,  $\sim 0.8 \text{ Å}$ .

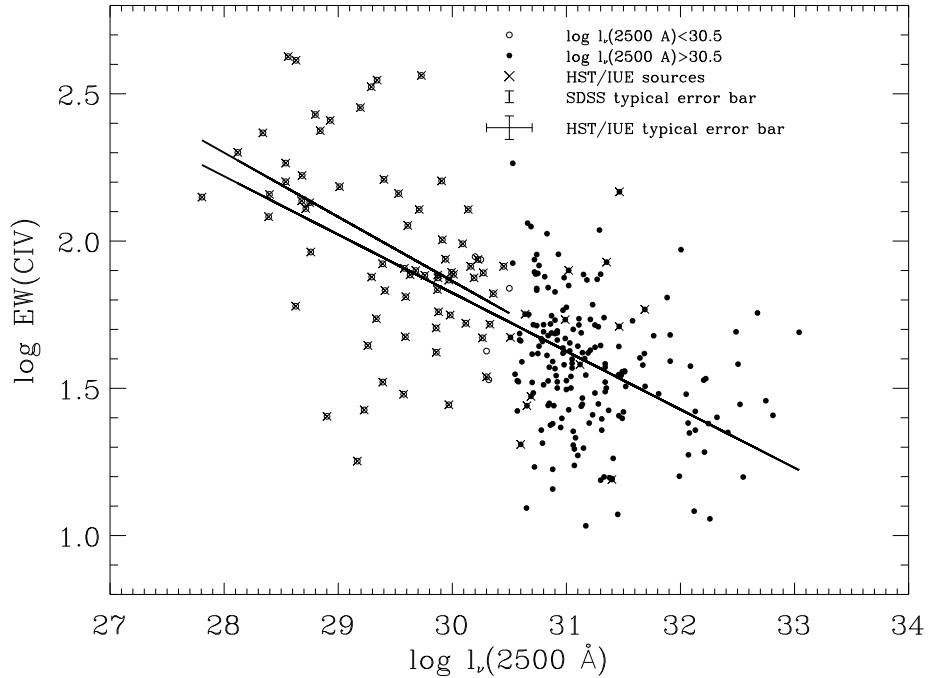


Fig. 3.3 The C IV BEff diagram for the combined sample. The EW(C IV) is given in  $\text{\AA}$ , and  $l_\nu(2500 \text{ \AA})$  is in units of  $\text{erg s}^{-1} \text{ Hz}^{-1}$ . Points are distinguished by luminosity; the short solid line is the best linear fit to low-luminosity points and the long one is the best linear fit to the entire sample.

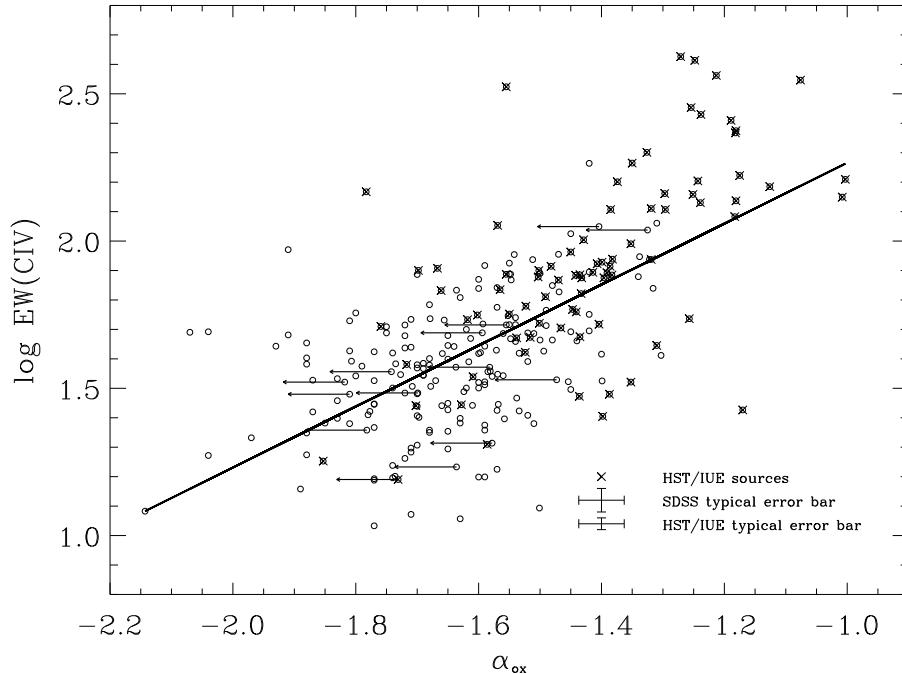


Fig. 3.4 The correlation between  $\text{EW}(\text{C IV})$  and  $\alpha_{\text{ox}}$ . Upper limits on  $\alpha_{\text{ox}}$  are marked with arrows. The solid line is the best linear fit using the EM algorithm.

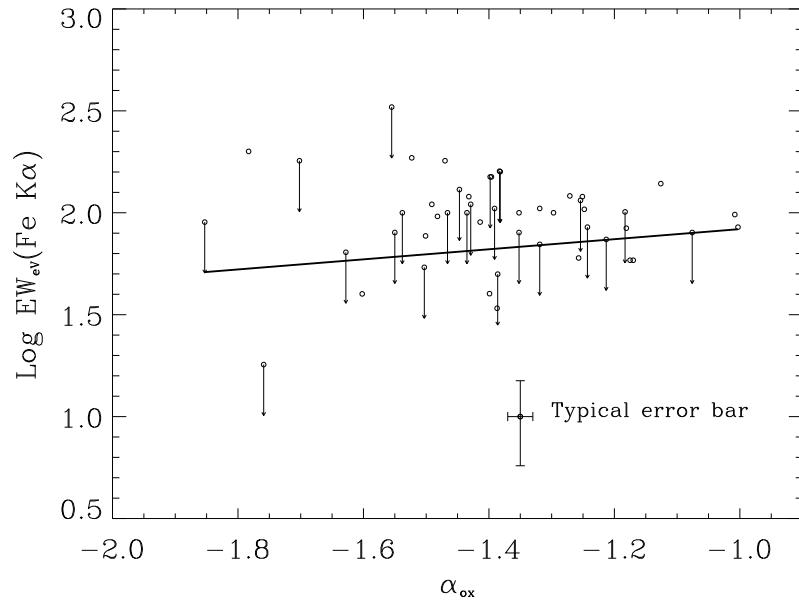


Fig. 3.5 The correlation between EW(Fe K $\alpha$ ) and  $\alpha_{\text{ox}}$  for the core sample. The solid line is a linear fit using the ASURV package for the censored data. The typical error bar of the data is displayed at the top-right corner of the plot.

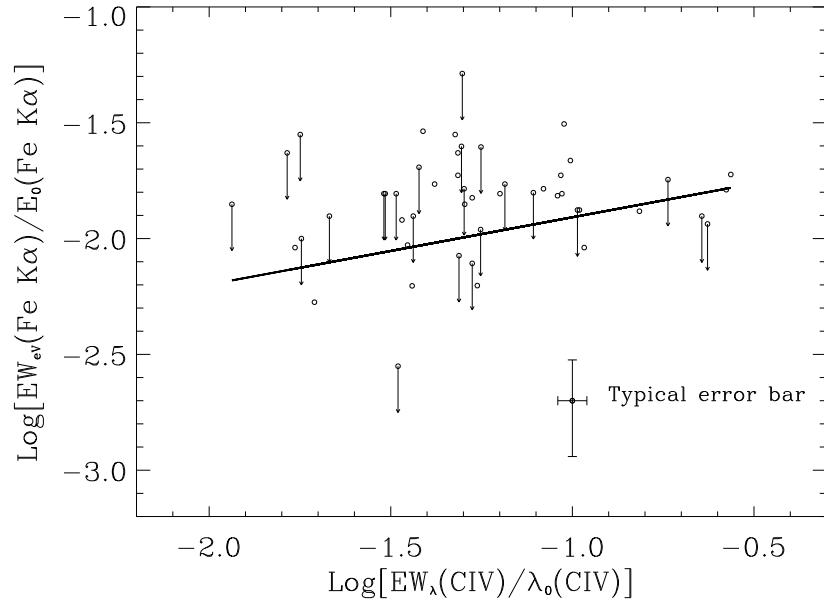


Fig. 3.6 Plot of  $\log [EW(\text{Fe K}\alpha)/E_0(\text{Fe K}\alpha)]$  vs.  $\log [EW(\text{C IV})/\lambda_0(\text{C IV})]$  for the core sample. Because the units of  $EW(\text{Fe K}\alpha)$  and  $EW(\text{C IV})$  are different, we divide them by central energy  $E_0(\text{Fe K}\alpha) = 6.4 \text{ keV}$  and central wavelength  $\lambda_0(\text{C IV})$  to make them dimensionless.

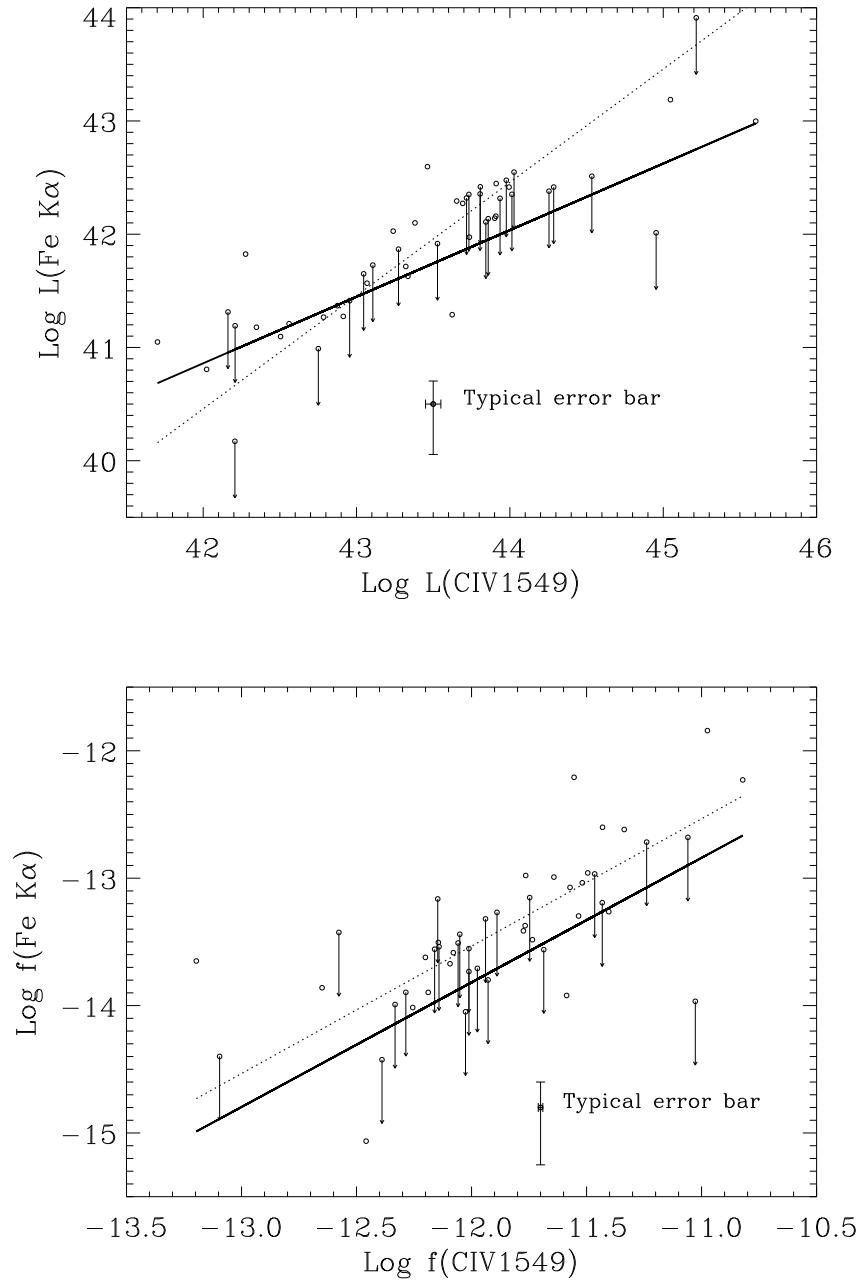


Fig. 3.7 Plot of  $\log L(\text{Fe K}\alpha)$  vs.  $\log L(\text{C IV})$  (upper panel) and  $\log f(\text{Fe K}\alpha)$  vs.  $\log f(\text{C IV})$  (lower panel) for the core sample. Upper limits are denoted as downward arrows. The solid lines are the best linear fits to the data using ASURV. For comparison, we show the best linear fits with a unity slope in dotted lines. The data symbols and typical error bars are labelled at the bottom right corners.

Table 3.1. Summary of samples.

Sample	SDSS	HST	IUE	Total	Redshift range	$\log l_\nu(2500 \text{ \AA})$ range
A	0	34	16	50	0.009–1.735	27.81–31.69
B	98	0	0	98	1.7–2.7	30.53–31.67
C	91	13	20	124	0.015–4.720	28.12–33.04
Combined	189	47	36	272	0.009–4.720	27.81–33.04

Table 3.2. UV properties of the combined sample.

Object	$z$	$\log l_\nu(2500 \text{ \AA})$ (erg s $^{-1}$ Hz $^{-1}$ )	EW(C IV) (\AA)	$\alpha_{\text{ox}}$	Flag <sup>1</sup>	UV/Optical Instrument	Sample
NGC4593	0.009	$27.81 \pm 0.15$	$141.0 \pm 23.2$	$-1.008 \pm 0.066$	0	HST	A
NGC3783	0.010	$28.40 \pm 0.08$	$144.0 \pm 9.4$	$-1.251 \pm 0.045$	0	HST	A
Mkn352	0.015	$28.12 \pm 0.15$	$200.0 \pm 12.9$	$-1.326 \pm 0.067$	0	IUE	C
Mrk1044	0.016	$28.63 \pm 0.04$	$60.1 \pm 1.3$	$-1.523 \pm 0.037$	0	HST	A
NGC7469	0.016	$28.72 \pm 0.28$	$129.0 \pm 25.1$	$-1.319 \pm 0.111$	0	HST	A
MCG8–11–11	0.020	$28.34 \pm 0.19$	$233.0 \pm 39.1$	$-1.182 \pm 0.082$	0	IUE	C
Mkn79	0.022	$28.75 \pm 0.09$	$135.0 \pm 11.3$	$-1.239 \pm 0.049$	0	IUE	C
Mrk335	0.026	$29.29 \pm 0.18$	$75.5 \pm 7.7$	$-1.503 \pm 0.077$	0	HST	A

Note. — Table 3.2 is published in its entirety in the electronic version of the *Astrophysical Journal*. The portion is shown here for guidance regarding its form and content.

<sup>1</sup> A value of “1” indicates that  $\alpha_{\text{ox}}$  for this object is an upper limit.

Table 3.3. UV and X-ray properties of Sample A objects.

Object	$z$	EW(Fe K $\alpha$ ) (eV)	$\log l_\nu(2 \text{ keV})$ ( $\text{erg s}^{-1} \text{ Hz}^{-1}$ )	$\log L(\text{Fe K}\alpha)$ ( $\text{erg s}^{-1}$ )	$\log f(\text{Fe K}\alpha)$ ( $\text{erg s}^{-1} \text{ cm}^{-2}$ )	$\log L(\text{CIV})$ ( $\text{erg s}^{-1}$ )	$\log f(\text{CIV})$ ( $\text{erg s}^{-1} \text{ cm}^{-2}$ )
NGC4593	0.009	98.0 $^{+21.0}_{-21.0}$	$25.18 \pm 0.09$	$41.05^{+0.09}_{-0.10}$	$-12.21^{+0.08}_{-0.11}$	$41.70 \pm 0.03$	$-11.55 \pm 0.03$
NGC3783	0.010	120.0 $^{+14.0}_{-14.0}$	$25.14 \pm 0.09$	$41.10^{+0.05}_{-0.05}$	$-12.98^{+0.12}_{-0.17}$	$42.50 \pm 0.01$	$-10.82 \pm 0.01$
NGC7469	0.016	105.0 $^{+25.0}_{-25.0}$	$25.28 \pm 0.09$	$41.18^{+0.09}_{-0.12}$	$\lesssim -13.19$	$42.35 \pm 0.03$	$-11.43 \pm 0.03$
Mrk1044	0.016	186.0 $^{+61.0}_{-61.0}$	$24.66 \pm 0.09$	$40.81^{+0.12}_{-0.17}$	$-11.84^{+0.22}_{-0.48}$	$42.02 \pm 0.01$	$-11.76 \pm 0.01$
Mrk335	0.026	$\lesssim 54.0$	$25.38 \pm 0.09$	$\lesssim 40.99$	$\lesssim -14.42$	$42.75 \pm 0.02$	$-11.43 \pm 0.02$
Mrk590	0.026	121.0 $^{+65.0}_{-51.4}$	$25.25 \pm 0.09$	$41.21^{+0.19}_{-0.24}$	$-13.41^{+0.17}_{-0.28}$	$42.56 \pm 0.01$	$-11.64 \pm 0.01$
Mrk290	0.030	58.5 $^{+56.5}_{-43.4}$	$25.62 \pm 0.09$	$41.27^{+0.29}_{-0.59}$	$-12.62^{+0.08}_{-0.09}$	$42.79 \pm 0.01$	$-11.52 \pm 0.01$
Mrk493	0.031	$\lesssim 101.0$	$25.31 \pm 0.09$	$\lesssim 41.19$	$\lesssim -13.55$	$42.21 \pm 0.03$	$-12.15 \pm 0.03$

Note. — Table 3.3 is published in its entirety in the electronic version of the *Astrophysical Journal*. The portion is shown here for guidance regarding its form and content.

Table 3.4. Hypothesis and linear fitting results.

Fig	Sample	$x$	$y$	$\rho^1 P_0^2$	$k^3$
3.3	Combined	$\log l_\lambda(2500 \text{ \AA})$	$\log \text{EW}_\lambda(\text{C IV})$	-0.559(< 0.001)	$-0.198 \pm 0.015$
3.3	Combined with $\log l_\nu(2500 \text{ \AA}) < 30.5$	$\log l_\lambda(2500 \text{ \AA})$	$\log \text{EW}_\lambda(\text{C IV})$	-0.465(< 0.001)	$-0.218 \pm 0.048$
3.4	Combined	$\alpha_{\text{ox}}$	$\log \text{EW}_\lambda(\text{C IV})$	0.755(< 0.001)	$1.035 \pm 0.075$
3.5	Sample A	$\alpha_{\text{ox}}$	$\log \text{EW}_{\text{eV}}(\text{Fe K}\alpha)$	0.104(0.470)	$0.247 \pm 0.230^4$

<sup>1</sup>Spearman rank correlation coefficient.<sup>2</sup>Significance levels of Spearman's rank correlation.<sup>3</sup>Slope from EM algorithm.<sup>4</sup>Calculated for Sample A using the Buckley-James method in the ASURV software package (Lavalley et al. 1992).

Table 3.5. Correlation and partial-correlation analysis results.

$x$	$y$	$z^a$	Spearman	Pearson	Kendall
Combined sample (279)					
log EW(C IV)	log $l_\nu$ (2500 Å)	$\alpha_{\text{ox}}$	...	...	-0.237 <sup>b</sup>
log EW(C IV)	log $l_\nu$ (2500 Å)		-0.599(<0.001)	...	-0.400(<0.001)
$\alpha_{\text{ox}}$	log EW(C IV)	log $l_\nu$ (2500 Å)	...	...	0.260
$\alpha_{\text{ox}}$	log EW(C IV)		0.607(<0.001)	...	...
Combined Sample Without Censored Data (258)					
log EW(C IV)	log $l_\nu$ (2500 Å)	$\alpha_{\text{ox}}$	-0.224	-0.311	-0.220
log EW(C IV)	log $l_\nu$ (2500 Å)		-0.580(< 0.001)	-0.605	-0.417(< 0.001)
$\alpha_{\text{ox}}$	log EW(C IV)	log $l_\nu$ (2500 Å)	0.332	0.258	0.284
$\alpha_{\text{ox}}$	log EW(C IV)		0.615(< 0.001)	0.587	0.450
Sample A (49)					
log EW(C IV)	log $l_\nu$ (2500 Å)	$\alpha_{\text{ox}}$	0.120( 0.291)	0.074( 0.214)	0.036
log EW(C IV)	log $l_\nu$ (2500 Å)		-0.304(0.001)	-0.304( 0.001)	-0.180( 0.001)
log EW(Fe K $\alpha$ )	log $l_\nu$ (2 keV)	$\alpha_{\text{ox}}$	... <sup>c</sup>	... <sup>d</sup>	< 0.001
log EW(Fe K $\alpha$ )	log $l_\nu$ (2 keV)		-0.230( 0.111)	...	...
$\alpha_{\text{ox}}$	log EW(Fe K $\alpha$ )	log $l_\nu$ (2 keV)	...	...	-0.130
$\alpha_{\text{ox}}$	log EW(Fe K $\alpha$ )		0.104( 0.470)	...	...

Note. — The number in the parentheses after a correlation coefficient is its significance level. Because we are considering the null hypothesis, a small number indicates a possibility of a strong correlation.

<sup>a</sup>  $z$  is the controlled parameter. If a  $z$  entry is not empty, we calculate the partial correlation of  $x$  and  $y$  while controlling for  $z$ ; otherwise, we only calculate the correlation of  $x$  and  $y$ .

<sup>b</sup> For the Kendall's partial  $\tau$  correlation coefficients, generally, the sampling distribution is unknown; therefore, the probability values are not available (Kendall 1938, 1970).

<sup>c,d</sup> Because EW(Fe K $\alpha$ ) is a set of censored data, we have to use a generalized correlation statistics that can deal with censored data to derive the coefficients. For the non-partial correlations, we used the ASURV software package (Lavalle et al. 1992) which only provides Spearman's  $\rho$  and Kendall's  $\tau$  correlation correlation coefficients. For the PCA, we only use the Kendall's partial  $\tau$  correlation statistics (Akritas & Siebert 1996).

Table 3.6. The RMS values of residuals after regression from different variables.

Independent variables	Dependent variable	RMS values
$l_\nu(2500 \text{ \AA})$	EW(C IV)	0.231
$\alpha_{\text{ox}}$	EW(C IV)	0.228
$l_\nu(2500 \text{ \AA}) + \alpha_{\text{ox}}$	EW(C IV)	0.217
EW(C IV)	$l_\nu(2500 \text{ \AA})$	0.747
$\alpha_{\text{ox}}$	$l_\nu(2500 \text{ \AA})$	0.645
EW(C IV) + $\alpha_{\text{ox}}$	$l_\nu(2500 \text{ \AA})$	0.615

Note. — To consistently compare the RMS values, we compute them using the combined sample without the censored data.

Table 3.7. Correlation and regression analysis for EW, emission line luminosity, and flux data between Fe K $\alpha$  and C IV.

Relations	$\rho(P)$ <sup>1</sup>	$k^2$
EW	0.319( 0.027)	0.291±0.131
$L(\text{line})$	0.529(<0.001)	0.588±0.079
$f(\text{line})$	0.551(<0.001)	0.978±0.188

<sup>1</sup>Correlations are tested using Spearman's  $\rho$  and the significance level ( $P$ ) is evaluated against the null hypothesis.

<sup>2</sup>We use the Buckley-James method to do linear regression and  $k$  is the slope. The computation is done using the ASURV software package (Lavalley et al. 1992)

## Chapter 4

# A Quasar Catalog with Simultaneous UV, Optical and X-ray Observations by *Swift*

### 4.1 Introduction

Variability is a ubiquitous phenomenon for quasars (e.g., Matthews & Sandage 1963; Smith & Hoffleit 1963; Gaskell & Klimek 2003; Wilhite et al. 2005, 2006; Meusinger et al. 2011) and has been observed from radio, infrared, UV/optical, X-ray and even  $\gamma$ -ray bands (e.g., Kovalev et al. 2002; Sakata et al. 2010; Park et al. 2010; Grupe et al. 2010; Rieger & Aharonian 2008). While variability provides considerable information on the length scales of the quasar central engine and can be utilized to estimate the mass of the central super massive black hole (e.g., Peterson et al. 2004), it is a significant source of scatter to many correlations. Specifically, flux observed at a given frequency may change over time, so when you combine fluxes observed in different wavebands, the combined spectral shape may not represent the spectral shape at a specific time. It is therefore essential to take simultaneous observations in multi-wavebands in order to understand the true spectral shape.

The dispersion in multi-waveband correlations produced by variability can be estimated. For example, it is estimated that at least 60% of the dispersion of the Baldwin Effect and at least 75% of the  $\text{EW}(\text{C IV}) - \alpha_{\text{ox}}$  relation can be attributed to variability (Wu et al. 2009). However, a program of simultaneous observations at multi-wavebands to verify that the dispersions can be reduced has not yet been performed. In this work, we will use simultaneous UV/optical and X-ray observations of serendipitous quasars to study quasar SEDs and dispersions caused by variability.

Because the *Swift* is designed for detecting and observing Gamma-ray bursts, there are only a few quasars at the pointing centers. However, the relatively large FOVs of both the UVOT and XRT allow coverages of a number of serendipitous objects. The SDSS DR5 quasar catalog contains 77,429 quasars in  $5,740 \text{ deg}^2$ ; the UVOT FOV is  $17' \times 17'$ , so there is approximately 1 quasar per UVOT field. Based on this estimation, we match the *Swift* pointings with SDSS DR5 quasar catalog with a matching radius of  $20'$ . In this work, we use the *Swift* pointing results from launch through June, 2008 ( $\sim 3.5$  years).

Catalogs using similar strategies were constructed by Tueller et al. (2008) and Grupe et al. (2010). Our work is unique in terms of the selection criteria and sample size. The sample in Tueller et al. (2008) contains 153 9-month hard X-ray (14–195 keV) selected local AGNs with a mean redshift of  $\sim 0.03$ . Because this sample is selected using the hard X-ray band which is much more penetrating to obscuration than UV/optical and soft X-ray photons, it provides a homogeneous sample to study UV/optical and/or X-ray faint AGNs. Using this sample, Winter et al. (2009) estimated the fraction of

“Hidden” AGNs (Ueda et al. 2007) in the local universe to be  $\sim 24\%$ . The sample in Grupe et al. (2010) contains 92 soft X-ray selected AGNs with redshift ranging from 0.002 to 0.349. Due to their selection criteria, objects in this sample are X-ray bright Type 1 AGNs with few ( $< 10\%$ ) radio loud (RL; e.g., 3C 273) and Seyfert 1.5 objects (e.g., Mkn 841; see Wilkes et al. 1999). Using this sample, they constructed composite AGN SEDs using simultaneous observations from *Swift*. By fitting these SEDs using two different models, they attempted to constrain AGN bolometric corrections (BCs). They found a strong correlation between UV and X-ray spectral indices, namely  $\alpha_{\text{UV}}$  and  $\alpha_X$ <sup>1</sup> for AGNs with  $\alpha_X > -1.6$ . Because their sample contains a large fraction of narrow line Seyfert 1 (NLS1) galaxies, they were also able to study differences of broad line Seyfert 1 (BLS1) galaxies and NLS1 galaxies in terms of  $\alpha_{\text{UV}} - L_{\text{bol}}/L_{\text{Edd}}$  relation.

As a result of the detection limits in their selected wavebands, both of the above samples are limited to nearby AGNs and relatively small sample sizes. Our sample is optically selected. The raw catalog after matching the SDSS quasar catalog and the *Swift* pointings contains 1034 quasars, which is almost an order of magnitude larger than either previous samples. The redshift of our sample ranges from 0 to  $\sim 5$ . Figure 4.1 shows the distribution of our quasar sample in the luminosity (represented by absolute magnitude)–redshift diagram. The distribution of our sample in this diagram is consistent with SDSS DR5 quasar catalog. This sample allows us to expand the quasar SED study to high redshift.

One of the prominent features in the AGN SED is the big blue bump (BBB) in the extreme UV (EUV) energy band, which is believed to be produced by the thermal emission from the accretion disk (e.g., Shields 1978). Because of strong Galactic and (possibly) intrinsic absorption, it is almost impossible to observe this wavelength region. On the other hand, this feature is likely to be an important contributor to quasar bolometric luminosity. Based on the observed data in UV/optical and X-ray bands, we can place some constraints to the flux contribution this feature. Grupe et al. (2010) compiled a catalog of  $\sim 100$  quasars with simultaneous observations from *Swift* and constrained the BBB. However, quasars in their sample are at low redshift ( $z < 0.4$ ). In contrast, over 50% of quasars in our sample are at  $z \gtrsim 1$ . As an extension of their work, we will attempt to constrain the BBB feature based on observations of quasars are much higher redshifts.

It should be emphasized that we are not repeating the work by Elvis et al. (1994) and Richards et al. (2006). Instead, we are attempting to use simultaneously observed data to constrain the BBB. This component, because of lack of observational data, is usually represented as a power-law connecting a UV flux point, such as 2500 Å, and an X-ray point, such as 2 keV (e.g., Richards et al. 2006). The slope determined by these two points is defined as  $\alpha_{\text{ox}}$  (Tananbaum et al. 1979), so that

$$\alpha_{\text{ox}} = \frac{\log l_\nu(2 \text{ keV})/l_\nu(2500 \text{ \AA})}{\log \nu(2 \text{ keV})/\nu(2500 \text{ \AA})} = 0.3838 \log [l_\nu(2 \text{ keV})/l_\nu(2500 \text{ \AA})]. \quad (4.1)$$

---

<sup>1</sup>In Grupe et al. (2010), both  $\alpha_{\text{UV}}$  and  $\alpha_X$  are defined based on  $f_\nu \propto \nu^{-\alpha_\nu}$ . In our work, we define them as  $f_\nu \propto \nu^{+\alpha_\nu}$ .

This parameter is generally used to characterize the spectral hardness between UV and X-ray band (e.g., Avni & Tananbaum 1982, 1986; Anderson & Margon 1987; Wilkes et al. 1994; Vignali et al. 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007). A long-standing problem about the measurement of  $\alpha_{\text{ox}}$  is that it is difficult to obtain simultaneous measurements of an object in the X-ray and UV/optical. Source variability introduces scatter into measurements of  $\alpha_{\text{ox}}$ . The simultaneous nature of our observation will remove the uncertainty.

This paper is organized as follows. In Section 4.2, we describe the observations and data processing; in Section 4.3, we present our data processing results, including the UV/optical light curves, and the composite SED for each quasar; in Section 4.4, we present our final catalog and describe the catalog format; in Section 4.5, we select a sample of quasars from this catalog to study the dispersion of the  $\alpha_{\text{ox}} - l_b(2500 \text{ \AA})$  relation. Finally, in Section 4.6, we summarize this catalog and our conclusions. Throughout this work, we adopt the following cosmology:  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ .

## 4.2 Observations and Data Processing

The catalog is constructed in the following steps.

1. Construct the raw catalog by matching SDSS quasar catalog to *Swift* pointings.
2. Process XRT data to obtain X-ray spectra and parameters.
3. Process UVOT data to obtain UV/optical photometry.
4. Supplement UV/optical data from other resources.
5. Create quasar SEDs.
6. Calculate parameters for each quasar and generate the final catalog.

The raw catalog is constructed by matching 3.5 years *Swift* pointings and SDSS DR5 quasar catalog and contains 1034 objects (see Section 4.1). We briefly introduce the *Swift* instruments before describing the processing steps in details.

### 4.2.1 Instruments

The *Swift* Gamma-ray Burst Explorer (Gehrels et al. 2004), with three major instruments on board, provides an ideal facility for observing simultaneous observations in UV/optical and X-ray bands. *Swift* instruments include the Burst Alert Telescope (BAT; Barthelmy et al. 2005), the X-ray Telescope (XRT; Burrows et al. 2005) and the Ultraviolet/Optical Telescope (UVOT; Roming et al. 2005).

The BAT is a highly sensitive, large field of view (FOV) instrument designed to provide critical GRB triggers and  $4'$  position accuracy. It has a  $1.4 \text{ sr}$  FOV (half coded), with an energy range of  $15\text{--}150 \text{ keV}$  for imaging with a non-coded response up to  $500 \text{ keV}$ . The XRT is a focusing X-ray telescope with  $23' \text{ FOV}$ ,  $18'' \text{ resolution}$  (half-power diameter), and  $0.2\text{--}10 \text{ keV}$  energy range. It is designed for measuring fluxes, spectra and light curves of GRBs and afterglows over a dynamic range covering more

than 7 orders of magnitude in flux. XRT supports three readout modes: 1) Imaging mode is only used to position bright sources up to  $7 \times 10^{-7}$  erg cm $^{-2}$  s $^{-1}$ . 2) Windowed timing mode can achieve high time resolution (2.2 ms) and bright source spectroscopy through rapid CCD readout, but has a relatively low positioning resolution; it is most useful for sources with flux below  $\sim 10^{-7}$  erg cm $^{-2}$  s $^{-1}$ . 3) Photon-counting mode permits full spectral and spatial information for low flux sources ranging from  $2 \times 10^{-14}$  to  $2 \times 10^{-11}$  erg cm $^{-2}$  s $^{-1}$ . In this work, we only use XRT observations in photon-counting mode because most objects in our catalog are serendipitous quasars with low X-ray flux. The UVOT (Roming et al. 2005) is based on the Optical Monitor on board the *XMM-Newton* space X-ray observatory (Mason et al. 1996). It is co-aligned with XRT with a 30 cm diameter primary mirror and a FOV of  $17' \times 17'$ . The six filters carried by UVOT and their basic characteristics can be found in Table 1 of Poole et al. (2008).

#### 4.2.2 XRT Data

We process the XRT data using the standard pipeline routine *xrtpipeline* Version 0.12.4 (Released on 2009-08-25). This routine stacks all the available XRT sky images for each observational segment (identified by an observational ID, hereafter OBSID) and generates a composite event file with an associated exposure map. In case of multiple event files for one OBSID, we use the one observed in photon counting mode and with the longest exposure. In most cases, the discarded event files are exposed for only tens of seconds and are essentially useless for our study.

In the second step, we define source and background region files for every XRT sky image of an individual object. The quasar coordinates are adopted from SDSS DR5 quasar catalog. The source and background region circles are defined in the sky image with the longest exposure. These regions are then visually inspected to ensure that they can be applied to each sky image with shorter exposures. While we keep the same source region circle, the position and size of the background region circle is adjusted so that it is applicable to all sky images covering this object. We exclude sky images in which the source region is not or only partially inside the FOV or not covered by all the observing frames. We also discard sky images in which the target source is severely contaminated by nearby X-ray sources. Depending on the brightness of the X-ray source, the source region radius is adjusted to be large enough to ensure that the photon density at the border is at the same level as the background region file. We correspondingly adjust the background region circle so that the radius is at least four times as large as the source region circle.

In the next step, event files in each OBSID are processed using *XSelect* (V2.4a). We use this software to filter out suspect photons and only keep photons between 0.3 and 10 keV in the observed frame. We then create an auxiliary response file (ARF) for each OBSID using *xrtmkarf* (Version 0.5.6). The ARF file deals with filter transmission, vignetting, effective area and point spread function (PSF) using the standard response matrix. It also automatically removes “hot columns”, which decrease the effective exposure time. The response matrix file (RMF) we use is *sxwpc0to12s0\_20010101v011.rmf*. The calibration database (CalDB) version is *CALDB-4.3.0*. We flag each quasar based

on the number of X-ray photon counts ( $N_{\text{Xph}}$ ):  $g$  (good with  $N_{\text{Xph}} > 100$ ),  $a$  (acceptable with  $10 < N_{\text{Xph}} \leq 100$ ), and  $w$  (weak with  $N_{\text{Xph}} \leq 10$ ).

On August 30th, 2007 (at 14:28 UT), the CCD substrate voltage was raised from  $V_{ss} = 0$  V to  $V_{ss} = 6$  V. The change was made in order to reduce the thermally-induced dark current in the CCD, which allows the XRT to collect useful science data at slightly higher temperatures (3 to 4°C) than was possible before the change. The first set of PC RMFs are un-broadened RMFs, and they are the same as those issued from launch to February 28th, 2007 (*swxpc0to12s0\_20010101v011.rmf* and *swxpc0s0\_20010101v011.rmf*). The second and the third sets of PC RMFs, including an epoch-dependent broadened kernel, should be used for data collected from March 1st, 2007 to August 31st, 2007 (*swxpc0to12s0\_20070301v011.rmf* and *swxpc0s0\_20070301v011.rmf*) and from September 1st, 2007 onwards (*swxpc0to12s6\_20070901v011.rmf* and *swxpc0s6\_20070901v011.rmf*). Because we are only using data collected in photon counting mode, we do not need to deal with the response matrix function of data collected in window timing mode. The details regarding to response function file changes can be seen from <http://www.swift.ac.uk/Gain\RMF\releases.html>. According to the changes above, strictly, we should group OBSIDs observed in the same period and process each period separately using their corresponding RMF files. However, because we only fit the X-ray continuum, the line profile broadening will not significantly affect our fitting results, so we do not distinguish RMFs in different periods.

Next, we read all the available event files and extract photons from the source and background regions, respectively. We also generate the ARF file for the composite event file using *addarf* (Version 1.2.6). The X-ray energy spectrum is binned using *grppha* (Version 3.0.1). The binning strategy and the statistical method we use are described in Table 4.1. Basically,  $\chi^2$  statistics are used if the X-ray photon count ( $N_{\text{Xph}}$ ) is greater than 100 so that on average, each bin contains at least 10 photons and there are at least 10 bins. If  $N_{\text{Xph}}$  drops below 100, we do not bin the spectrum so as not to lose spectral information, but use Cash statistics (Cash 1979). If  $N_{\text{Xph}} < 10$ , we do not generate energy spectra but only estimate flux upper limits. At the end of this step, 103 objects are flagged as good ( $g$ ), 296 are flagged as acceptable ( $a$ ) and 406 are flagged as weak ( $w$ ).

The X-ray energy spectrum is fit using *XSPEC* (Arnaud 1996; Version 12.5.1n). For each object with  $N_{\text{Xph}} \gtrsim 10$ , we fit the X-ray spectrum with four models (Table 4.2). The H I Galactic column density,  $N_{\text{H,G}}$ , is obtained using the HEASARC *nh* command which calculates  $N_{\text{H,G}}$  based on the Leiden/Argentine/Bonn (Lab) survey of Galactic H I (Hartmann & Burton 1997; Kalberla et al. 2005). We visually inspect the fitting quality and select the model that produces the best constraint and quality. During this visual inspection process, we must balance parameter constraints and degree of freedom (DOF). If parameter constraints from a more flexible model are too loose, we favor a simpler model with a low DOF because it provides tighter parameter constraints. For example, in most of cases, the intrinsic column density  $N_{\text{H,i}}$  from Model C is consistent with zero so we choose Model B, because this model produces much tighter constraints to the flux level and photon index. In general, we consistently apply the following rules to select the model.

1. If  $N_{\text{Xph}} < 30$ , use Model A. In this case, it is impossible to constrain the intrinsic absorption and the constraint on  $\Gamma$  is very poor ( $\delta\Gamma \gtrsim 0.5$ ). Therefore, we decide to fit the flux levels and fix  $\Gamma = 2$ .
2. If  $30 \leq N_{\text{Xph}} < 100$ , there are two possible situations.
  - If the intrinsic column density  $N_{\text{H},i}$  is inconsistent with zero and well constrained, choose Model D. Although  $N_{\text{H},i}$  is also set to be free in Model C, we do not choose this model because constraints  $N_{\text{H},i}$  and other parameters are very poor due to a small  $N_{\text{Xph}}$ .
  - If  $N_{\text{H},i}$  is consistent with zero, and  $\Gamma$  is well constrained ( $\delta\Gamma < 0.5$ ), choose Model B. If  $\Gamma$  is not well constrained, choose Model A.
3. In case of  $N_{\text{Xph}} \geq 100$ , if both  $\Gamma$  and  $N_{\text{H},i}$  are well constrained, use Model C. Otherwise, simply follow Rule 2.

The visual inspection process is performed by three people and a consensus is reached to avoid discriminations and inconsistencies. The number of each model we finally choose is shown in Table 4.2. For each model fit (A,B,C,D), we present an example as shown in Figure 4.5.

Based on the spectral fitting results, we calculate a number of parameters including the photon count rate, the observed flux between 0.3 and 10 keV, the unabsorbed flux between 0.3 and 10 keV, and the hardness ratio. For objects with spectra, we shift the data into the rest frame and calculate the monochromatic luminosity at 2 keV. The numerical values of these physical quantities are tabulated in the final catalog (see Section 4.4).

#### 4.2.3 Weak X-ray Sources

We process weak X-ray sources (flagged as  $w$  in final catalog) with  $N_{\text{Xph}} < 10$  separately because a fraction of these objects may not be detected by XRT and we can only estimate their flux upper limits.

We first use *confidlev* (Kraft et al. 1991) to determine whether an X-ray source is detected or not. There are three input parameters of this program: the total number of photon counts in the source region  $N_{\text{Xph,tot,srcreg}}$ , the background photons counts in the source region  $N_{\text{Xph,bkg,srcreg}}$  and a confidence level. We scale the background photons from the background region  $N_{\text{Xph,bkg,bkgreg}}$

$$N_{\text{Xph,bkg,srcreg}} = N_{\text{Xph,bkg,bkgreg}} \frac{1}{R_T} \frac{1}{R_A}$$

in which  $R_T = T_{\text{bkgreg}}/T_{\text{srcreg}}$  is the ratio of average exposure times in source and background regions and  $R_A = A_{\text{bkgreg}}/A_{\text{srcreg}}$  is the ratio of areas in source and background regions. We use the  $3\sigma$  as the criteria to judge if an X-ray source is detected so the confidence level is 0.9973. Based on the results of *confidlev*, 98 out of 406 weak X-ray sources are detected by XRT.

The X-ray flux values of detected X-ray sources are calculated using Model A as bright X-ray source and PSF correction is automatically performed by *Xspec* using the ARF file. For undetected sources, we use flux scaling and interpolation methods to obtain the PSF-corrected flux upper limits. For undetected sources with none-zero photons received, we can still use *Xspec* to calculate the PSF corrected flux between 0.3 and 10 keV  $F_{\text{PSF},\text{Xspec}}$ <sup>2</sup>, the flux upper limit can be scaled from photon counts so that

$$F_{\text{PSF},\text{Xspec,lim}} = F_{\text{PSF},\text{Xspec}} \frac{N_{\text{Xph,lim}}}{N_{\text{Xph}}}$$

in which  $N_{\text{Xph,lim}}$  is the source photon count upper limit output by *confidlev*.

If an X-ray source has zero photon counts received or if the received photon count is small (e.g.,  $N_{\text{Xph}} < 2$ ), *Xspec* is unable to build a model to fit input data. In these cases, we first calculate the flux without PSF correction using *pimms* (Mukai 1993) and then obtain the PSF-corrected flux using interpolation. The *pimms* software allows us to calculate the absorbed and unabsorbed flux between 0.3 and 10 keV *without PSF correction*, i.e.,  $F_{\text{abs,noPSF}}$  and  $F_{\text{unabs,noPSF}}$  by specifying the count rate upper limits. Note that for one object, the fraction before and after the PSF correction is constant, regardless of absorbed or unabsorbed flux, i.e.,

$$\frac{F_{\text{abs,PSF}}}{F_{\text{abs,noPSF}}} = \frac{F_{\text{unabs,PSF}}}{F_{\text{unabs,noPSF}}} = \text{const}$$

Therefore, the only parameter we need to know if the PSF-corrected flux with absorption  $F_{\text{abs,PSF}}$ . This can be obtained by scaling the count rate limits based a detected reference X-ray source, i.e.,

$$F_{\text{PSF,int,lim}} = \text{CR}_{\text{lim}} \frac{F_{\text{PSF,Xspec}}^0}{\text{CR}^0} \quad (4.2)$$

in which  $\text{CR}^0$  and  $F_{\text{PSF,Xspec}}^0$  are the count rate and *real* flux with absorption of the detected reference X-ray source, and  $\text{CR}_{\text{lim}}$  is the count rate upper limit of the undetected X-ray source. Because the PSF correction is dependent on the offset angle  $\theta$  and photon energy  $\epsilon$  (Moretti et al. 2005), the reference object must be located at a similar offset angle and have a similar spectral shape as the target source. In addition, the reference object should be located at the same offset angle in all sky images so that the difference of PSF correction across different images is negligible. To simplify the process to search for these reference objects to match a particular target object, we use a more general approach by generating a curve of  $F_{\text{PSF,Xspec}}^0/\text{CR}^0$  as a function of offset angle  $\theta$  based on reference objects and interpolate this curve to obtain the corresponding  $F_{\text{PSF,Xspec}}/\text{CR}$  for undetected using Equation (4.2).

The reference object candidates are selected from two sets

1. X-ray sources flagged with *g* and *a* whose X-ray spectra are modeled using Model A and Model B. For X-ray sources modeled with Model B, we only select sources

---

<sup>2</sup>The uncertainty is so large that it is consistent with zero.

whose photon indices are consistent with  $\Gamma = 2$  within uncertainties. There are 263 candidates in this set.

## 2. Detected X-ray sources flagged with $w$ . This set includes 98 sources.

For consistency, the selected reference objects must have the same source region aperture (otherwise, the PSF corrections are different). Because a majority of sources has a source region radius of  $30''$  (Figure 4.2), X-ray sources whose source region apertures other than this value are excluded. For each object, the offset angles of all sky images must be consistent within  $\delta\theta < 0.1'$ . Using these two criteria, the final set of reference objects contains 153 objects. Their offset angle distribution is presented in Figure 4.3.

In Figure 4.4, we plot the  $F/\text{CR}$  vs.  $\theta$  for all of reference objects. The dispersion is due to other factors that control the PSF correction, e.g., photon energy distribution. Because we wish to perform a *mean* interpolation for these extremely X-ray faint sources by neglecting other factors, we bin the data points in Figure 4.4 with bin size of  $1'$ . The red boxes in Figure 4.4 represent the curve after binning, and the curve is tabulated in Table 4.3.

For undetected X-ray sources whose  $N_{Xph} < 2$  and source region apertures are not  $30''$ , their flux limits are estimated by re-extracting events and X-ray spectra with source region radius equal to  $30''$ . If the original source region radius is smaller than  $30''$ , this may reduce the  $S/N$  of the source, but will still provide an order of magnitude estimate of flux upper limit. After calculating the PSF-corrected unabsorbed flux between 0.3 and 10 keV, we can easily calculate the rest-frame flux limit at 2 keV and convert it to monochromatic luminosity limit at 2 keV.

### 4.2.4 UVOT Data

Instead of using a pipeline to obtain a co-added event file as was done with the XRT data, we process each individual UVOT sky image for each object. The composite photometry can then be obtained by summing over photon counts from each individual image and dividing it by the total exposure time.

First, we must identify defective sky images. Because we are using serendipitous observations, it is inevitable that some sources are too close to the edge of the FOV<sup>3</sup>. The goal of this process is to select a set of sky images so that the default source and background regions can be directly applied on them.

The source region is defined using a circle with a radius of  $3.^{\prime}0$ , which is recommended in the UVOT photometric calibration (Poole et al. 2008). Since a  $5.^{\prime}0$  radius aperture, which contains  $85.8 \pm 3.8\%$  of the PSF, was used for calibrating the UVOT, an aperture correction is applied to the data when running *uvotsource* (Poole et al. 2008). The inner and outer radii of the background region are  $r = 27.^{\prime}5$  and  $R = 35.^{\prime}0$  respectively. This is the standard background region used to construct the first GRB afterglow catalog (Roming et al. 2009). The large difference between the source and background

---

<sup>3</sup>Objects are not necessarily always within the FOV because our matching radius is larger than UVOT FOV and/or because UVOT was working under other modes.

radii ensures that the background is at least 50 times as large as the source region to provide an accurate background subtraction.

We then flag each sky image based on the position of the object and the aspect keyword value. We only use the images whose aspect values equal to DIRECT. The other images (flagged as -1) may not have correct aspect corrections and may lead to inaccurate photometric results. The detailed image flagging strategy is tabulated in Table 4.4 and depicted in Figure 4.6. First, if the object is out of the FOV, images are flagged as -1 or 5 and discarded. In addition, the object cannot be too close to the edge because of vignetting effects (flagged as 100+) or photon losses (flagged as 4) due to the PSF. The minimum distance to the edge is  $d_E = r = 27.^{\circ}5$  and the minimum distance to the corner is  $d_V = 100''$ . If  $r < d_E < R = 35.^{\circ}0$  (flagged as 10+), we may need to shift the background region slightly, so these images are passed to visual inspection process. We accept images with  $d_E > R$ .

Due to relatively small photon counts or positioning uncertainties, a fraction of object images are not symmetric or well aligned with the default source region circle with a radius of  $3.^{\circ}0$ . It is necessary to co-align the quasar image with the center of the source region circle because the aperture correction process assumes a symmetric photon loss outside the pre-defined aperture. To co-align the locations, we extract a region of  $11 \times 11$  pixels (about  $5'' \times 5''$ ) centered on the SDSS quasar coordinates. If the total photon count inside this region is less than 20, we regard this image as “faint” and assign a flag of “1”. Even if a single image is regarded as “faint”, it will contribute when stacked with other images of the source (e.g., Lehmer et al. 2007). If an image is not flagged as “faint”, we attempt to fit the source with a two-dimensional Gaussian profile. The background is determined by averaging the photons inside the region centered at the SDSS coordinate and bounded by two squares with sides of  $21 \times 21$  pixels and  $11 \times 11$  pixels. If the photon count inside the  $11 \times 11$  pixel region is less than three times of the background level, this image is also treated as “faint” (flagged as “1”). Otherwise, we fit the  $11 \times 11$  pixel region with a 2-D Gaussian profile initially centered at the SDSS quasar coordinate. The new Gaussian centroid is then compared with the original SDSS quasar coordinate. We denote their separation as  $\delta$  (see Figure 4.6).

There are three categories of images.

- If  $\delta \leq 0.^{\circ}618$ , the two coordinates are regarded as consistent and the default source region circle is used. This image is flagged as “0” and is passed to the next step.
- If  $0.^{\circ}618 < \delta \leq 3.^{\circ}0$ , this image is flagged as “2” and the source region circle is centered at the Gaussian centroid.
- If  $\delta > 3.^{\circ}$ , this image is flagged as “3”. This large offset could be caused by the unrecorded aspect problem, which requires manual correction or non-Gaussian photon distribution. Less than 1% sky images are flagged as “3”, so we drop these images from our analysis.

In the next step, we carefully define the default source and background regions for each object. For each object, we select the sky image with the longest passband central wavelength and the longest exposure. We preferentially use V band, because (at least for local AGNs), stellar emission (and therefore the host galaxy contribution)

is larger in optical than in UV. If the V band is not available or its exposure is too short ( $\lesssim 100$  seconds), we examine the B band. If B is unavailable, we check U band, and then UVW1, UVM2 and UVW2. For each object, we attempt to exclude all stars, galaxies and bad pixels in the background annulus. If the source region is significantly contaminated or most of the background region must be masked, we remove this object from the UVOT analysis.

The procedures above can be performed automatically, but occasionally one must manually adjust the results. For example, we found a number of images with stellar ghost rings. Because we only use sky images of one waveband with the longest exposure, these images may not be viewed and excluded in the region define process. This implies that we must inspect all sky images to completely exclude images with nearby stars, stellar ghost rings, galaxy spiral arms, comets, and bad pixels. Due to the difference of aspect and exposure time, the influence of these factors may vary for each sky image even for a given object. In addition, although the `ASPCORR` keyword is set to `DIRECT`, the image may still suffer from uncorrected aspect problems so that a single point source may have multiple images or even smeared images. Consequently, it is necessary to visually inspect each sky image ( $> 26,000$  unique images) to perform the second order region customization and image selection. After the visual inspection process, 3183 (1%) UVOT sky images were excluded from our UVOT image set (Table 4.5). Figure 4.7 shows some examples of removed sky images.

Our goal is to construct a catalog that contains the deepest sky image in each filter for each object, but we are also interested in obtaining the time-resolved photometry. Therefore, we use *uvotsource*<sup>4</sup> (HEADAS Version 6.10) to calculate photometry on *individual* sky images using the curves of growth from the *Swift* CALDB<sup>5</sup> for the aperture correction model. The composite photometry is calculated using the photon counts of all available images. For a given waveband, the mean magnitude is calculated as

$$\langle m \rangle = Z_{\text{pt}} - 2.5 \log \langle R_{\text{LSS}} \rangle \quad (4.3)$$

In Equation 4.3, the  $Z_{\text{pt}}$  is the zero-point magnitude (the magnitude when the count rate is 1 photon s<sup>-1</sup>). The values are taken from Table 6 of Poole et al. (2008).  $R_{\text{LSS}}$  is the source count rate with coincidence-loss corrections, aperture corrections and large-scale sensitivity corrections applied. The associated uncertainty in this magnitude measurement is

$$\delta_{\langle m \rangle} = \sqrt{\Delta_{Z_{\text{pt}}}^2 + \left( \frac{2.5}{\ln 10} \langle R_{\text{LSS}} \rangle \right)^2 \Delta_{R_{\text{LSS}}}^2} \quad (4.4)$$

in which

$$\langle R_{\text{LSS}} \rangle = \frac{\sum R_{\text{LSS},i} \cdot T_i}{\sum T_i}, \quad \text{and} \quad \Delta_{\langle R_{\text{LSS}} \rangle} = \frac{\sqrt{\sum (T_i \Delta_{R_{\text{LSS},i}})^2}}{\sum T_i} \quad (4.5)$$

---

<sup>4</sup><http://heasarc.gsfc.nasa.gov/ftools/caldb/help/uvotsource.html>

<sup>5</sup><http://heasarc.gsfc.nasa.gov/docs/heasarc/caldb/data/swift/uvota/index.html>

$T_i$  is the exposure time in image  $i$ . We also calculate the flux density at the effective wavelength for each filter for each object by multiplying the LSS photon count rate by the conversion factors in Table 10 (conversion from power-law spectra) of Poole et al. (2008).

The above equations can be applied to an individual sky image as well as a group of sky images. This process produces three measurements for each filter: photometry for each individual sky image, for all sky images in an OBSID, and for all OBSIDs of a filter. This information allows the construction of the light curve for each object at each available waveband. An example of the UVOT data for a quasar is displayed in Figure 4.9.

A total of 675 objects in our sample are detected by UVOT. We also calculate flux upper limits for objects covered by UVOT but without detection. In this work, we only use the UVOT-detected objects.

### 4.3 Quasar SEDs

We classify objects in the database into four types (Table 4.6) based on data availability. In the description below, “useful data” includes both detections and non-detections.

- Type A: The 637 objects that have useful data from both UVOT and XRT. We will construct SEDs for this set of objects. These objects do not necessarily have detections in all bands. Among these objects, 345 have detections by UVOT *and* XRT; they will be the focus of our analysis.
- Type B: the 38 objects in this type only have useful data from UVOT but *not* XRT; we will not create full band SED plots for objects in this type but we will construct the SED in UV/optical band.
- Type C: The reverse of Type B, objects that have only useful data from XRT but not UVOT. We will only create the XRT energy spectrum for these objects. There are 168 objects this type.
- Type D: The 191 objects objects are bereft of any useful data from UVOT or XRT. They are not included in the final catalog.

#### 4.3.1 Supplementary Data

As the object redshift increases, more UVOT wavebands are shifted into EUV band shortward of Ly $\alpha$ ; as a consequence, they cannot be used to measure the rest-frame UV power-law. The SDSS and 2MASS data, observed at longer wavelengths, can be used to extend the available measurements to UV and optical bands in the quasars’ rest-frames. We supplement our UV photometry with the five bands ( $ugriz$ ) from SDSS DR7 quasar catalog and three bands ( $J, H, K_s$ ) of 2MASS (Cohen et al. 2003; Skrutskie et al. 2006) survey, when available. In the best cases, we have 14 photometric data points in UV/Optical wavebands: six from *Swift*/UVOT, five from SDSS, and 3 from 2MASS (Table 4.7).

### 4.3.2 Initial SED Plots

The monochromatic luminosity corresponding to each UVOT filter can be easily computed after shifting the flux density to the quasar's rest-frame. The frequency band widths are calculated from rest-frame FWHMs of corresponding wavebands (Table 4.7).

The quasar catalog provides magnitudes that we must convert to flux densities. Instead of the *asinh* magnitude used in the general SDSS photometric measurements (Lupton et al. 1999), we convert the SDSS flux densities from corresponding band magnitude using the Pogson magnitude (Pogson 1857) which has a much simpler analytical expression  $m_1 = -2.5 \log f_1/f_0$ , where  $f_0$  is the zero-magnitude flux. The Pogson magnitude system deviates from the asinh system for faint objects. According to *asinh softening parameters (b coefficients)* table in the SDSS DR5 photometric calibration document<sup>6</sup>), the difference between Pogson and asinh magnitudes is less than 1% for objects brighter than  $g = 22.60$  and  $u = 22.12$ . In Figure 4.8, we plot the apparent magnitudes vs. redshift of all UVOT detected objects for all six UVOT filters. In Panel (b) and (c), we plot the magnitude limits of SDSS  $u$  and  $g$  bands as dash-dotted lines, which indicates that the majority of our objects are much brighter than limit. so the difference in Pogson and asinh systems is negligible. There is one object (SDSSJ122740.85+440604.7) whose B band magnitude is fainter than 22.60 and two objects (SDSSJ020316.37–074832.1, and SDSSJ133613.62+025703.8) whose U band magnitudes are fainter than 22.12. They are all at high redshift ( $z > 2.5$ ) with extremely low X-ray photon counts ( $N_{\text{Xph}} < 10$ ), so they are not included in 345 quasars of the main sample. For each waveband in SDSS, we adopt the central wavelength and FWHM from Fukugita et al. (1996). For each waveband in 2MASS, we adopt the isophotal central wavelengths and bandwidths from Cohen et al. (2003).

For X-ray sources with less than 100 photon counts, we use a different binning strategy in *XSPEC* from the one presented in Table 4.1 to avoid large error bars. We use the binning command `setplot rebin` to re-bin the spectrum until each bin has a detection at least as large as  $1\sigma$  and no more than 100 bins maybe so combined. Error types are set to `quad` which sums in quadrature the errors on the original bins. The `rebin` command only affects the plot appearance but not the fitting results (see Figure 4.5 for examples), but produces a clear representation of the average flux levels and associated uncertainties. Because the relatively large uncertainties in the X-ray energy spectrum, we calculate the lower and upper error bars of each flux point, instead of applying standard error propagation.

The initial quasar SEDs are generated based on all available UV/optical photometric data points and X-ray energy spectra. Before shifting into the rest-frame, we perform Galactic reddening correction to all available wave band flux densities using the standard  $E(B-V)$  dependent extinction curve of Fitzpatrick (1999). Values of  $E(B-V)$  are calculated following Schlegel et al. (1998). An example of an SED is shown in Figure 4.11.

These initial SEDs cannot be used to perform SED fitting due to two main reasons.

---

<sup>6</sup>[http://www.sdss.org/dr5/algorithms/fluxcal.html#asinh\\_table](http://www.sdss.org/dr5/algorithms/fluxcal.html#asinh_table)

- Broad emission lines can contribute significantly to broad band filter measurements, which may lead to incorrect SED shapes. For example, a UV band covering Ly $\alpha$  can alter the UV slope up to a value of 0.2.
- The dates of the observation from SDSS, 2MASS, and *Swift* for a given object differ, often by several years in the rest-frame. The data may need to be shifted to mitigate the effect of variability.

### 4.3.3 Emission Line Correction

Emission line corrections have been performed by subtracting the broad line contribution based on their average equivalent width (EW) (e.g., Elvis 2011, in preparation). Because we have SDSS UV/optical spectra for all quasars, we use a more sophisticated method that convolves the response function  $R(\lambda)$  of each filter with the observed spectrum  $f_{\text{tot}}(\lambda)$  and the power-law only spectrum  $f_{\text{pl}}(\lambda)$  to calculate the emission line correction factor.

$$\text{EC} = \log \frac{F_{\text{tot}}}{F_{\text{pl}}} = \log \int f_{\text{tot}}(\lambda) R(\lambda) d\lambda - \log \int f_{\text{pl}}(\lambda) R(\lambda) d\lambda \quad (4.6)$$

Because the quasar UV power-law usually extends from Ly $\alpha$  to  $\sim 5600$  Å (e.g., Vanden Berk et al. 2001), we only need to perform emission corrections to filters covered within this wavelength range. However, low-redshift quasar spectra do not cover Ly $\alpha$  and high-redshift quasar spectra do not cover 5600 Å. In these cases, the emission line corrections are performed based on the composite spectrum in Vanden Berk et al. (2001). Although the shape of the composite spectrum may not be exactly the same as the real spectrum, it does provide a *mean* correction.

To estimate the size of the errors introduced by performing emission line corrections using the composite spectra, we study the emission line correction trends as a function of redshift for different filters (see Figure 4.12). From these plots it is clear that the emission line correction is typically less than 0.1 dex. The real spectrum corrections are generally distributed around the composite spectrum, meaning they are generally in agreement. To view their differences more clearly, we plot the distributions of differences between these two correction,  $\Delta\text{EC} = \text{EC}_{\text{real}} - \text{EC}_{\text{composite}}$ , in Figure 4.13. These distributions indicate that on average, the composite spectrum emission line correction is consistent with the real spectrum emission line correction; the dispersion of these distributions are around 0.05. Therefore, we apply a systematic uncertainty of 0.05 to photometric points corrected using composite spectra. Figure 4.15 displays some examples of emission corrections for quasars at different redshifts. After emission line correction, the photometric slopes agree well with spectroscopic slopes in UV band with a median value of  $-0.43$ . In the upper panel of Figure 4.14, we plot these two slopes for objects with SED fitting. Most data points are distributed along the line representing  $\alpha_{\nu,\text{ph}} = \alpha_{\nu,\text{sp}}$  with some scatter. We argue that this scatter is mostly caused by low redshift quasars (typically  $z < 0.8$ ) in which the 2200 Å “line-free” rest-frame continuum point is not covered by SDSS spectra. In these cases, the uncertainty of the UV spectroscopic slope  $\alpha_{\nu,\text{sp}}$  may be larger than for high redshift quasar spectra. The value

of  $\alpha_{\nu,\text{sp}}$  can be more accurately measured for objects at higher redshifts, which explains why the dispersion of  $\alpha_{\nu,\text{ph}} - \alpha_{\nu,\text{sp}}$  is much smaller (*bottom panel*, Figure 4.14).

#### 4.3.4 Photometric Shift

The unique feature of our dataset is the simultaneous observations in UV and X-ray bands. Because of quasar variability, photometric observations in different dates may have different flux levels. The photometric data in SDSS and 2MASS are not observed in the same dates as the *Swift* UVOT. To make best use of these data, we perform some *pseudo*-simultaneous data points by shifting the observed flux levels to be consistent with the *Swift* UVOT photometry, assuming the UV/optical and IR SED shapes remain unchanged. With additional data points, we can make better constraints to UV slopes and luminosities while still approximately maintaining the simultaneous property of the dataset. The photometric shift follows strategies described below (see Table 4.8 and Figure 4.16).

- Because all quasars have five SDSS measurements we always interpolate or extrapolate the *ugriz* data to obtain fluxes at the UVOT filter effective wavelengths, then shift all SDSS and 2MASS photometry to match UVOT.
- We prefer to match in the UVOT U band. The *Swift* U band is very close to SDSS *u* band, which significantly reduces uncertainties introduced by interpolation/extrapolation. The U band, being bluer than the V and B bands, is less contaminated by host galaxy light.
- If *Swift* U band photometry is not available, we interpolate the SDSS *u* and *g* bands to match the *Swift* B band. If neither U nor B band is available, we interpolate the SDSS *r* and *g* band to match the *Swift* V band.
- If none of U, B, and V bands are available, we extrapolate the SDSS *g* and *u* band to match one of UVW1, UVM2 and UVW2 bands, using the available UVOT filter with the longer effective wavelength. Matching these filters is only done as a last resort; the significant extrapolation required inevitably introduces considerable photometric uncertainties. Further more, these three bands frequently lie in the Lyman forest where the quasar SED suffers from severe intrinsic and intervening absorption (e.g., Rauch 1998). This part of the continuum cannot be approximated as a single power-law.
- When selecting filters, we require both of the SDSS filters and the UVOT filter to fall in the UV range (between Ly $\alpha$  and 5600 Å) or within the EUV range, simply because the UV power-law cannot be extended to EUV region.

Examples of photometric shifting results are presented in Figure 4.17.

### 4.3.5 Error Analyses

For UVOT photometry, we adopt the photometric uncertainties produced by *uvot-source* and perform error propagation assuming these errors follow a Gaussian distribution. The error bars of SDSS and 2MASS fluxes are converted from magnitude uncertainties in SDSS DR7 quasar catalog. The frequency bandwidth plotted on the SED are converted from the corresponding FWHM of each filter.

For the XRT data, we use the parameter uncertainties produced by the `error` command. Because the X-ray photon counts follow Poisson distribution, we must calculate the upper and lower error bars separately.

### 4.3.6 SED Models

Because of the unavailability of the EUV data, which covers the BBB, the exact bolometric luminosities strongly depends on the model used to fit the SED. Traditionally, emission in this “gap” is represented by assuming a power-law continuum with a slope  $\alpha_{\text{ox}}$ . In this work, we consider a model which provides an upper limit of BBB in the canonical case. This shape is inspired from the presence of a soft X-ray excess over a flat X-ray component reported by Arnaud et al. (1985) in *EXOSAT* spectra of Seyfert 1 galaxy Mkn 841 and further observed and studied by Walter & Fink (1993) and Gierliński & Done (2004). Specifically, the latter study found that this soft X-ray excess can be well fit by a black body of energy 0.1 – 0.2 keV. This result motivates us to *constrain* this feature using a “bump” shape. In the following context, we call the first model the *exponential decay model* (hereafter EXP model). In addition, we use another SED model by connecting the high energy limit of the UV power-law at Ly $\alpha$  and the lower energy limit of the X-ray power-law at 0.3 keV. We call it the *triple power-law model* (hereafter TPL model) (Grupe et al. 2010). In either model, the total flux is contributed by a UV and an X-ray component.

$$f_{\nu,\text{tot}} = f_{\nu,\text{UV}} + f_{\nu,\text{X}}. \quad (4.7)$$

- EXP model. The two components are

$$f_{\nu,\text{UV}} = 10^{\beta_{\text{UV}}} \nu^{\alpha_{\text{UV}}} e^{(-h\nu/kT_B)^{\gamma_{\text{UV}}}} \quad (4.8)$$

$$f_{\nu,\text{X}} = 10^{\beta_{\text{X}}} \nu^{\alpha_{\text{X}}} e^{(-h\nu/kT_X)^{\gamma_{\text{X}}}} \quad (4.9)$$

The UV component is a power-law multiplied by an exponential decay term. The power-law slope  $\alpha_{\text{UV}}$  and scale factor  $\beta_{\text{UV}}$  are obtained by fitting photometric data points covered in the UV region (from 5600 Å to Ly $\alpha$ ). Canonically  $\alpha_{\text{UV}} \approx -0.4$  (e.g., Vanden Berk et al. 2001). The exponential term and the power-law term create a “bump” in the EUV region. The value of  $\gamma_{\text{UV}}$  controls how fast the flux declines given a value of  $T_B$  which is obtained from SED fitting. We choose to fix  $\gamma_{\text{UV}} = 1.5$  so that at  $\nu < \nu(\text{Ly}\alpha)$  the contribution from the exponential term is negligible and the total curve agrees well with the UV power-law given the best fit of  $T_B$ . For example, if we adopt  $\gamma_{\text{UV}} = 1$  Grupe et al. (2010), a typical value of  $T_B \sim 4$  Ryd as a result of the SED fitting produces a discrepancy between the UV power-law and the total SED curve by  $\sim 0.1$  dex at  $\nu(\text{Ly}\alpha)$ .

The X-ray component is also a power-law multiplied by an exponential decay term. The power-law slope  $\alpha_X$  and scale factor  $\beta_X$  are obtained by fitting the X-ray energy spectra using *XSPEC*. The X-ray decay energy  $T_X$  is fixed at 0.3 keV. We use the same value of  $\gamma_X = -8/3$  which assumes neutral hydrogen absorption (Longair 1992).

We emphasize that both of the UV and X-ray components are only mathematical expressions. For example, the value of  $T_B$  does not reflect the accretion disk temperature. Our goal is to place a reasonable upper limit to the BBB which contributes a significant fraction of the bolometric luminosity (e.g., Mathews & Ferland 1987; Zheng et al. 1997; Laor et al. 1997; Telfer et al. 2002; Shang et al. 2005). Therefore, the specific mathematical form is not important.

Because the XRT observes between 0.3 and 10 keV, these measurements are shifted to higher energy range for high redshift quasars, i.e., we lack the soft X-ray data points for these objects. This affects our SED fitting because if we only fit the observed data points, the decay energy of the bump will be shifted to higher energy band as the redshift becomes higher, which leads to additional flux contribution. To solve this problem, we supplement some artificial data points by extending the X-ray power-law from the minimum energy of the real data down to 0.3 keV. We use 0.3 keV as the lowest energy for the X-ray power-law based on previous X-ray studies of quasars. These investigations found that the X-ray power-law can extend from  $\sim 0.1$  keV to  $\gtrsim 10$  keV (e.g., Turner & Pounds 1989; Laor et al. 1997; George et al. 2000; Page et al. 2005; Young et al. 2009). A “soft X-ray excess” below  $\sim 0.3$  keV is found in the X-ray spectra of AGNs and quasars (e.g., Walter & Fink 1993; Crummy et al. 2006; Gierliński & Done 2004). We then fit all the data points with this SED model using the Levenberg-Marquardt algorithm, leaving  $T_B$  the only free parameter. The median value of  $T_B$  is 5.96 Ryd.

- TPL model. In this model, the UV and X-ray power-laws are fit in the same way as the EXP model. However, instead of multiplying an exponential decay term to each power-law, we simply connect the UV power-law at Ly $\alpha$  and X-ray power-law at 0.3 keV and define the slope as  $\alpha_{UVX}$ . The distribution of  $\alpha_{UVX}$  is presented in Figure 4.18.

In a canonical situation, e.g.,  $\alpha_{UV} = -0.4$  and  $\Gamma = 2$ , the EXP model produces a bump in the EUV region which produces an upper limit of the bolometric luminosity while the TPL model yields a lower-limit of the BBB. This may not be the case for strong X-ray absorption quasars. An example is shown in Panel (a) of Figure 4.19. In these cases, the absorption leads to a flat UV slope in addition to low X-ray emission in soft X-ray band. In these cases, the EXP model does not produce a “bump” but a “dip” in the EUV region and the corresponding integrated bolometric luminosity is not reliable. These quasars are flagged as “red” and will not be included into our cleaned sample defined below. Some examples of fitted SEDs using these two models are presented in Figure 4.19.

### 4.3.7 Bolometric Luminosity and Black Hole Mass

To obtain black hole masses, we must measure broad emission line widths which can be done with the SDSS spectra. We employ the same software package we used in previous SDSS spectral analyses (e.g., Wu et al. 2009; Vanden Berk et al. 2011, in preparation). We remove bad pixels based on the flags from the SDSS pipeline and perform Galactic reddening corrections to all the spectra using the  $E(B - V)$  dependent extinction curve of Fitzpatrick (1999). Values of  $E(B - V)$  are calculated following Schlegel et al. (1998). Each spectrum is shifted to the quasar's rest-frame. We first estimate the power-law continuum by connecting two empirical line-free regions. Next, we perform the joint fit to the underlining continuum using a single power-law in the UV band (between the Ly $\alpha$  emission line and  $\sim 5500$  Å), the small blue bump (between  $\sim 2000$  and  $\sim 4000$  Å), and the iron emission forest. We apply the UV iron template from Vestergaard & Wilkes (2001) and the optical iron template from Véron-Cetty et al. (2004). After we subtracting the underlying continuum, each broad emission line is fit by a single or multiple Gaussian profiles. Because one of the goals of this step is to compare the spectral fitting results using our soft package with those using the spectral fitting pipeline in Shen et al. (2008b), we do not limit the fitting spectra for 346 quasars with SEDs. Instead, we select 923 quasars from the raw catalog which contains 1034 quasars. The other quasar spectra are not fit for several reasons: (1) they are identified as BAL quasars; (2) their spectra have low  $S/N$ ; (3) their spectra are almost featureless (e.g., a BL Lac object); (4) the emission lines are extremely broad (e.g., double peaked H $\beta$ ).

Because of the SDSS spectra have a fixed observed wavelength region, the redshift range of the quasars in our sample does not allow us to select a set of features that are defected in all of our quasars. As a result, we need to use different black hole mass estimators at different redshifts. We adopt these black hole estimators from Shen et al. (2008b). Basically, at  $z < 0.7$ , we use H $\beta$  plus  $l_\lambda(5100$  Å), at  $0.7 < z < 1.9$ , we use Mg II plus  $l_\lambda(3000$  Å), and at  $z > 1.9$ , we use C IV plus  $l_\lambda(1350$  Å). The black hole mass is calculated using the following equation (Shemmer et al. 2006a; Shen et al. 2008b)

$$\log \left( \frac{M_{\text{BH}}}{M_\odot} \right) = a + b \log \left( \frac{\lambda L_\lambda}{10^{44} \text{ ergs s}^{-1}} \right) + 2 \log \left( \frac{\text{FWHM}}{\text{km s}^{-1}} \right) \quad (4.10)$$

in which  $(a, b) = (0.66, 0.53)$  for quasars with  $z < 0.7$ ,  $(a, b) = (0.505, 0.62)$  for quasars with  $0.7 < z < 1.9$  and  $(a, b) = (0.672, 0.61)$  for quasars with  $z > 1.9$ . The emission lines, continuum points and  $(a, b)$  values for different redshifts and their original references are tabulated in Table 4.9.

The bolometric luminosity is calculated using BC factors at 5100 Å for H $\beta$ , 3000 Å for Mg II, and 1350 Å for C IV using the composite SED for a sample of SDSS-DR3 quasars constructed by Richards et al. (2006), so that  $L_{\text{bol,SDSS}} = \text{BC} \cdot \lambda l_\lambda$ . The BC values are tabulated in Table 4.9.

As a check on our results, we compare our important correlations and distributions with Shen et al. (2008b). We first plot the black hole masses and bolometric luminosities of 923 quasars (Figure 4.20). The locations of quasars are consistent with the locations of quasars in Fig. 11 in Shen et al. (2008b), indicating that most quasars are accreting in a sub-Eddington level. The few quasars fall below  $R_{\text{Edd}} = 1$  limit (they are accreting

in super-Eddington level) are narrow line Seyfert 1 (NLS1) galaxies (Shen et al. 2008a; Grupe et al. 2010). Figure 4.21 shows the black hole mass as a function of redshift. This plot is consistent with Fig. 10 in Shen et al. (2008b). Finally, in Figure 4.25, we present the distributions of FWHM for the three emission lines that are crucial parameters for black hole mass estimation. The median values of  $\log \text{FWHM}/(\text{km s}^{-1})$  are 3.62, 3.63 and 3.75 for H $\beta$ , Mg II and C IV, respectively. The distributions and median values are again consistent with the results in Shen et al. (2008b). One can clearly see a potential source of bias in the calculation of black hole masses as the C IV FWHMs are systematically larger than those of H $\beta$  and Mg II. However, we still use the black hole masses estimated by C IV because this offset is small compared to the dispersion of the black hole mass distribution (see Figure 4.21).

The black hole masses are calculated using the emission line FWHMs and interpolated continuum luminosities at corresponding wavelengths from fitting the UV photometry. Their values are tabulated in the final catalog. In Figure 4.22, we compare the black hole masses calculated using two different estimators. It is clear that the black hole masses estimated from Mg II and H $\beta$  are consistent, but the C IV estimator, produces a black hole mass typically 0.2 dex higher.

The bolometric luminosities are calculated by integrating our two model curves from 5600 Å to 20 keV. In general, emission in this region contributes the majority of the quasar luminosity unless the quasar shows strong excess emission in the infrared band (Richards et al. 2006).

In Figure 4.26, we show the differences between the bolometric luminosities by integrating two SED models and the bolometric luminosities calculated using BC. The EXP model and the TPL model produce  $L_{\text{bol}}$  on average higher and lower than the  $L_{\text{bol}}$  from BC, respectively. Because the bump emission from the EXP model is more sensitive to the UV and X-ray observed spectral shapes than the TPL model, the EXP model has a larger dispersion. A fraction of the EXP models produce less bolometric luminosity ( $\log L_{\text{bol,EXP}} - \log L_{\text{bol,SDSS}} < 0$ ) than the TPL models, a fraction of these objects are quasars with strong intrinsic absorption (the black shaded area in Figure 4.26). The reason why the TPL model produces less bolometric luminosity on average than the BC method is because the bolometric correction by Richards et al. (2006) includes the infrared wavebands, which contributes about 37% of the total flux. After correcting contribution from this component, the entire histograms of either EXP and TPL model should move positively by  $\sim 0.2$  dex. Consequentially, the TPL model on average produces a bolometric luminosity consistent with that from the BC and the EXP model on average over produces  $L_{\text{bol}}$  by  $\sim 0.3$  dex.

#### 4.4 Catalog Description

Our results are tabulated in a catalog, available in the online edition, which contains 843 quasars with  $0.0129 \leq z \leq 4.5766$  and  $-30.24 \lesssim M_i \lesssim -22.01$ . There are 805 objects covered by UVOT, 675 covered by XRT and 637 covered by both UVOT and XRT observations. The X-ray detection rate is  $399/675 = \sim 59\%$ . Among these objects, 345 objects have both *detections* in UVOT and XRT, so that we are able to plot their SEDs from UV to X-ray band. These objects constitute our parent sample to

evaluate the flux contribution from the BBB in the next chapter. This catalog contains parameters directly measured from UVOT and XRT and quantities derived based on these measurements such as the black hole mass and bolometric luminosity. Each column in this catalog is described in Table 4.10.

#### 4.4.1 Comments on Individual Objects

The following sources possess special spectroscopic or photometric features or have unusual classifications.

- *SDSSJ102738.53+605016.5*. The optical spectrum of this object exhibits extremely broad and double-peaked H $\beta$ , and is cataloged as a double-peaked emission line quasar in Wu & Liu (2004). They derived the FWHM of H $\beta$  of  $\sim 16,200 \text{ km s}^{-1}$  and black hole mass of  $\log(M_{\text{BH}}/M_{\odot}) = 9.649$ .
- *SDSSJ141927.49+044513.8*. This object has a featureless UV spectrum with a few narrow absorption lines. It is a BL Lac object in the catalogs of Collinge et al. (2005), Plotkin et al. (2008) and Massaro et al. (2009). The narrow absorption lines are most likely to be intervening.
- *SDSSJ090821.01+045059.4*. The Mg II  $\lambda 2798$  emission line is barely detected, H $\beta$  is extremely weak and the two nearby [O III] lines appear very prominent, indicating that the BLR is obscured and only the NLR is seen. This object has been identified as a strong radio source (e.g., Griffith et al. 1995) but the exact classification is not determined, yet.
- *SDSSJ161742.53+322234.3*. This is the strong ratio quasar 3C 332. The H $\beta$  line is relatively weak, but the H $\alpha$  line exhibits a very prominent double peak structure with a FWHM of  $19600 \text{ km s}^{-1}$  measured by Strateva et al. (2006) and  $23200 \text{ km s}^{-1}$  measured by Wu & Liu (2004). The black hole mass is large,  $\log(M_{\text{BH}}/M_{\odot}) = 9.334$  (Wu & Liu 2004). The H $\alpha$  line also displays long-term profile variability which can be explained by a low, smooth, secular change in disk illumination (Gezari et al. 2007).
- *SDSSJ083148.87+042939.0*. This is a FSRQ (a subtype of blazar, see Section 1.1) with peak radio flux of  $\sim 1 \text{ Jy}$ . The optical spectrum exhibit a featureless continuum with an extremely weak H $\alpha$  emission line.
- *SDSSJ021702.66–082052.3*. This is another FSRQ as cataloged in Massaro et al. (2009). The optical spectrum exhibits a featureless continuum with very weak Mg II and H $\beta$  emission lines and a few narrow lines such as [O III].
- *SDSSJ074625.86+254902.1*. This is a high redshift FSRQ (Massaro et al. 2009) with  $z = 2.979$  and an MeV blazar discovered by *Swift* (Tueller et al. 2008; Jolley et al. 2009; Sambruna et al. 2006) and thus in the BAT selected AGN catalog by Winter et al. (2009). It is also observed by *Suzaku* (Watanabe et al. 2009). The UV spectrum contains broad emission lines such as Ly $\alpha$ , C IV and C III]  $\lambda 1908$ , which is unusual for a blazar.

- *SDSSJ154929.43+023701.1.* This object is classified as a FSRQ (Healey et al. 2007) with high polarization (Scarpa & Falomo 1997). The optical spectrum from SDSS, however, contains prominent Mg II, H $\beta$ , and even H  $\gamma$  lines. Even some weak forbidden lines are prominent, e.g., the two [Ne V] lines around 3400 Å. This is another case in which a blazar has a regular broad line quasar spectrum.
- *SDSSJ121826.51+294846.5.* This object, also known as Mkn 766 and NGC 4253, is a local Seyfert 1.5 galaxy ( $z = 0.013$ ) and is resolved in the UVOT image. Because of the strong host galaxy contamination, it is difficult to isolate the AGN component. It is not included in our UVOT processing list. The X-ray spectrum is complicated; it cannot be fit by any model we prescribed in Table 4.2. This object was previously studied in detail in the *Swift* AGN catalog by Grupe et al. (2010).
- *SDSSJ162901.30+400759.9.* This object is classified as a blazar in Massaro et al. (2009) or FSRQ by Falcone et al. (2004) but also classified as a NLS1 galaxy by Bade et al. (1995). The optical spectrum exhibits strong H $\alpha$ , H $\beta$ , and a complex of low ionization Fe II emission.
- *SDSSJ083740.24+245423.1.* This object is a blazar and classified as a FSRQ by Healey et al. (2007). The UV spectrum exhibits strong Mg II and C III]  $\lambda 1908$  emission lines.
- *SDSSJ101405.89+000620.3.* This quasar is classified as a Seyfert 1.8 by Dong et al. (2005). The optical spectrum exhibits a strong and double peaked H $\alpha$  but weak H $\beta$ . The excess emission over the power-law around these two lines is likely to be contributed by host galaxies.
- *SDSSJ213638.58+004154.1.* This  $z = 1.9414$  FSRQ has a strong radio jet (Healey et al. 2007; Liu & Zhang 2002; Massaro et al. 2009). The UV spectrum contains strong Mg II, C IV, and C III]  $\lambda 1908$  emission lines.
- *SDSSJ170231.06+324719.6.* This is a typical NLS1 (Zhou et al. 2006b; Grupe et al. 2010).
- *SDSSJ135516.54+561244.7.* This is a typical NLS1 (Zhou et al. 2006b; Grupe et al. 2010).
- *SDSSJ092703.01+390220.8.* This is a FSRQ with strong radio jet (Healey et al. 2007; Liu & Zhang 2002). The UV / optical spectrum also contains prominent Mg II, H $\beta$  and H  $\gamma$  emission lines.
- *SDSSJ074110.70+311200.2.* This is an FSRQ (Healey et al. 2007; Liu & Zhang 2002).
- *SDSSJ101810.98+354239.4.* This is an FSRQ (Healey et al. 2007).
- *SDSSJ103303.70+411606.2.* This is an FSRQ (Healey et al. 2007).

- *SDSSJ094215.12+090015.8*. This is quasar with double peaked H $\alpha$ . The FWHM of the broad component is extremely wide and reaches  $\sim 40,600 \text{ km s}^{-1}$ , which is the broadest known (Wang et al. 2005). The extremely broad H $\alpha$  suggests that the emission region is close to the black hole  $r \sim 80\text{--}100 r_g$  (Strateva et al. 2006).
- *SDSSJ142921.87+540611.2*. This object was classified as an FSRQ by Healey et al. (2007), and Massaro et al. (2009), and listed as a lensed quasar candidate in King et al. (1999). Browne et al. (2003) rejected the lensing hypothesis based upon surface brightness and spectral indices criteria.
- *SDSSJ081331.28+254503.0*. This is a confirmed lensed quasar (Williams et al. 2008; Congdon et al. 2010; Inada et al. 2010); it is excluded from our analyses because the amplified fluxes.
- *SDSSJ101541.14+594445.2*. This is a RL quasar exhibiting two pairs of radio lobes in an X-shape (Cheung 2007). The origin of the wings in these radio sources is unclear.

## 4.5 Correlation Analyses

In this section, we investigate the  $\alpha_{\text{ox}} - l_\nu(2500 \text{ \AA})$  relationship using a selected sample of Type 1 quasars from our catalog. This correlation has been described in a series of papers (Avni & Tananbaum 1982; Marshall et al. 1984; Tananbaum et al. 1986; Anderson & Margon 1987; Wilkes et al. 1994; Pickering et al. 1994; Avni et al. 1995; Vignali et al. 2003; Strateva et al. 2005; Steffen et al. 2006; Just et al. 2007); as quasars become more luminous, their relative UV and X-ray SEDs become softer, i.e., less X-ray emission with respect to UV emission. In the study of 372 type 1 quasars in Just et al. (2007; hereafter the J07 sample), the UV/optical data are from SDSS DR3 quasar catalog (Schneider et al. 2005), IUE (Boggess et al. 1978; Kondo et al. 1989) and COMBO-17 survey (Wolf et al. 2003) which covers the E-CDF-S (Wolf et al. 2004). The X-ray data are from *ROSAT*, *Chandra*, and *XMM-Newton*. As a result of this variety of observations, the UV/optical and X-ray data for an object frequently span a time scale of years. Quasar variability inevitably introduces scatters to  $\alpha_{\text{ox}} - l_\nu(2500 \text{ \AA})$  relation. In this section, we attempt to demonstrate that this scatter can be reduced by using simultaneous observed UV and X-ray data.

### 4.5.1 Sample Selection

In this work, we use 637 quasars with simultaneous observations by both UVOT and XRT. The sample size vs. exposure cut-off is presented in Figure 4.27 and the X-ray detection rate vs. exposure cut-off is presented in Figure 4.28. It is seen from Figure 4.28 that the X-ray detection rate is very low if we use the entire parent sample. Even if we supplement our dataset with that in Grupe et al. (2010), the detection rate is still less than 70%, i.e., over 30% of X-ray flux values are censored data (upper limits). These censored data can be handled by the ASURV software package but the large fraction of non-detections makes the results unreliable. On the other hand, if we use a very

long exposure time, the detection rate rises to  $> 95\%$ , but only a few objects remain in the sample and the statistics are very poor. Therefore, the selection of X-ray exposure time cut-off should balance the sample size as well as the X-ray detection rate. In this thesis, we use 20,000 s as the XRT exposure cut-off. The sample size with this constraint contains about 130 objects and the X-ray detection rate is increased to about 90%. We define this sample the *catalog parent* sample. In the following analysis, we supplement this sample with the AGNs in Grupe et al. (2010), which contains 88 *Swift* observed AGNs<sup>7</sup>. We define this sample *supplemental sample* or the *G10* sample. The parent sample is mixed with different types of quasars. To obtain a *cleaned catalog sample*, we exclude objects that fall into any of the following four categories. The number of quasars in each category is tabulated in Table 4.11. The cleaned catalog sample contains 89 objects with an X-ray detection rate 90%. There are 22 objects from the G10 sample after applying the 20 ks exposure cut-off. Because the G10 sample is X-ray selected, all objects are detected by XRT.

- RL quasars. They frequently have substantial X-ray flux from the radio jet, which leads to a higher  $\alpha_{\text{ox}}$  than those of radio quiet quasars (Worrall et al. 1987; Brinkmann et al. 2000).

We adopt the “radio loudness” ( $R^*$ ) defined by the ratio of monochromatic luminosities at rest-frame 5 GHz and 2500 Å (Stocke et al. 1992; della Ceca et al. 1994; Lu et al. 2007; Miller et al. 2009), e.g.,  $R^* = l_\nu(5 \text{ GHz})/l_\nu(2500 \text{ Å})$ . This definition differs slightly from the  $R$  criterion, which is defined using the *observed* radio flux at 5 GHz and  $B$  band power (e.g., a Parkes radio quasar; Kellermann 1989). Because the former provides a more consistent estimate of the *intrinsic* ratio of radio to optical power among quasars at different redshifts, we adopt  $R^*$  throughout the paper.

For most of quasars in the parent sample, we calculate the  $k$ -corrected  $l_\nu(5 \text{ GHz})$  from the 20 cm peak flux from the Very Large Array (VLA) FIRST Survey (Becker et al. 1995; White et al. 1997) cataloged in the SDSS DR7 quasar catalog. The equation to convert a radio flux from the observed frequency  $\nu$  to the  $k$ -corrected frequency  $\nu_1$  assuming a constant power-law index  $\alpha_r$  is

$$\log f_{\nu_1} = \log f_\nu + \alpha_r \log \frac{\nu_1}{\nu} - (1 + \alpha_r) \log (1 + z). \quad (4.11)$$

In our case,  $\nu_1 = 5 \text{ GHz}$ ,  $\nu = 1.5 \text{ GHz}$  (corresponding to 20 cm). Here, we assume an average value of radio spectral index  $\alpha_r = -0.5$  (e.g., Kellermann et al. 1989; Komossa et al. 2006; Lu et al. 2007).

For objects not detected by the FIRST survey (but still covered), we estimate radio luminosity upper limits using the sensitivity limit of the FIRST survey, which is 1 mJy, which provides upper limits of  $R^*$ . If  $\log R^* < 1$ , we classify this object as RQ. If  $\log R^* \geq 1$ , we cannot simply claim it as RL because  $R^*$  is a *lower* limit. In this situation, we refer to other resources of radio surveys, such as NRAO VLA Sky Survey (NVSS; Condon et al. 1998).

---

<sup>7</sup>There are 92 objects in total, but 4 objects were not observed by UVOT, so they are excluded.

For objects not covered by the FIRST survey, we first check the published literature and see if this object was detected in previous radio observations. If no observation is found, we examine 15' NVSS image for strong nearby radio source. If we cannot confirm the quasar as RQ, we conservatively exclude it from the cleaned sample.

We preferentially adopt  $l_\nu(2500 \text{ \AA})$  from SDSS spectral fitting results, unless 2500 Å is not covered, in which situation we use  $l_\nu(2500 \text{ \AA})$  calculated from the SED fitting.

- Lensed quasar. The flux from these quasars are amplified, so these objects are removed from the cleaned sample (e.g., Just et al. 2007) because of flux uncertainty. There is only one confirmed lensed quasar in our sample: SDSSJ142921.87+540611.2 (see Section 4.4.1).
- Blazars, including BL Lac objects and FSRQs, usually have large amplitude variations in UV/optical (e.g., Rani et al. 2010) as well as strong radio and X-ray emissions (e.g., Beckmann et al. 2003; Padovani et al. 2007). As a result, the values of  $\alpha_{\text{ox}}$  can vary significantly over time. The UV/optical spectrum of BL Lac objects usually have a featureless continuum without any emission lines (e.g., Blandford & Rees 1978; Kollgaard 1994; Plotkin et al. 2008; Abdo et al. 2010). Many of them are RL and their emissions are believed to be relativistically beamed from the jet (Urry & Padovani 1995). Although some FSRQs exhibit similar broad emission line features as Type 1 quasars, their radio emission is still beamed and variable. These objects are also excluded from the cleaned sample.
- Reddened quasars. Our sample contains a number of quasars with shallow UV/optical slopes and/or strong soft X-ray decline (e.g., Figure 4.23). The relatively flat UV/optical spectra are attributed to dust reddening (Richards 2001; Richards et al. 2003; Hopkins et al. 2004), although observations of individual objects suggest that some slopes could be intrinsically steep (e.g., Hall et al. 2006). The standard reddening correction is to measure the Balmer decrement and to assume an intrinsic H $\alpha$ /H $\beta$  value, e.g., 2.86 (Osterbrock & Ferland 1989; Peterson 1997). However, this value assumes a low density region and optically thin emission lines (Baker & Menzel 1938), which may be applicable to planetary nebulae but not quasar BLRs, because the gas density of the latter is much higher and radiative transfer effects are not the same as in low density regime (e.g., Hao et al. 2005). In this work, we only flag these objects and exclude them from the cleaned sample.

Following Richards et al. (2003), we use the relative color to define the “dust-reddened” quasars. The *relative color* is calculated by subtracting the median colors of quasars at the redshift of each quasar from measured colors of each quasar. As argued by Richards et al. (2003), the quantity  $\Delta(g-i) = (g-i) - \text{Median}(g-i)_z$  is an excellent redshift-independent surrogate for the photometric spectral index. In Figure 4.24, we plot  $\Delta(g-i)$  vs. redshift for all the objects in our raw catalog. Objects to the right of the dashed line are flagged as “dust-reddened” quasars. These quasars, which comprise 8.5% of the raw catalog, are excluded from our Cleaned sample.

### 4.5.2 The $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$ Relation

In this section, we compare the dispersions between the J07 sample and our cleaned catalog sample. We show that the dispersion can be reduced using simultaneous UV and X-ray observations. Because the exact values of dispersion could be method-dependent, we use two methods in our analyses.

To enlarge our sample size, we supplement our cleaned catalog sample with the sample in Grupe et al. (2010). The G10 sample also contains simultaneous observations from *Swift* UVOT and XRT. The G10 sample is soft X-ray selected and is composed of low redshift AGNs ( $z \lesssim 0.4$ ). Therefore, these objects are mostly low luminosity AGNs and thus complements to our cleaned sample, which are relatively intrinsically bright. Finally, the G10 sample primarily consists of Type 1 AGNs, with only a few exceptions. After removing the overlap with our sample and two RL objects in G10, we construct a combined sample which consists of 111 objects. We will present two sets of results based on both of the cleaned catalog sample and the combined sample.

Because our sample contains censored data, we use ASURV software package (Lavalley et al. 1992) to perform statistical analysis. This package include the Expectation-Maximization (EM) and the Buckley-James (BJ) algorithms, which we can use to perform linear regression and dispersion estimation. The methods differ in that the EM algorithm estimates the residual assuming a Gaussian distribution while the BJ algorithm assumes the Kaplan-Meire distribution. The correlation and regression results are tabulated in Table 4.12. The results indicate that

- The cleaned catalog and the combined sample both exhibit strong correlations between  $\alpha_{\text{ox}}$  and  $l_{\nu}(2500 \text{ \AA})$ , although the absolute values of correlation coefficients are a bit lower than the J07 sample.
- The slope of the cleaned sample ( $-0.1532 \pm 0.0125$ ) is steeper than the J07 sample, the intercept is also larger, but they are both consistent within  $1 \sigma$  error bars. The combined sample has a shallower slope than the cleaned catalog sample which is caused by the G10 sample. The possibility that the  $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$  relation is non-linear is proposed in the study of Wilkes et al. (1994) and Anderson et al. (2003). As discussed in Wilkes et al. (1994), the difference in slopes is likely caused by the varying host galaxy contribution to the  $l_{\nu}(2500 \text{ \AA})$  measurement at low redshift. This may be the reason for the G10 sample since host galaxy contribution is not corrected<sup>8</sup>. Strateva et al. (2005) found that their sample does not offer significant evidence for a non-linear  $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$  relation. Although they obtained a shallower slope of  $-0.09 \pm 0.02$  for the low luminosity ( $l_{\text{UV}} < 30.5$ ) and  $-0.13 \pm 0.02$  for the high luminosity ( $l_{\text{UV}} > 30.5$ ) sample, they argue that the difference in slopes is likely an artifact of the addition of five outlier AGNs, which are probably X-ray-absorbed Seyfert galaxies at  $z < 0.22$ . From Figure 4.29, we do not see low luminosity “outliers”, but we still cannot exclude the possibility that the shallower slope is caused by host galaxy contamination. In addition, we also

---

<sup>8</sup>We do not correct for host galaxy light in our sample, but since the majority of our objects are luminous quasars at higher redshift than G10, the host galaxy contamination is much smaller.

note that the J10 sample contains a number of the most luminous quasars with  $l_{\text{UV}} > 31.5$ , but excluding these high luminosity quasars does not reduce the slope significantly (see regression results of J07T<sup>9</sup> in Table 4.12). A careful removal of host galaxy contribution is probably necessary for further check, e.g., by using *GalFit* (Peng et al. 2002, 2010).

- From Table 4.12, it is clear that the standard deviation of the cleaned catalog sample is smaller than the J07 sample by about 13%-19%. This dispersion is reduced by 18%-25% using the combined sample because the G10 sample has a higher degree of simultaneity.
- The dispersion of the  $l_{\nu}(2500 \text{ \AA})-\alpha_{\text{ox}}$  relation does not exhibit a luminosity dependence (see Figure 7 of Just et al. 2007). From Figure 4.29, we do not see the dispersion showing any evident dependence on luminosity, either.

## 4.6 Conclusions

We have compiled a serendipitous quasar catalog with simultaneous UV/optical and X-ray observations by *Swift*. The catalog is generated by matching the  $\sim 3.5$  year *Swift* pointings from November, 2004 to June, 2008 to the SDSS DR5 quasar catalog. For each object, the sky images observed by either UVOT or XRT are carefully selected to ensure image quality. In this work, we derive the composite UVOT photometry and XRT energy spectra by stacking all archival data to generate the deepest sky images. The resultant SEDs reflect the time-averaged shape of quasar emission with simultaneous observations at multi-wavebands. The final catalog contains 843 objects (see Table 4.6). There are 637 objects ( $\sim 76\%$ ) that have UVOT and XRT observations, 168 objects ( $\sim 20\%$ ) that only have XRT data and 38 objects that only have UVOT data ( $\sim 4\%$ ). Among all the 675 objects with X-ray coverage, 399 ( $\sim 59\%$ ) are detected. We construct SED plots for all objects with both XRT and UVOT data. In this thesis, we focus on 637 objects with both X-ray and UV observations. We supplement UVOT photometry with SDSS and 2MASS data if available. All the photometric points are corrected for the effects of emission lines and those from SDSS and 2MASS shifted to match the flux levels of *Swift* UVOT.

We fit SEDs using the EXP and TPL models, attempting to constrain the flux contribution from the BBB. In most cases, the EXP models creates a bump in the EUV region, producing an upper limit of the BBB emission, while the TPL model connects the SED points at 1216 Å and 0.3 keV, producing a lower limit of the BBB emission. After correcting the contribution from IR emission, the TPL model produces bolometric luminosities consistent with those estimated from BC while the EXP model produces bolometric luminosities on average 0.3 dex higher than the TPL model.

We further select a cleaned sample comprised of 89 typical RQ quasars. We also supplement this cleaned sample with 22 nearby AGNs from Grupe et al. (2010). We re-visit the  $\alpha_{\text{ox}}-l_{\nu}(2500 \text{ \AA})$  relation presented by Just et al. (2007), which is the best

---

<sup>9</sup>Low luminosity sample  $\log l_{\nu}(2500 \text{ \AA}) \leq 31$  in J07.

sample so far to investigate this correlation. We use two methods to compare the intrinsic scatters of the  $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$  relationship of our and the J07 sample using EM and James-Buckley methods. These two methods consistently indicate that the dispersion based on our sample is reduced by 13% to 19% using the cleaned catalog sample and 18% to 25% usinig the combined sample. This is the first time the reduced scatter in this relation has been seen. The reduction amount is on average smaller than expected ( $\sigma_{\alpha_{\text{ox}}} = 0.15$  as estimated by Strateva et al. 2005), but it can be potentially increased by using observations spanning a relatively short period of time. Firmly establishing the  $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$  emission in AGNs is an important step toward understanding energy generation mechanisms of AGNs. Our work further verified this correlation and validate that this correlation is intrinsically even tighter by reducing or eliminating scatters result from variabilities.

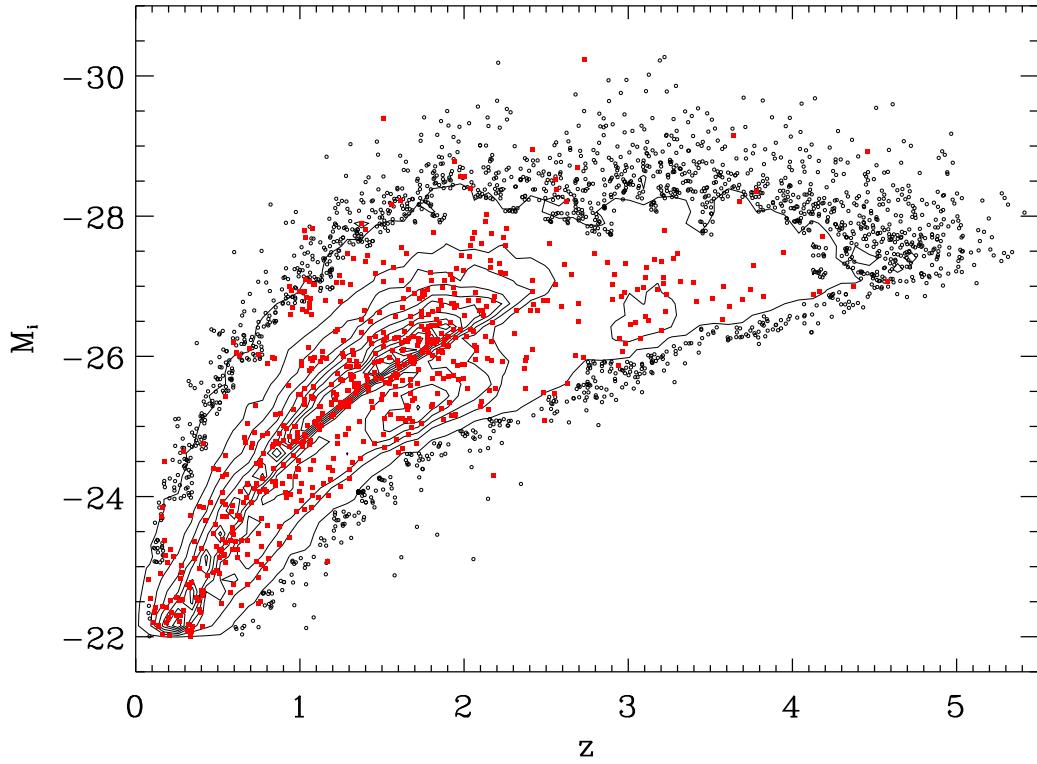


Fig. 4.1 Distribution of our sample on in the redshift vs. SDSS  $i$  band absolute magnitude diagram. The SDSS DR5 quasar catalog objects are represented by open circles. Their distribution is represented by a set of linear contours when the density of open circles in this two-dimensional space exceeds a certain threshold and circles begin to overlap. Objects in the *Swift* quasar catalog are represented by red filled squares. The luminosity limit is because SDSS DR5 selects quasars more luminous than  $M_i = 22.0$ .

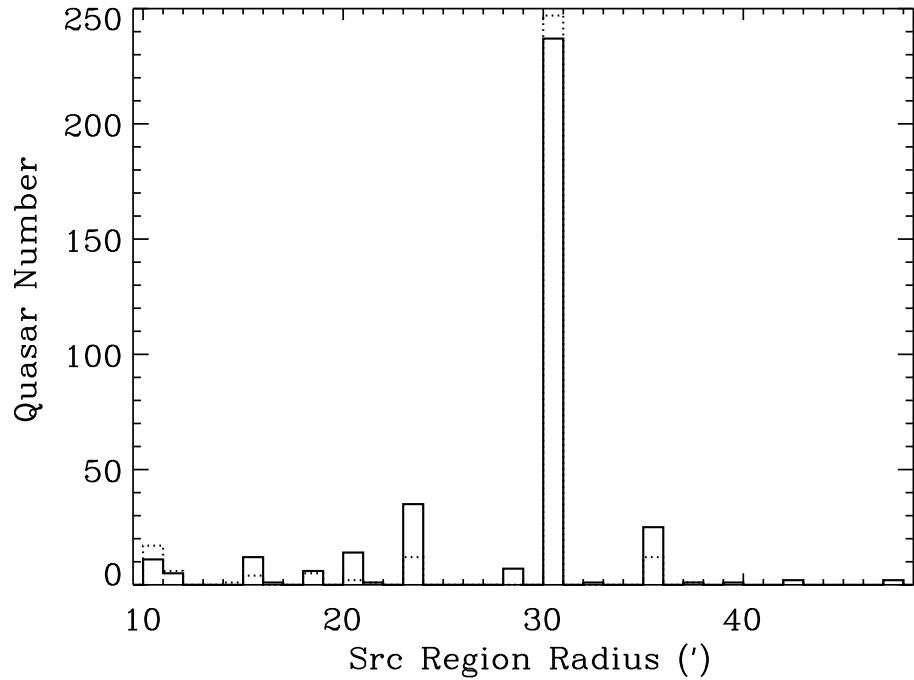


Fig. 4.2 Distribution of source region radii for reference object candidate (solid), including 263 objects bright X-ray sources modeled with Model A and B ( $\gamma \sim 2$ ) and 98 objects detected weak X-ray sources. The dotted histogram shows the distribution of all undetected objects (308 objects). We use  $30''$  aperture size for most objects.

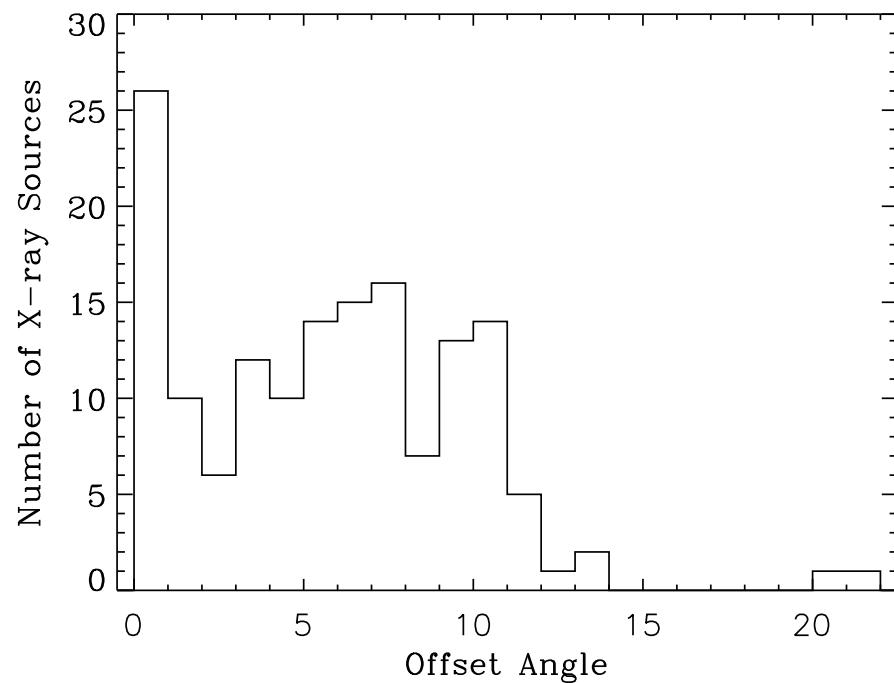


Fig. 4.3 Distribution of offset angle  $\theta$  (in arc minute) for 153 reference objects selected.

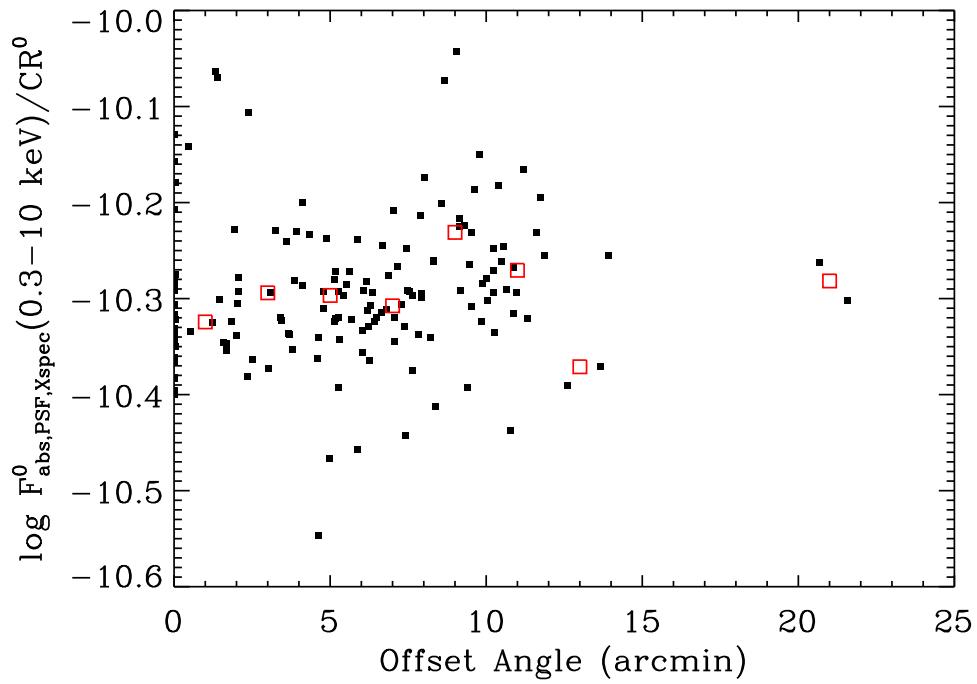


Fig. 4.4 Ratio of X-ray flux between 0.3 and 10 keV to count rate vs. offset angle  $\theta$  for reference objects (black squares) and binned curve (red boxes). The bin size is set to be  $1'$ . Note that the  $y$ -axis is in logarithmic scale.

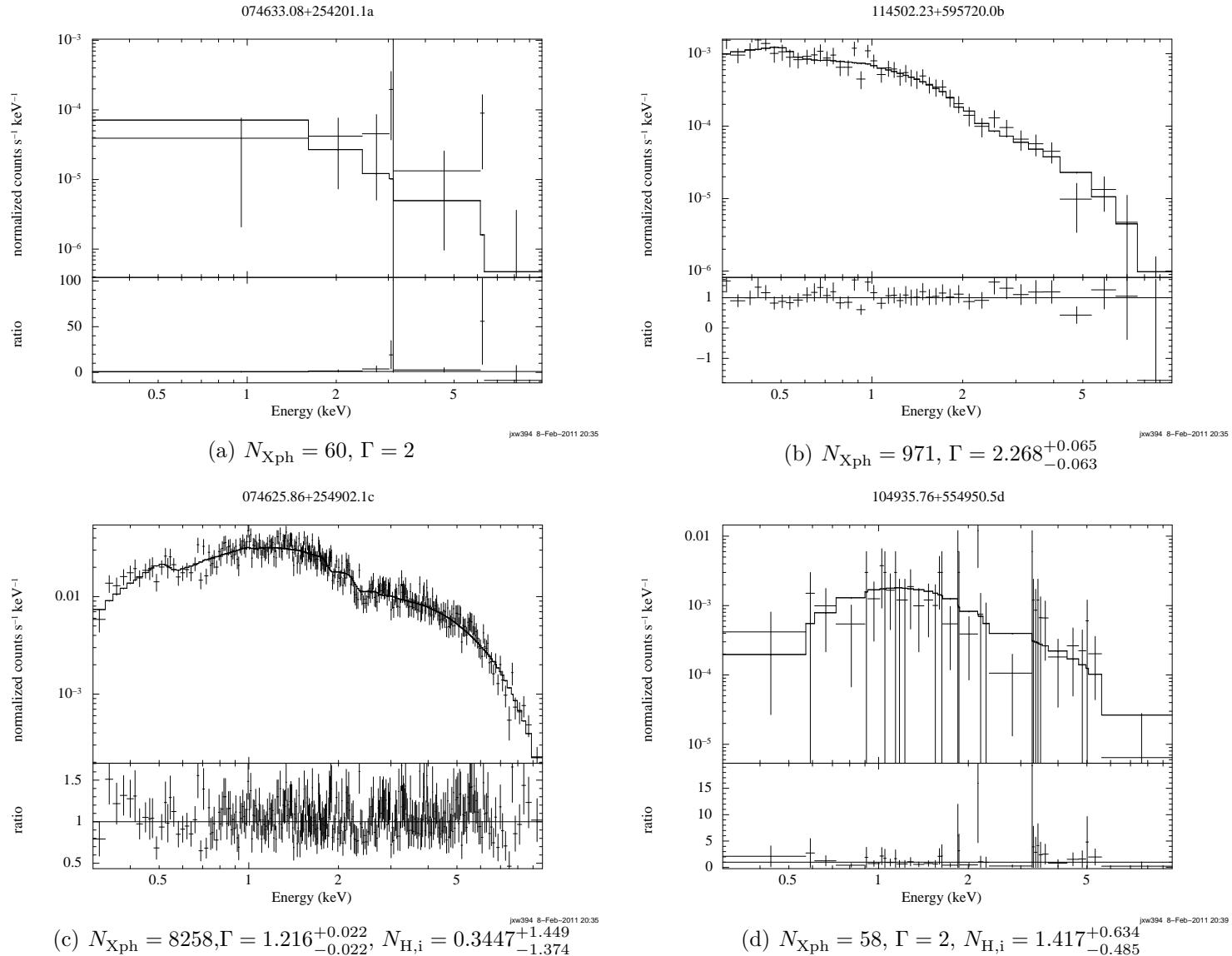


Fig. 4.5 Four examples of fitting the observed XRT spectra. Column densities ( $N_{\text{H,i}}$ ) are in units of  $10^{22} \text{ cm}^{-2}$ . Spectra shown in (b) and (c) are binned as Table 4.1. Spectra shown in (a) and (d) rebinned using Xspec command `setplot rebin` (see Section 4.3.2)

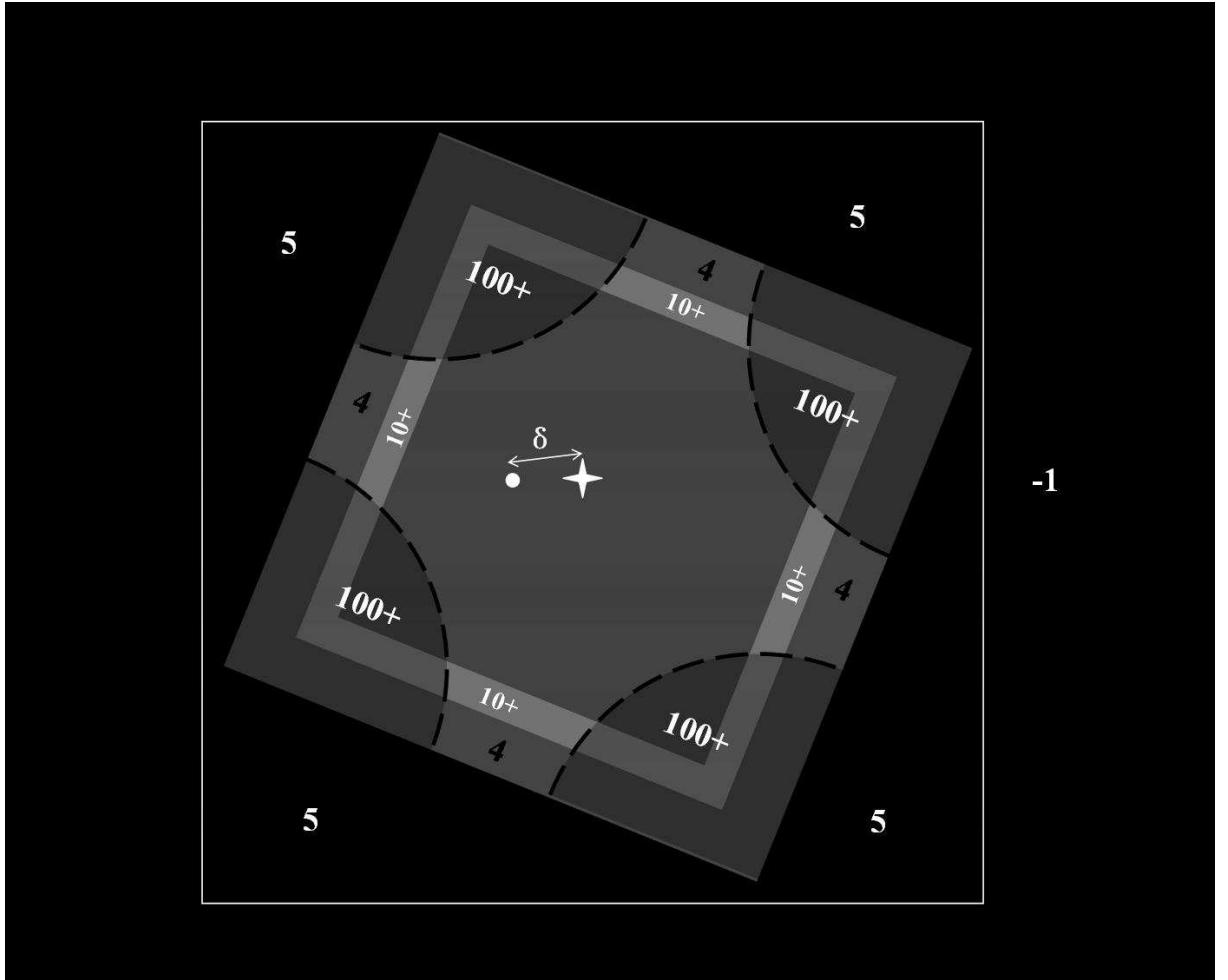


Fig. 4.6 UVOT sky image flagging. The white rectangle represents the image file. The largest tilted square is the FOV. The distance between the second and the largest tilted square is  $r = 27''.5$ , and the distance between the third and the most biggest tilted square is  $R = 35''.0$ . The radius of the sector at each corner is  $100''$ . The separation between the SDSS quasar coordinates (white dot) and the two dimensional Gaussian centroid (white star) is denoted as  $\delta$ . The figure is not drawn in scale. Objects falling into the various regions are flagged as the numbers in the figure. Objects covered by the sectors and the grey band areas are added by 100 and 10, respectively.

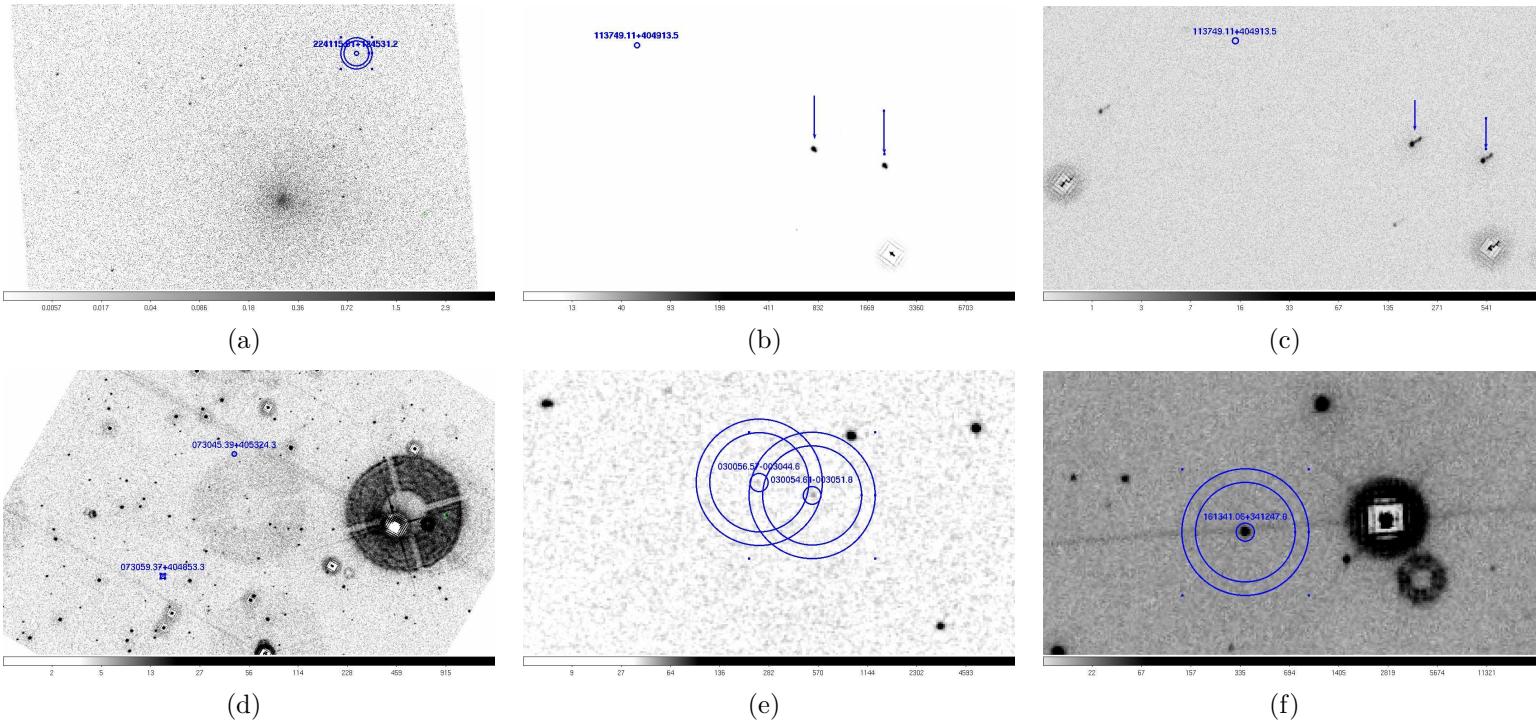


Fig. 4.7 Examples of removed UVOT sky images in visual inspection. In which panel, the image has a white background and black UV/optical sources. The small circle is the  $3''$  source region; the annular is the background region. (a) A very short exposure with a comet and a quasar from our sample; (b) An example of an image which suffered from a dual aspect solution either made by the spacecraft or by ground software. These images were flagged specially apart from images where the spacecraft pointing was unstable because it is possible to fix *some* of these images if they were taken originally in event mode; (c) An example of spacecraft pointing instability; (d) One quasar is embedded in the bright portion of the halo. Another is embedded in the less bright reflection of the halo; (e) One object falls within the background circle of another one; (f) A charge trail from a bright star runs right through the source.

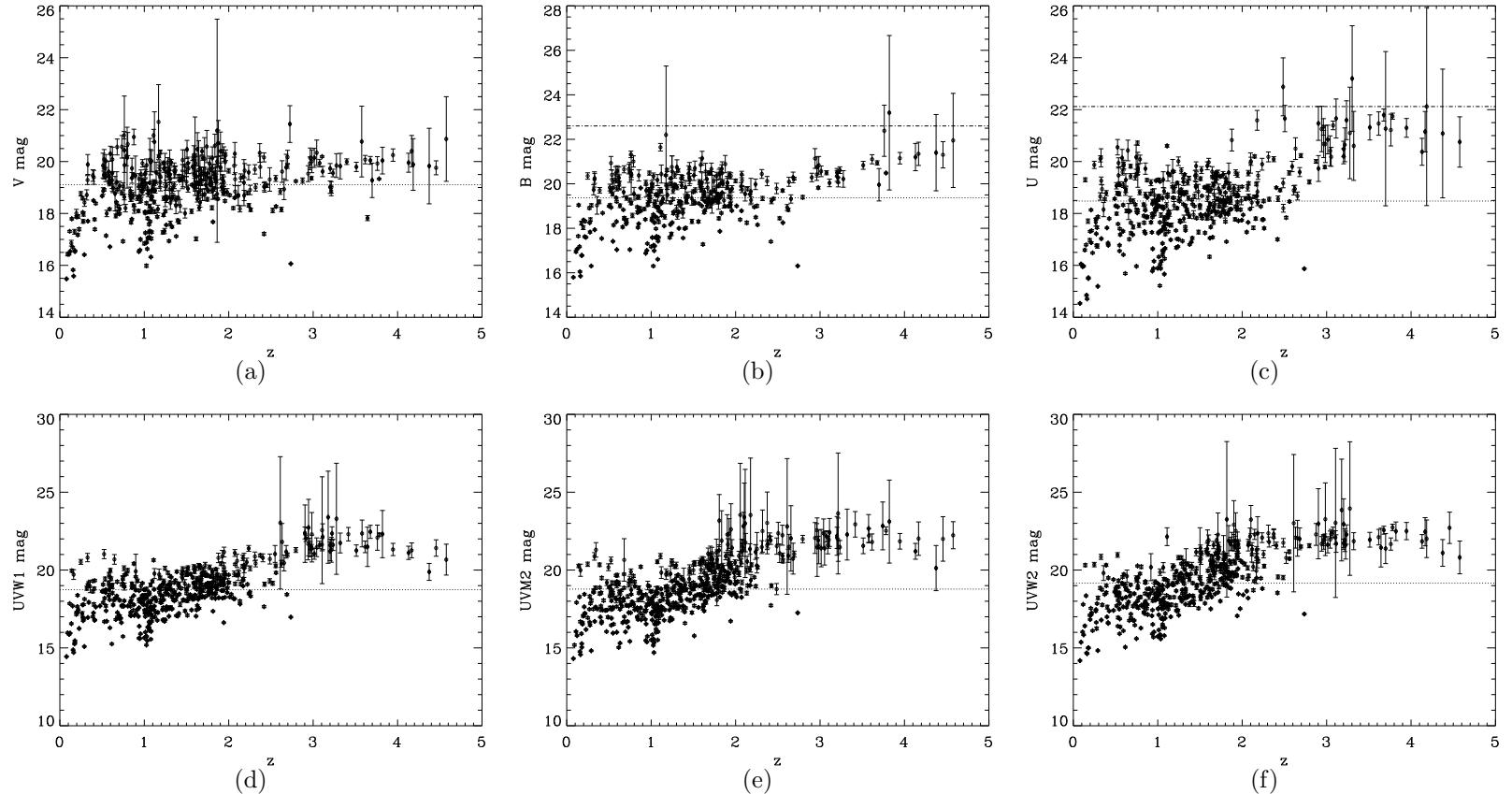


Fig. 4.8 Apparent magnitudes vs. redshift for quasars in our sample. In each panel, each data point is represented by an open circle with an error bar. The dotted line represents the median value of magnitude of a band. In Panel (b) and (c), we over plot a dot-dashed line in which plot which represents the SDSS  $g$  and  $u$  band magnitude limits. The difference between the Pogson and asinh magnitudes is less than 1% for objects brighter than these limits. The UVOT U and B bands effective wavelengths are close to the effective wavelengths of SDSS  $u$  and  $g$  band.

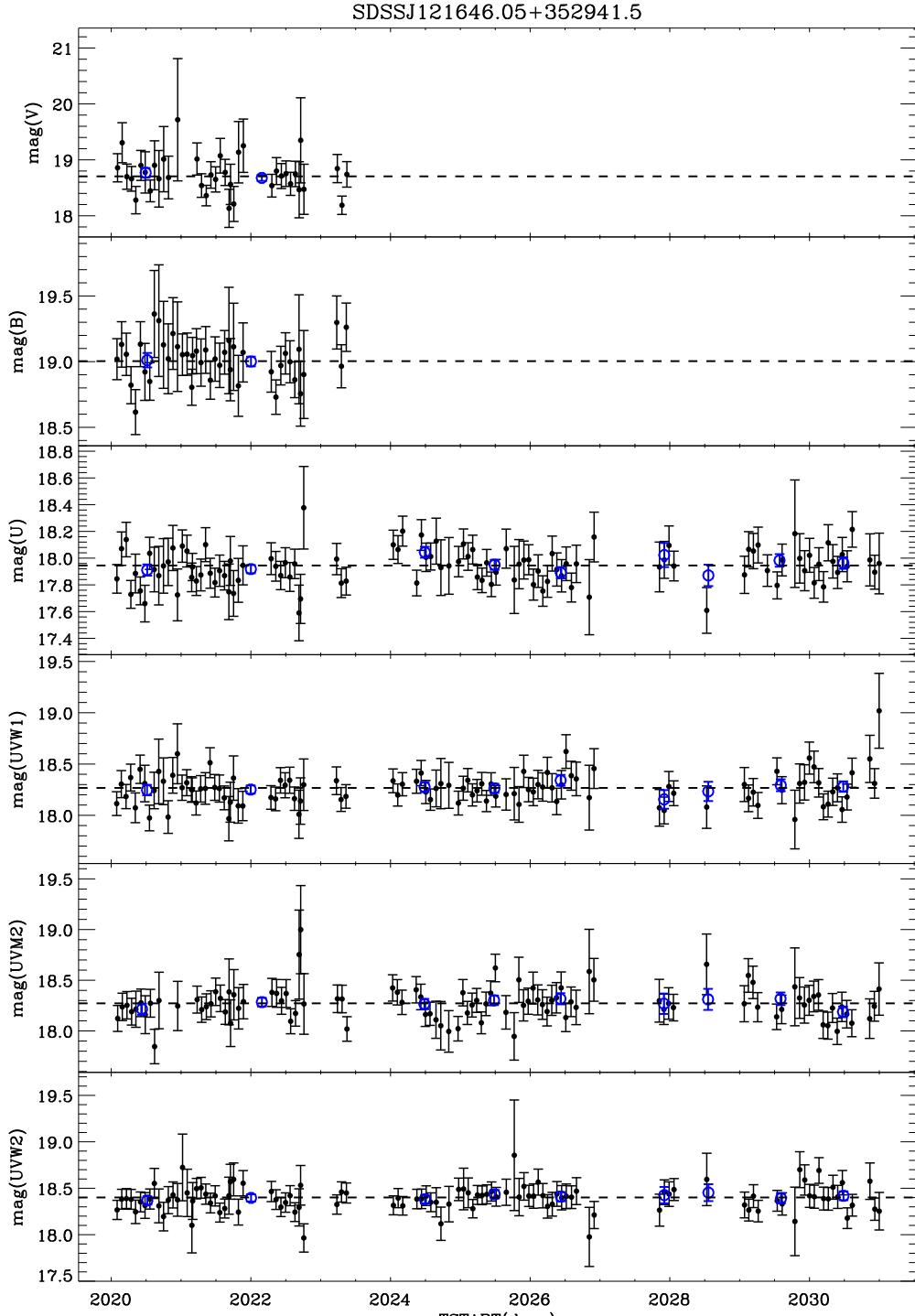


Fig. 4.9 The light curve for SDSSJ121646.05+352941.5 in all six UVOT filters. The black dots are apparent magnitudes for each sky image, and the red circles are averaged apparent magnitudes over an OBSID. The dashed line represents the apparent magnitude of all the sky images in a certain filter. The x-axis is the mission elapsed time in days.

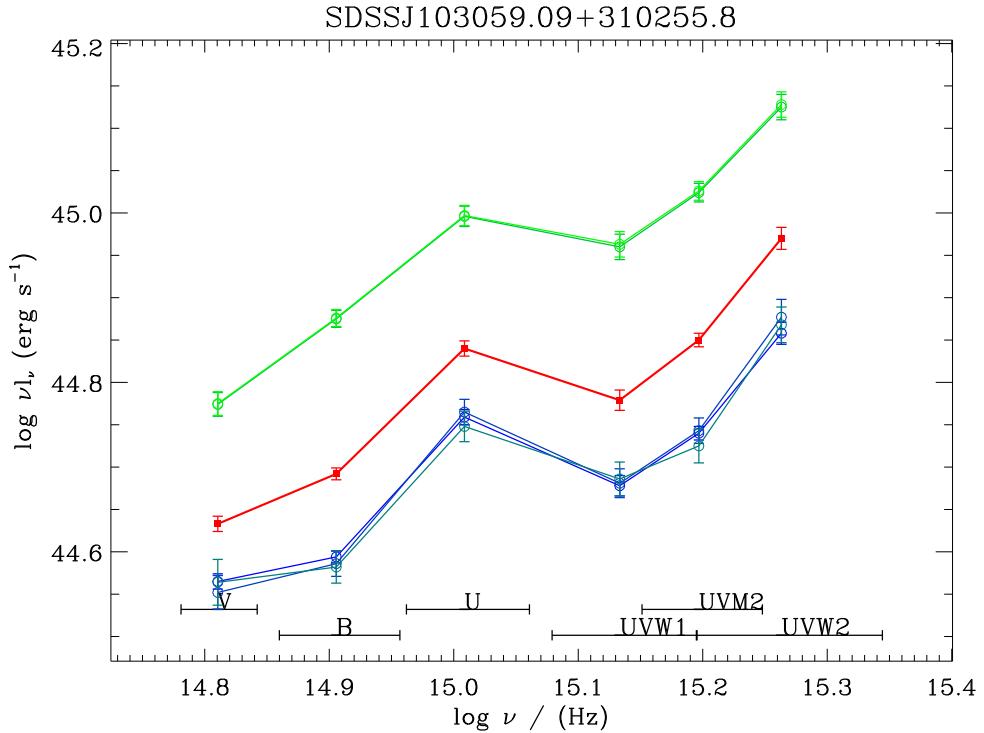


Fig. 4.10 An example showing the SED in the UV band of a quasar changes as a function of time. The color gradually changes from blue to green as the rest-frame time elapses. The rest-frame times (in days) of each curve relative from the earliest (bluest) curve are 0 (earliest), 2.21, 2.91, 1051.22 and 1051.22 (the last two were almost observed in the same cycle). The red curve is plotted based on the arithmetic mean of monochromatic luminosity at each waveband. The range of a waveband is shown as the FWHM of that band.

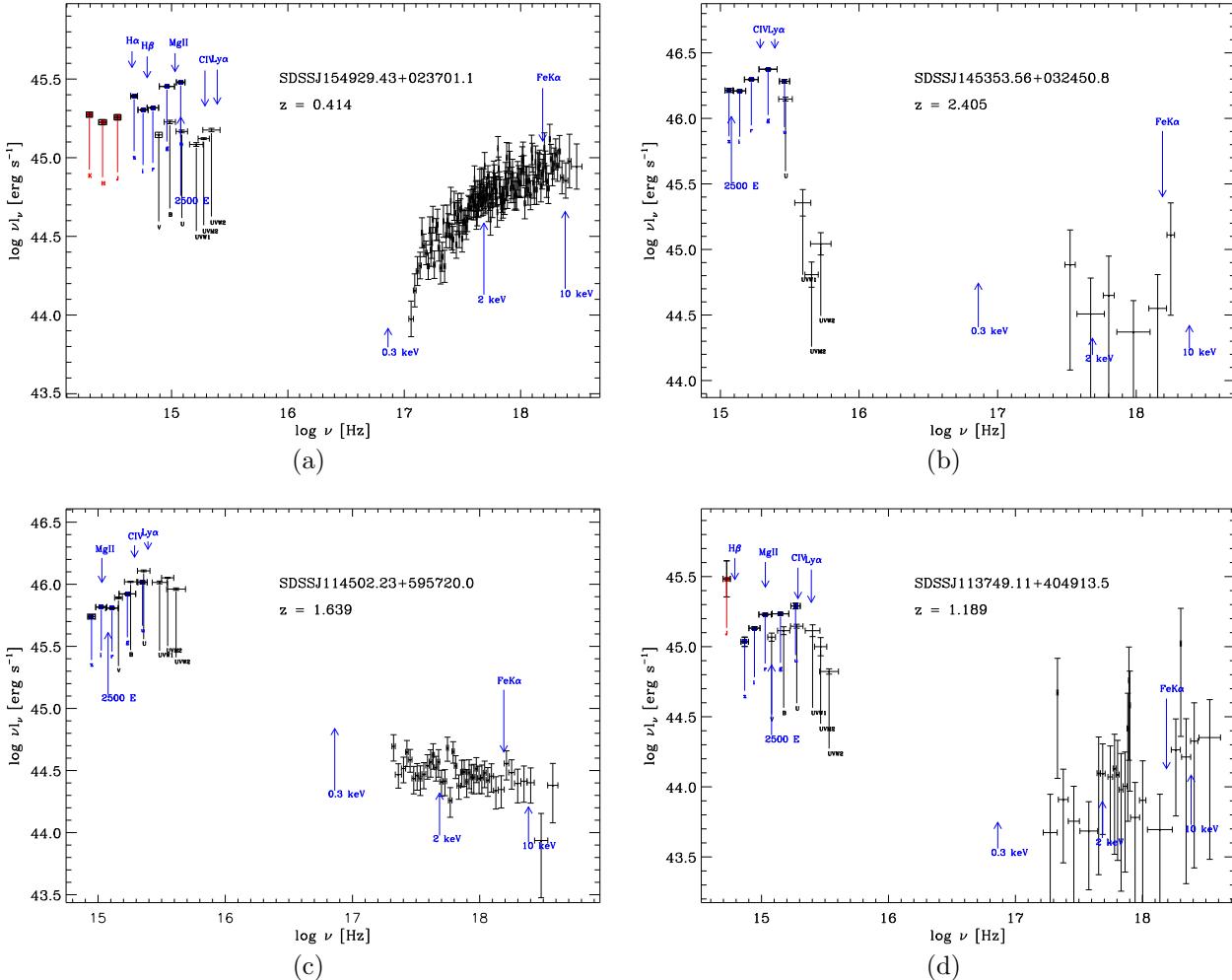


Fig. 4.11 Examples of initial SEDs, showing data from *Swift* UVOT (black), SDSS (blue) and 2MASS (red). There are clear flux offsets between SDSS and UVOT measurements.

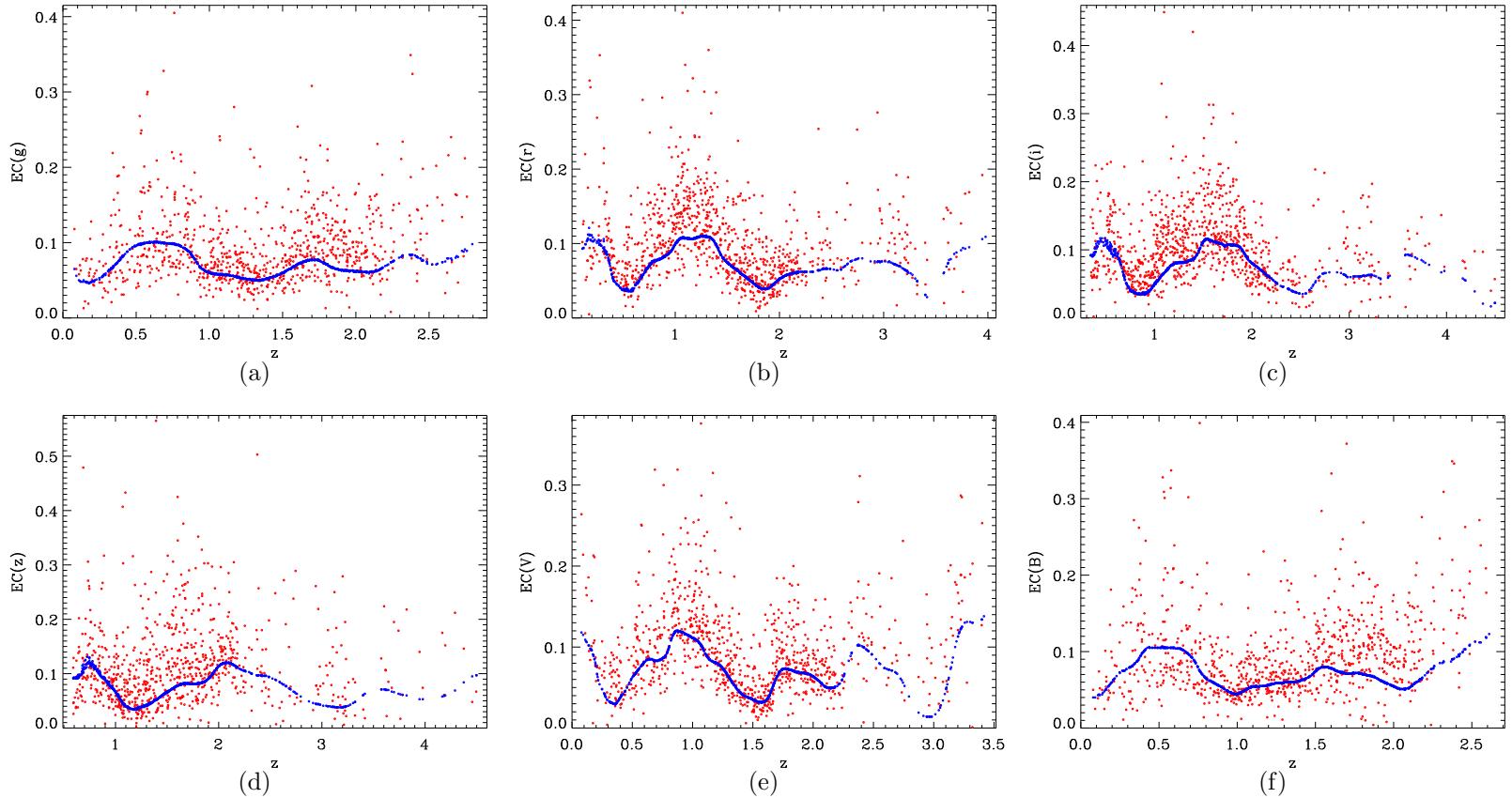


Fig. 4.12 Emission line corrections (EC) as a function of redshift for different filters. Red circles are EC performed on real spectra. Blue circles are EC performed on the composite spectrum by (Vanden Berk et al. 2001) shifted to the real spectrum redshift. The EC represented by blue circles varies with redshift as an emission line feature is shifted within the coverage of a filter. These plots indicate that the EC based on real spectra are in general consistent with EC based on composite spectra. The deviations between them are evaluated in Figure 4.13.

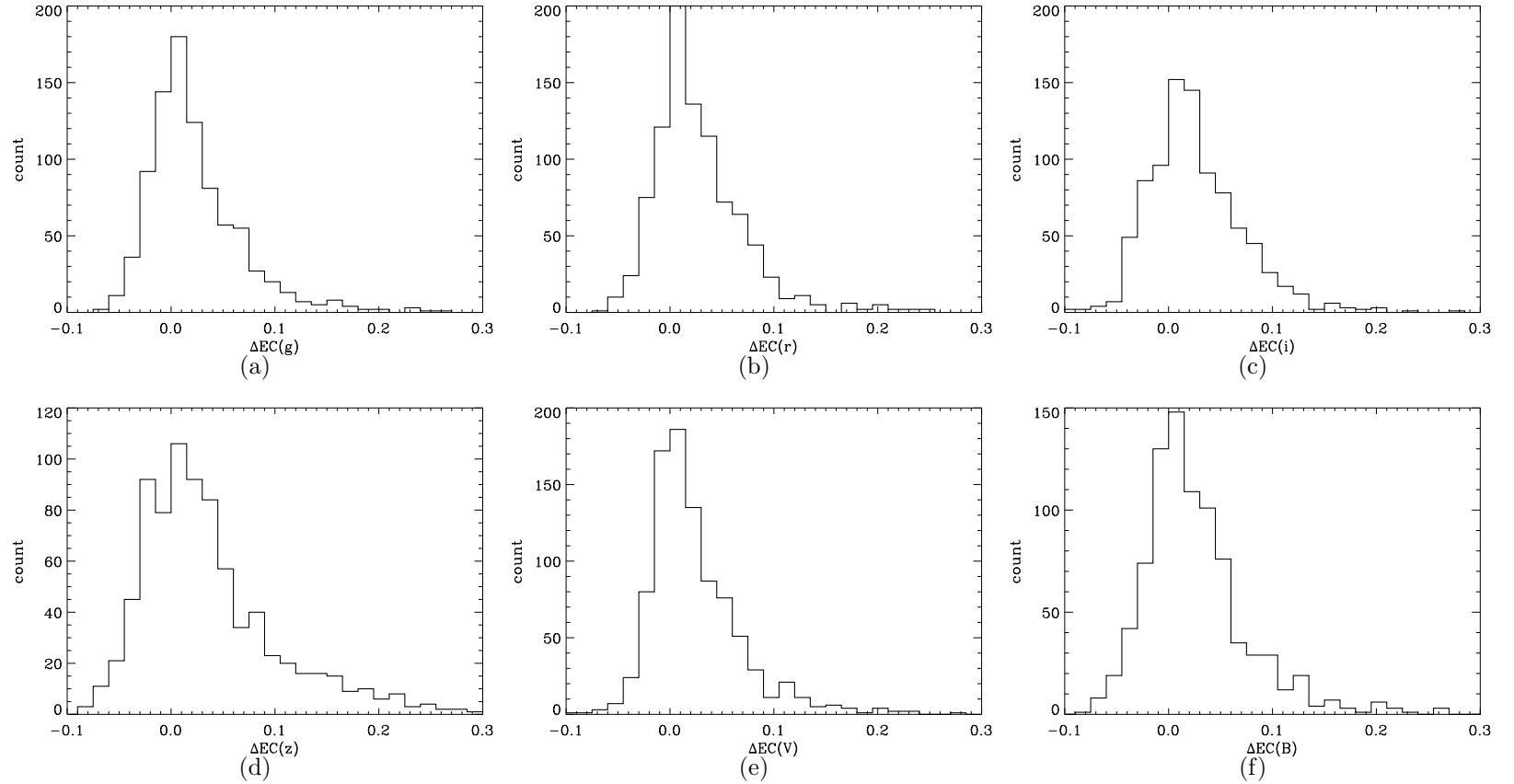


Fig. 4.13 Distributions of emission line correction differences  $\Delta EC = EC_{\text{real}} - EC_{\text{composite}}$  at different wavebands corresponding to Figure 4.12. These figures illustrate that we can in general obtain consistent ECs based on real and composite spectra at different bands. The dispersion of their difference is typically 0.05.

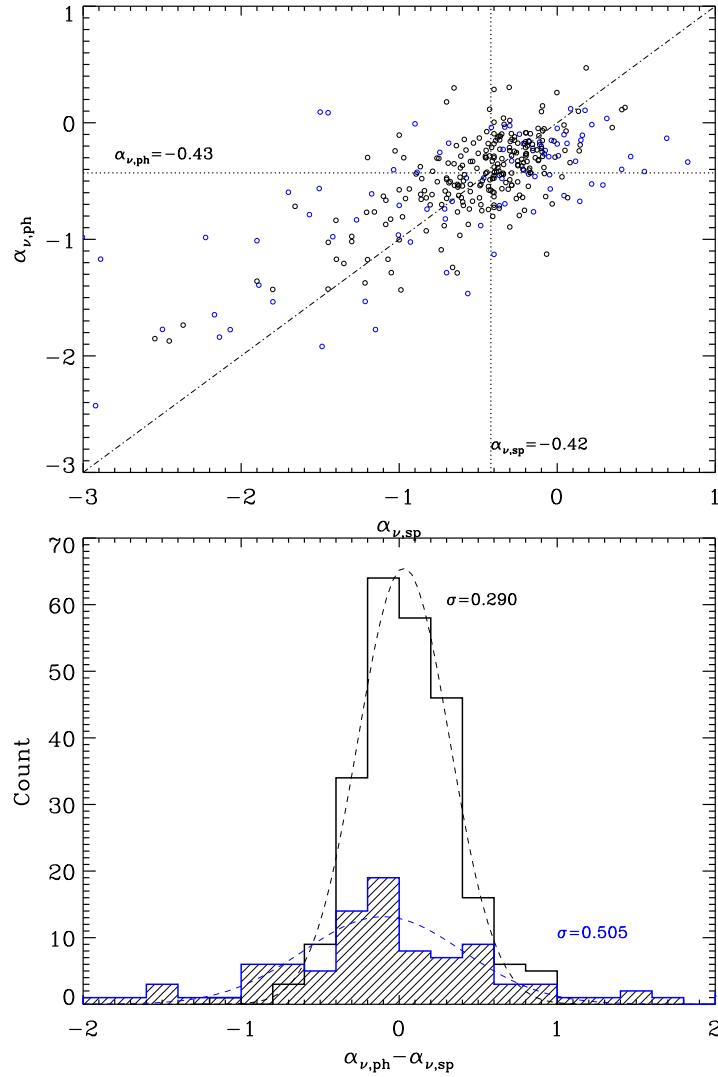


Fig. 4.14 *Upper*: photometric UV slope  $\alpha_{\nu,\text{ph}}$  vs. spectroscopic UV slope  $\alpha_{\nu,\text{sp}}$ . Low redshift objects ( $z < 0.8$ ) are in blue and objects in higher redshift ( $z > 0.8$ ) are in black. The dash dotted line is  $\alpha_{\nu,\text{ph}} = \alpha_{\nu,\text{sp}}$ . The dotted lines represent the median values of  $\alpha_{\nu,\text{ph}} = 0.43$  and  $\alpha_{\nu,\text{sp}} = 0.42$ , respectively. *Lower*: Distribution of differences between  $\alpha_{\nu,\text{ph}}$  and  $\alpha_{\nu,\text{sp}}$  for low redshift ( $z < 0.8$ , in blue) and higher redshift ( $z > 0.8$ , in black) objects. We also fit these two histograms with Gaussian profiles. The low redshift sample has a dispersion ( $\sigma \approx 0.5$ ) much larger than the higher redshift sample dispersion ( $\sigma \approx 0.3$ ).

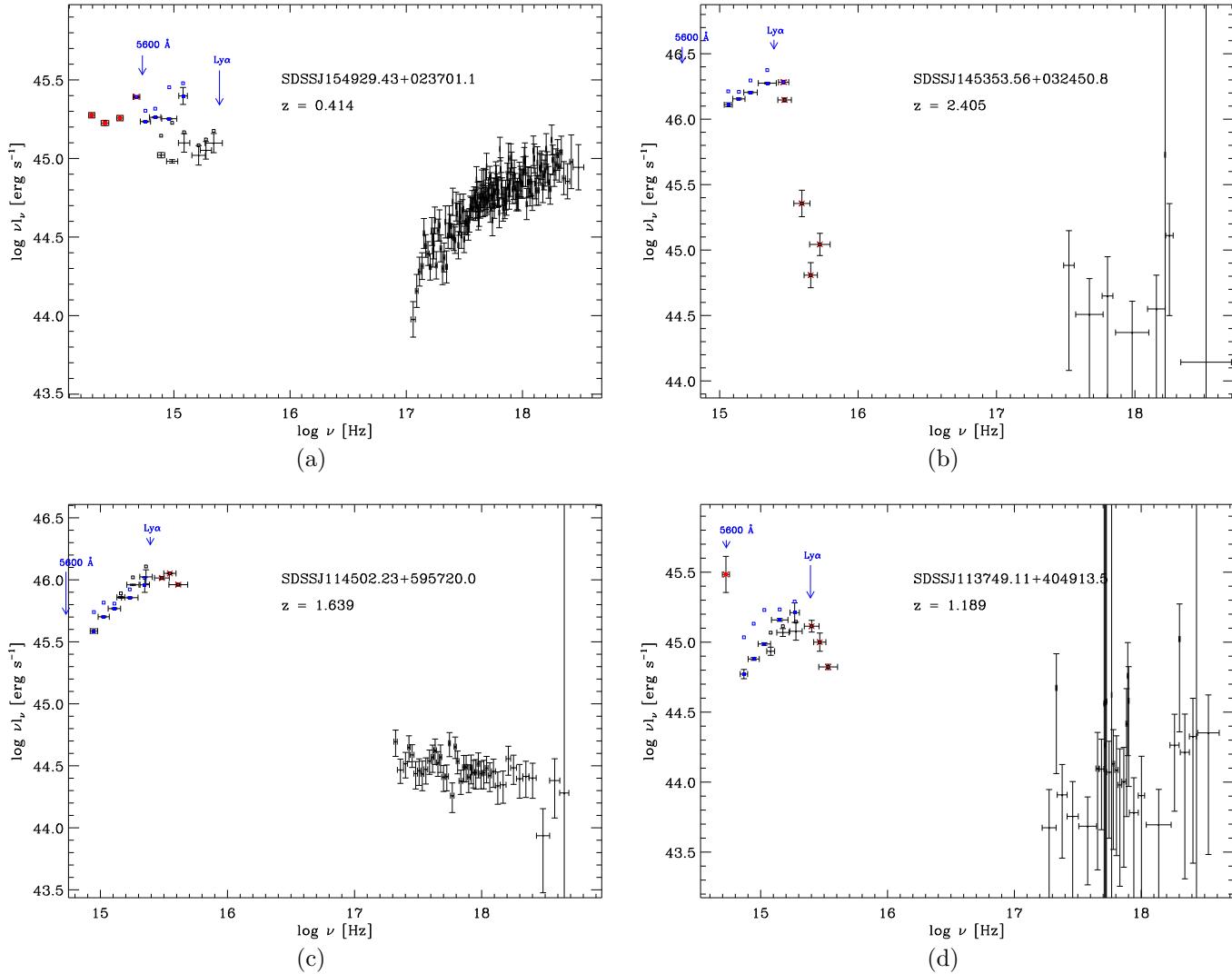


Fig. 4.15 Examples of emission line corrected SEDs. The X-ray data points ( $\log \nu > 17$ ) remain the same. In the UV band  $\log \nu 5600 \text{ \AA} < \log \nu < \log \nu (\text{Ly } \alpha)$ , data points from UVOT are represented by black crosses with error bars; SDSS data points are represented by blue squares and 2MASS data points are represented by red squares. Open squares are photometry before corrections, while filled squares are photometry after correction. Photometric points outside the  $5600 \text{ \AA}$ - $\text{Ly}\alpha$  region are not corrected and flagged by red crosses.

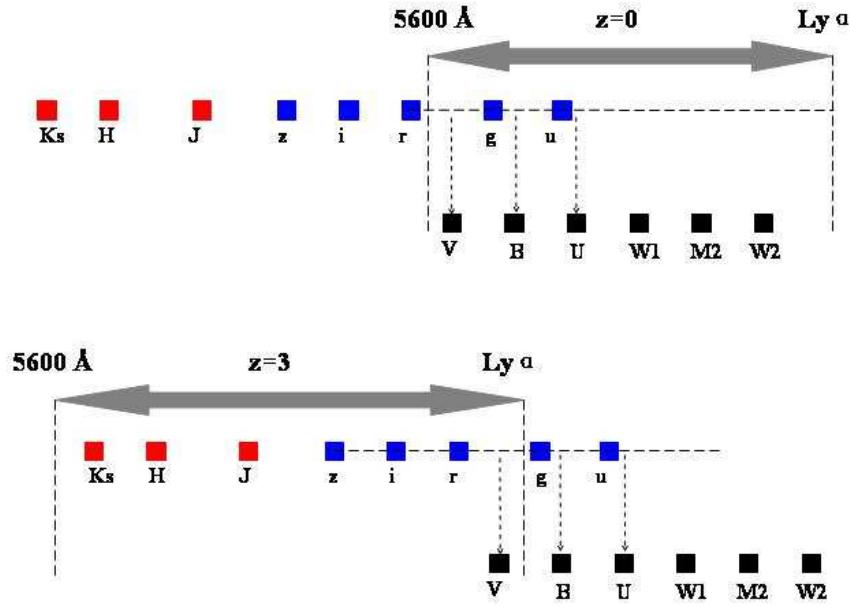


Fig. 4.16 Illustration of photometric shift strategy at  $z = 0$  and  $z = 3$ . In either case, SDSS photometries are interpolated or extrapolated at the wavelengths of UVOT filters. The SDSS and 2MASS data are then shifted by the difference of real and interpolated luminosities.

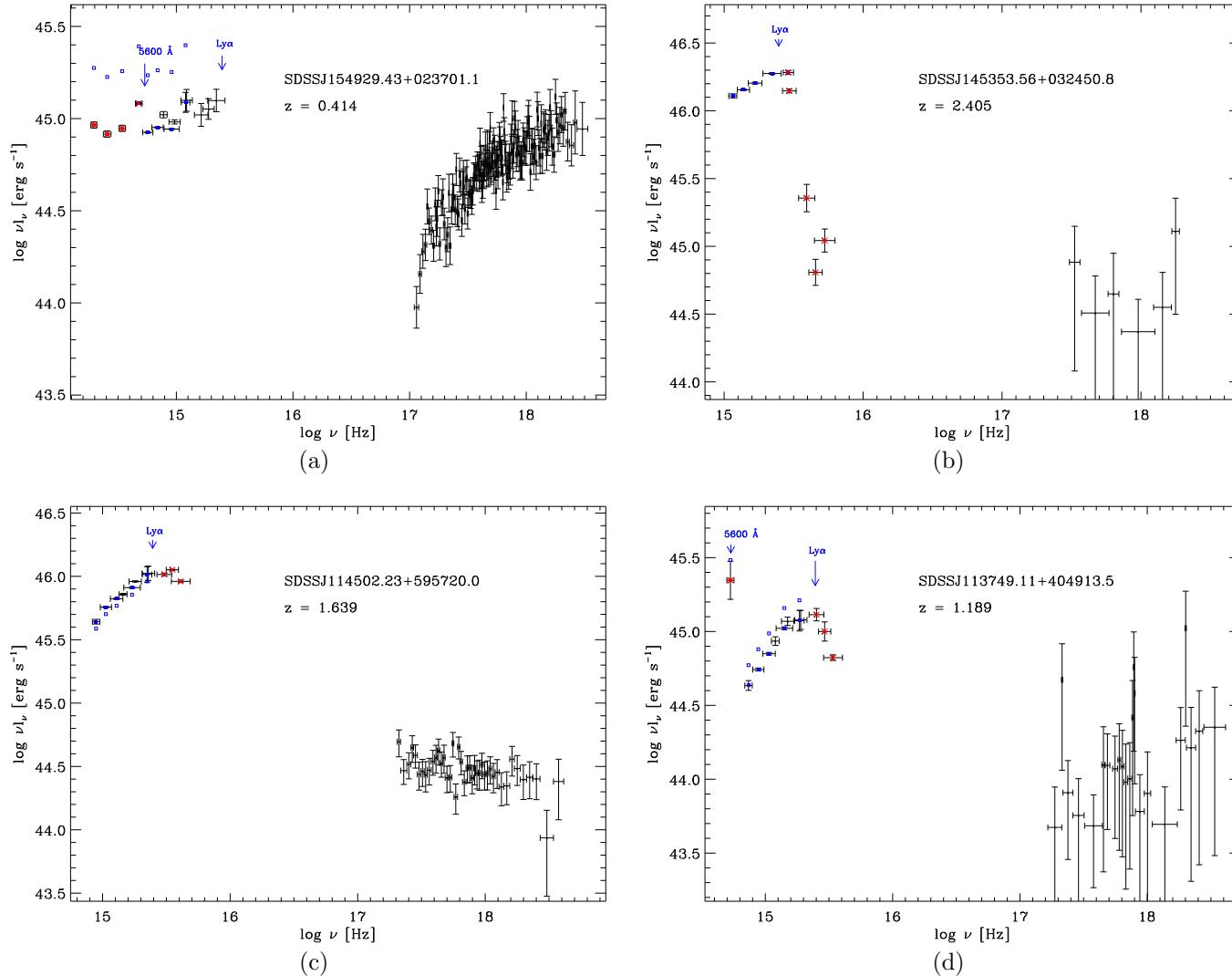


Fig. 4.17 Photometric shifted SEDs for four quasars. The X-ray data points ( $\log \nu > 17$ ) remain the same. In the UV band  $\log \nu 5600 \text{ \AA} < \log \nu < \log \nu(\text{Ly } \alpha)$ , data points from UVOT are represented by black crosses with error bars; SDSS data points are represented by blue squares and 2MASS data points are represented by red squares. Open squares are photometry before photometric shifts, while filled squares are photometry after photometric shifts. Photometric points outside the  $\text{Ly}\alpha$ - $5600 \text{ \AA}$  region are not corrected and flagged by red crosses. SDSSJ145353.56+032450.8 is not corrected because it does not satisfy any of conditions in Table 4.8.

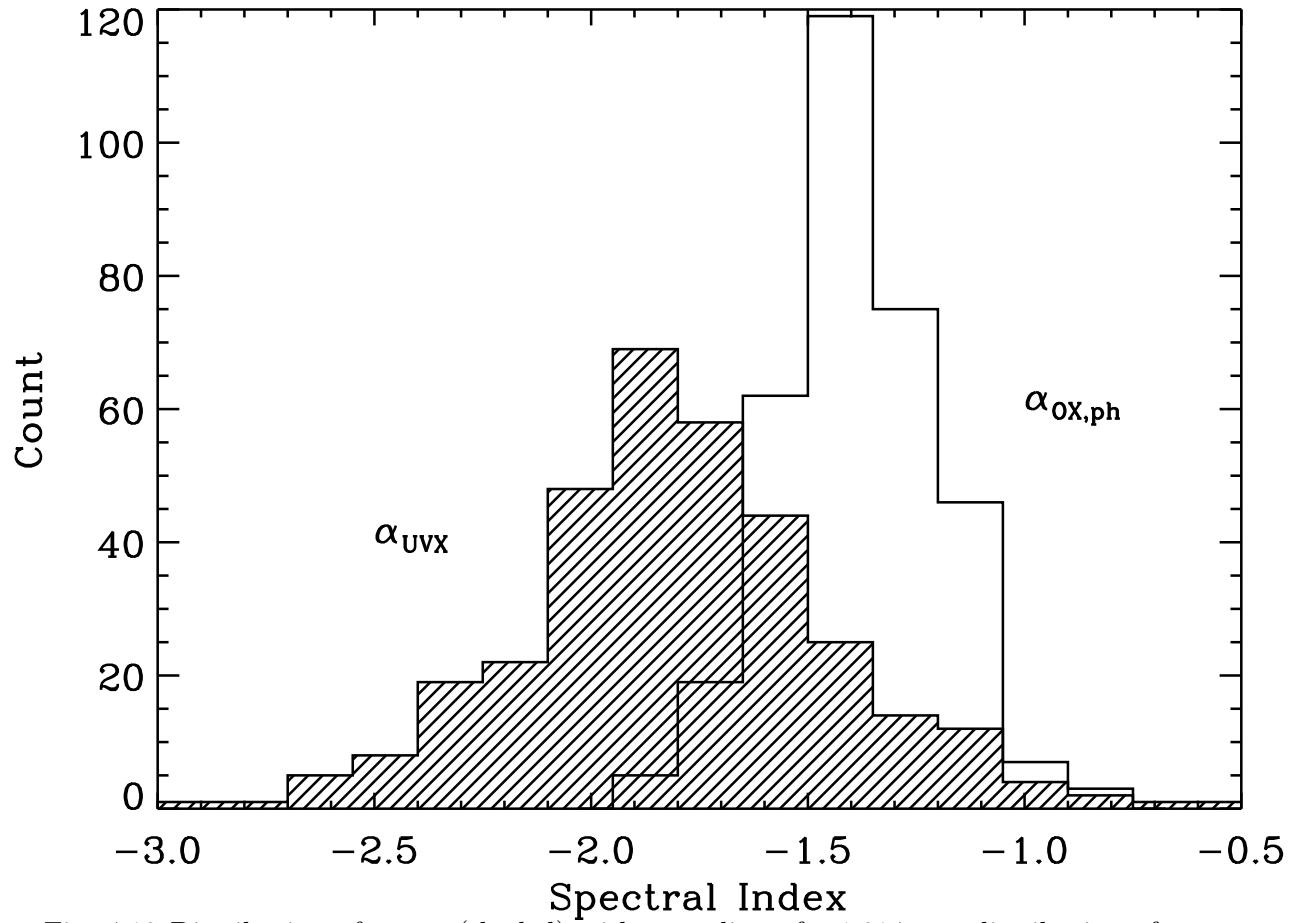


Fig. 4.18 Distribution of  $\alpha_{\text{UVX}}$  (shaded) with a median of  $-1.814$  vs. distribution of  $\alpha_{\text{OX,ph}}$  ( $l_\nu(2500 \text{ \AA})$ ) calculated using photometric data; unshaded region) with a median of  $-1.391$ .

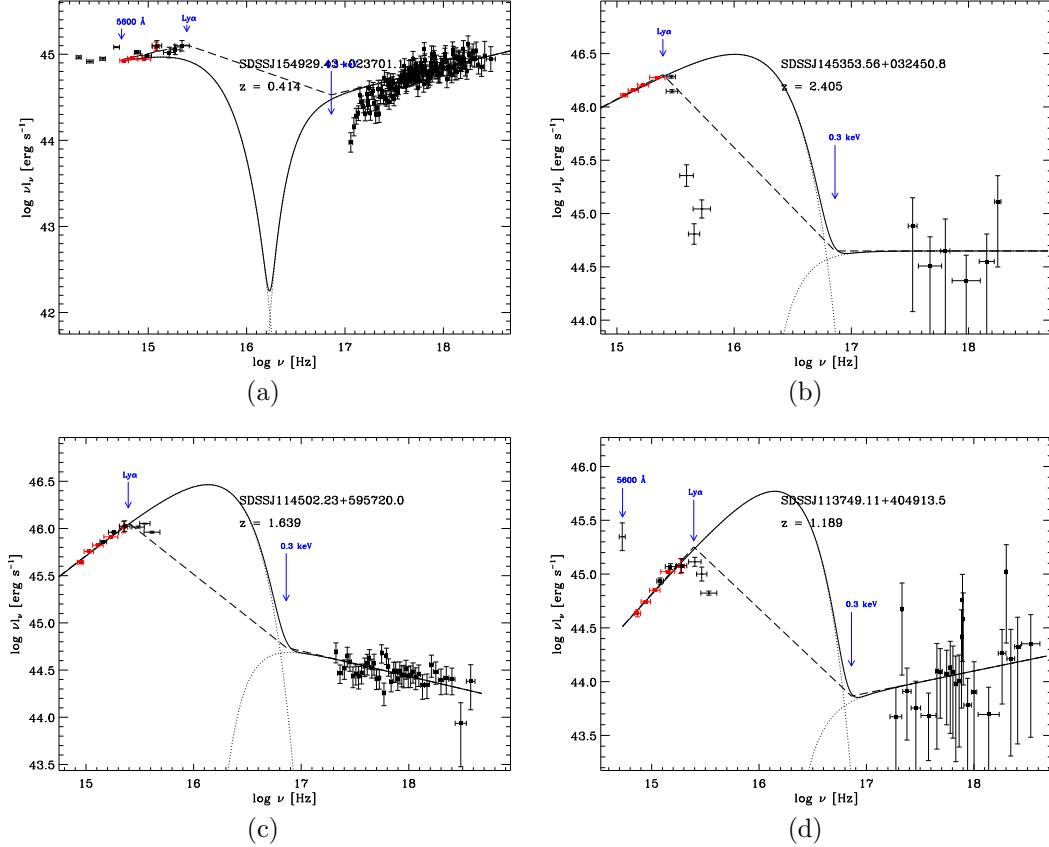


Fig. 4.19 Examples of SEDs fitting in EXP and TPL models. Data points without filled squares are beyond our definition of UV region and thus not used in power-law fitting. *Swift* UVOT data are plotted in black while SDSS data are plotted in red. Black curves are synthesize SED using EXP model and dashed lines represent TPL models. Dotted curves are UV and X-ray components for the EXP model. The SDSSJ154929.43+023701.1 is a case with strong UV and X-ray absorptions, in which the EXP model is no longer an upper limit of BBB emission.

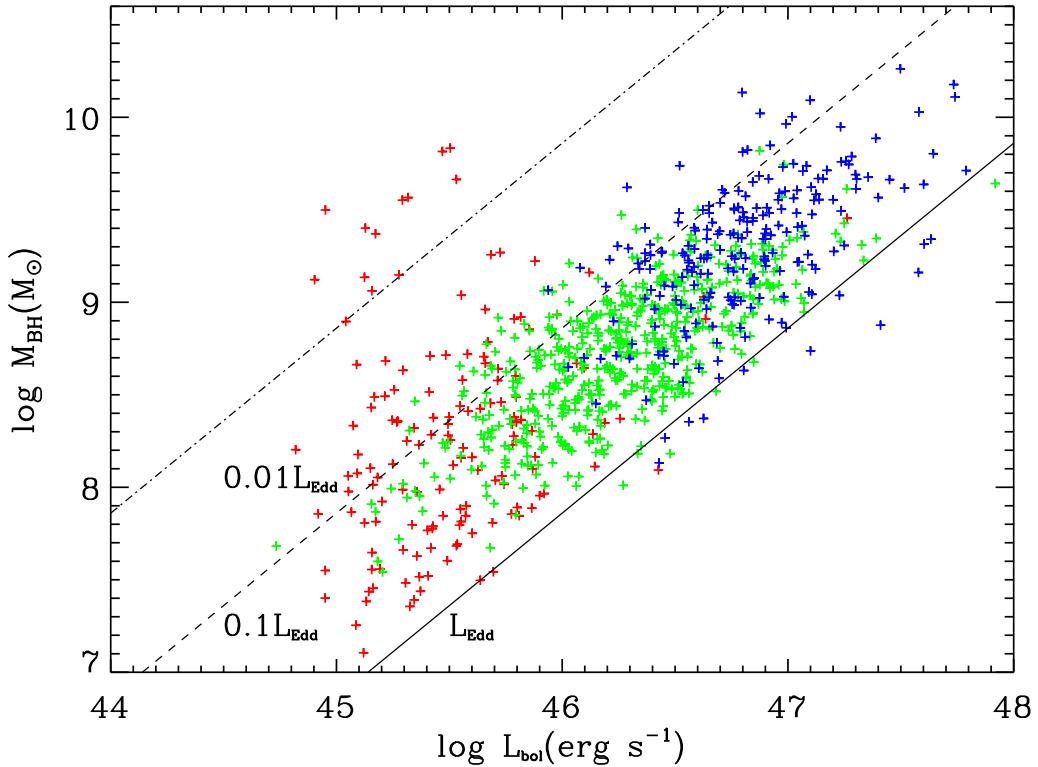


Fig. 4.20 Black hole mass as a function of bolometric luminosity for 923 quasars in our catalog. Quasars are color-coded based on their redshift ranges following the same convention as Shen et al. (2008b): red for  $z < 0.7$ , green for  $0.7 < z < 1.9$ , and blue for  $z > 1.9$ . We also plot solid, dashed and dash-dot lines when the Eddington ratio is 1, 0.1 and 0.01.

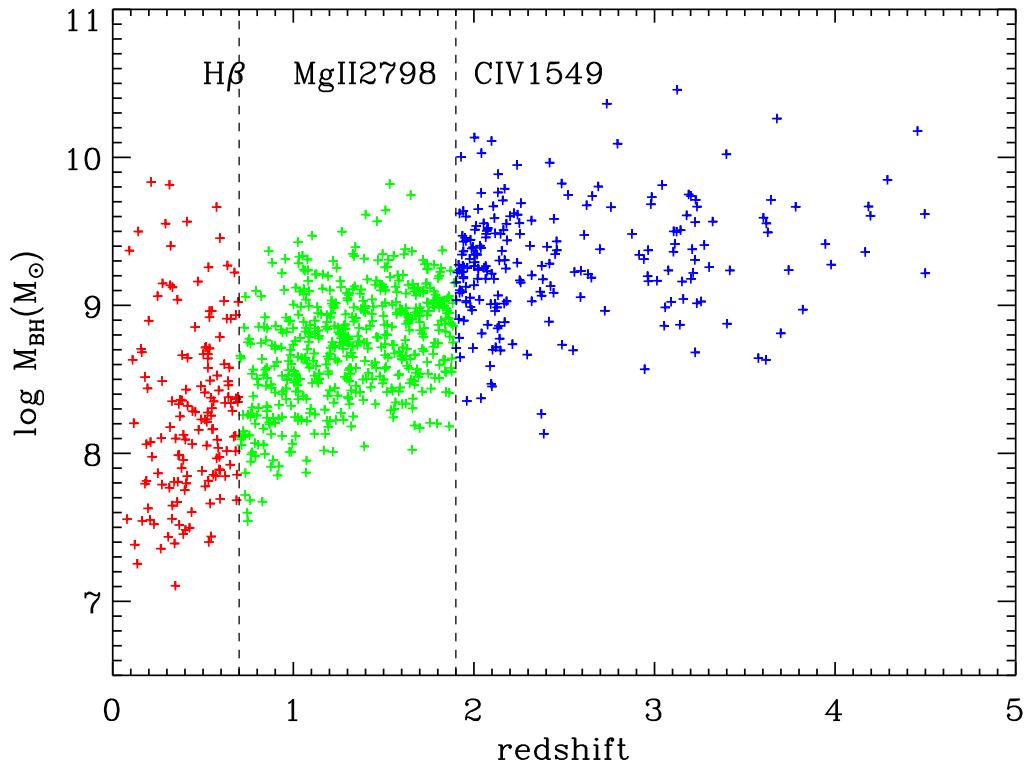


Fig. 4.21 Black hole mass as a function of redshift for 923 objects in our catalog. The black hole masses are measured based on quasar spectra in SDSS DR7 quasar catalog. Quasars are color-coded based on their redshift ranges (and the emission lines and continuum points used for calculation). Low redshift ( $z < 0.7$ ) are plot in red; intermediate redshift ( $0.7 < z < 1.9$ ) are plot in green; high redshift ( $z > 1.9$ ) objects are plot in blue.

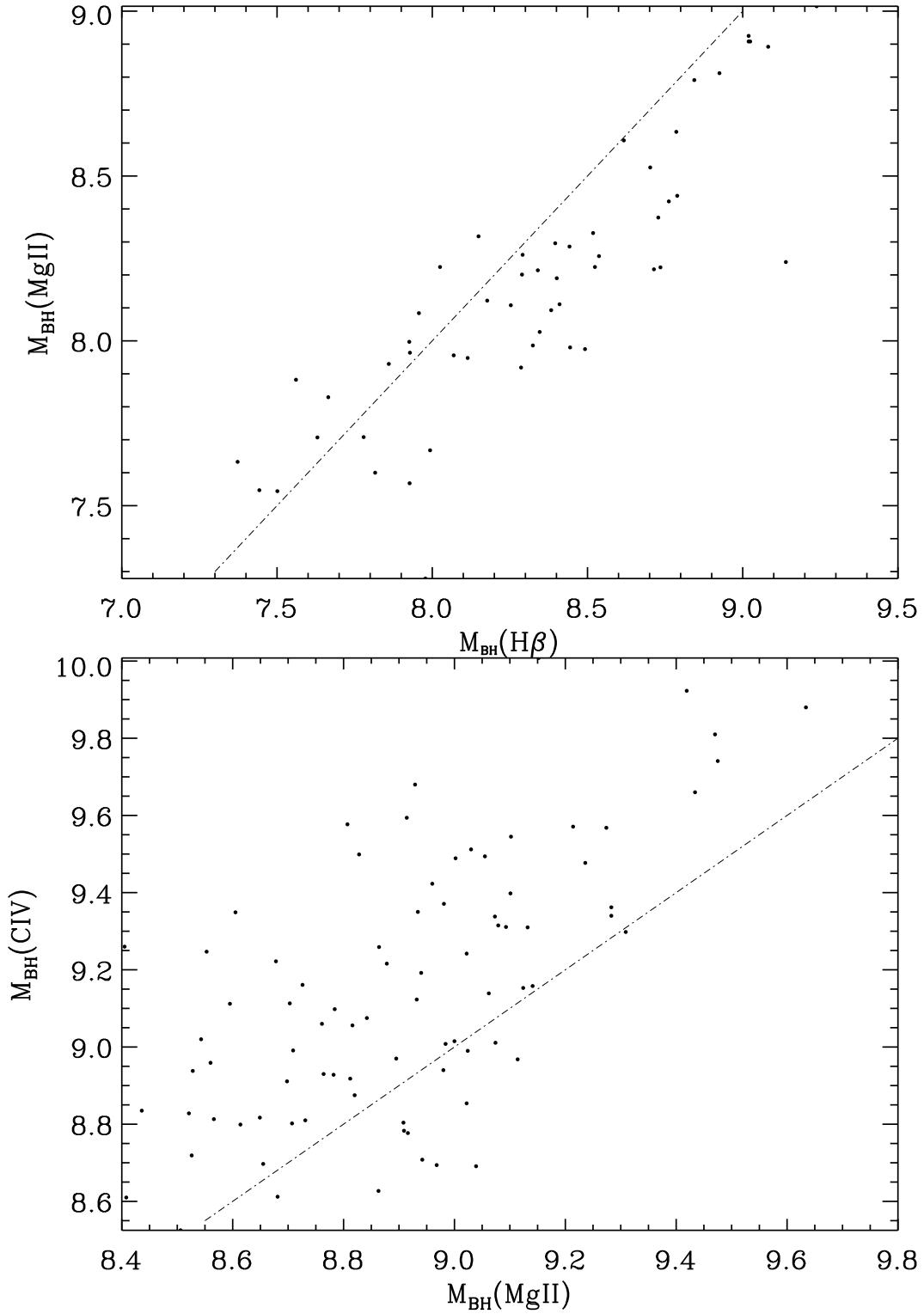


Fig. 4.22 Black hole masses (in logarithmic scale) calculated using different estimators. *Upper:* black hole masses of 53 quasars whose spectra cover both Mg II and H $\beta$ . *Lower:* black hole masses of 83 quasars whose spectra cover both Mg II and C IV. In either panel, the dash-dotted line represents the equality of values from two different estimators.

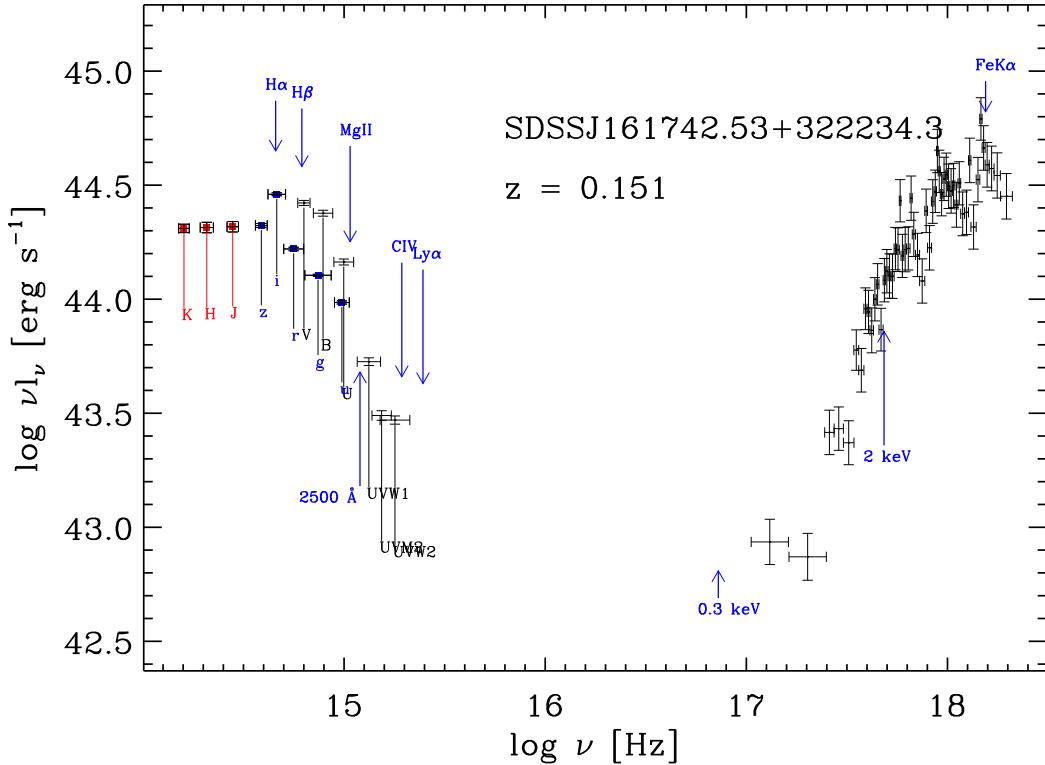


Fig. 4.23 An SED of a reddened quasar (SDSSJ161742.53+322234.3) as observed by *Swift* (black), SDSS (blue), and 2MASS (red). The band passes are listed below the photometric data points, and the positions of important emission lines and continuum wavelength/energy points are marked. The SDSS and 2MASS data points are not corrected for emission lines and not shifted to match the *Swift* flux level. Galactic reddening corrections have been applied to all UV/optical data points. The intrinsic column density  $N_{\text{H},i} \sim 10^{22} \text{ cm}^{-2}$ , which is about two orders of magnitude higher than the Galactic column.

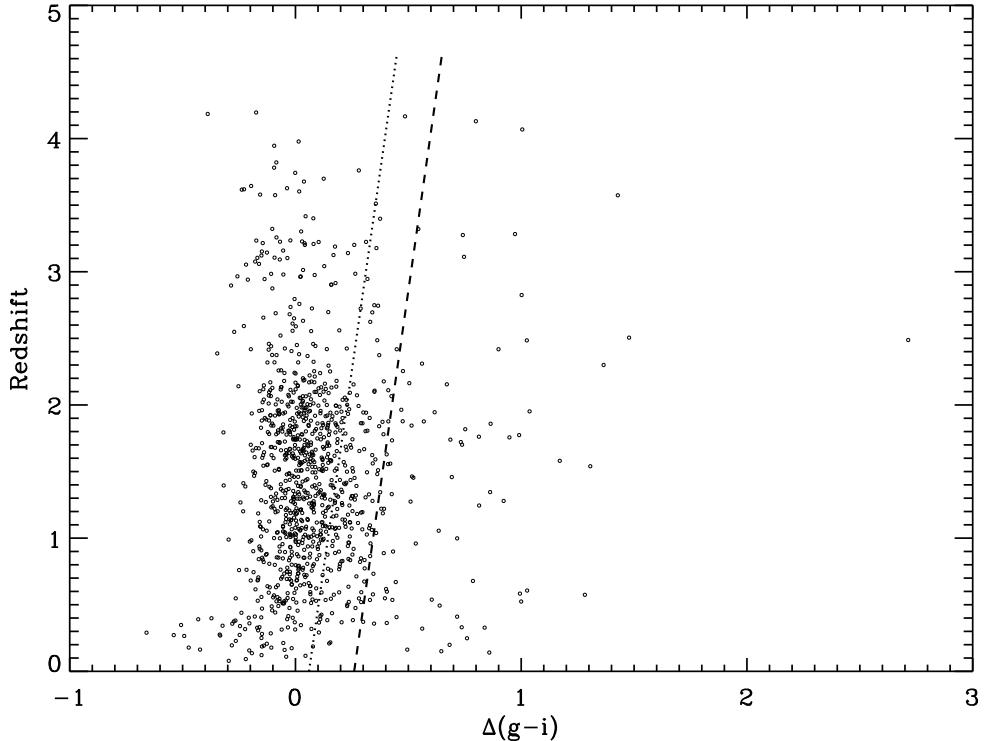


Fig. 4.24 The relative color  $\Delta(g - i)$  vs. redshift diagram of all the objects in the raw catalog. The dotted line shows the effect of SMC-type reddening as a function of redshift with  $E(B - V) = 0.04$  (see Richards et al. (2003) for the choice of 0.04 as the value of  $E(B - V)$ ). The dash line is the dotted line shifted by 0.2 to match the dust-reddening quasar definition of Richards et al. (2003). Quasars to the right of the dashed line can be considered to be dust-reddened.

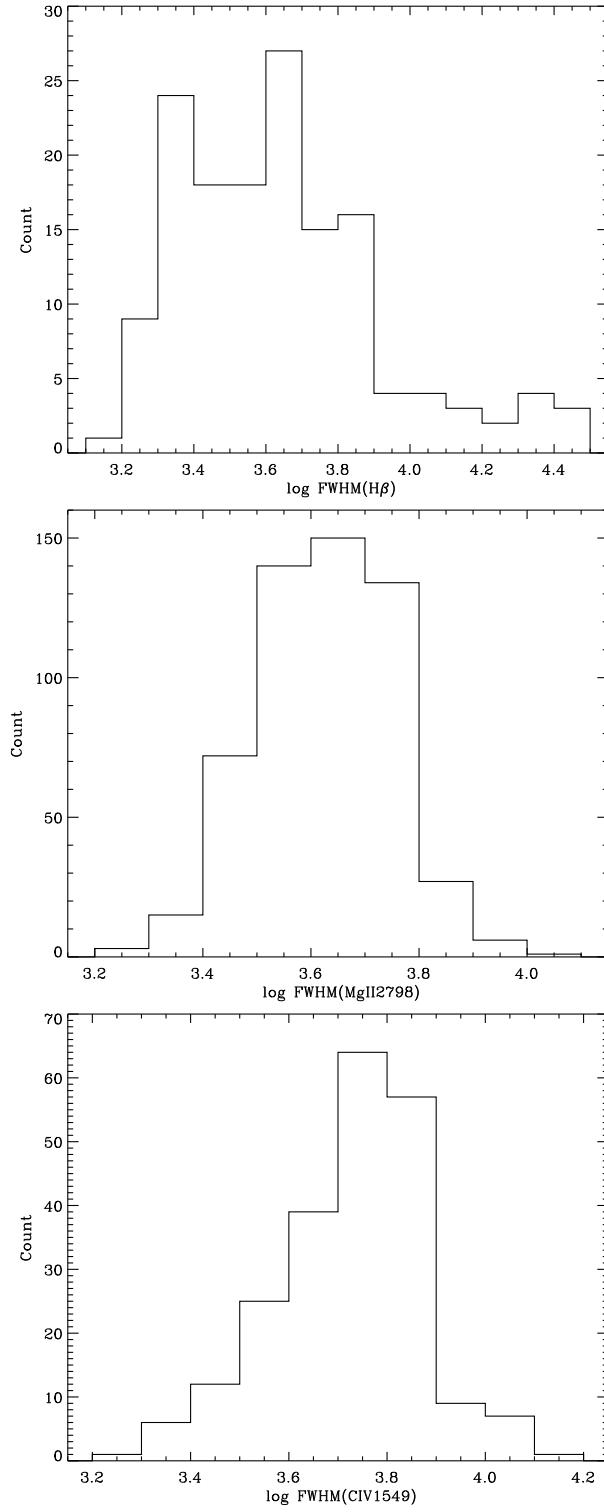


Fig. 4.25 Distributions of FWHM for H $\beta$ , Mg II, and C IV, for all objects in our raw catalog whose SDSS spectra are fit using our software package. The H $\beta$  emission line is fit using three Gaussian profiles; Mg II and C IV are fit using two Gaussian profiles. The histograms indicate that while H $\beta$  and Mg II have similar median values, the peak of C IV distribution is evidently higher, suggesting the bias of using C IV as a black hole mass estimator.

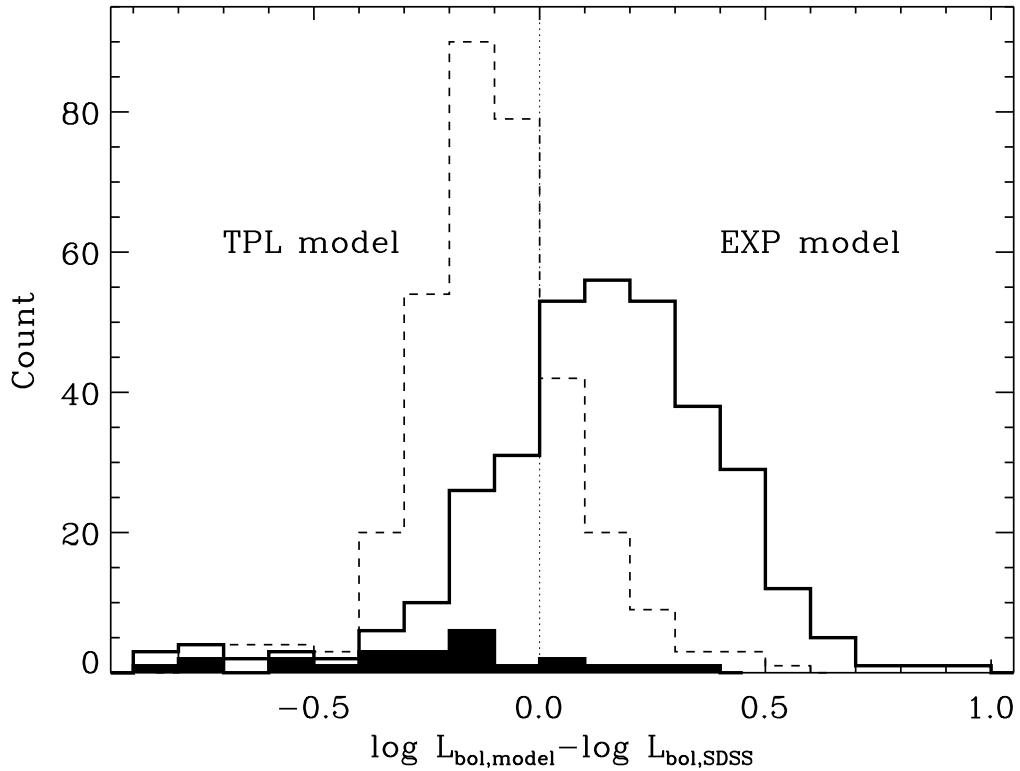


Fig. 4.26 Distributions of  $L_{\text{bol}}$  integrated from the EXP (thick solid line) and TPL (thin dashed line) models with respect to  $L_{\text{bol}}$  calculated using BC correction. The black shaded region under the EXP model histogram represents objects we flagged as “red” which suffer from strong intrinsic absorption (see the “Reddened quasars” in Section 4.5.1).

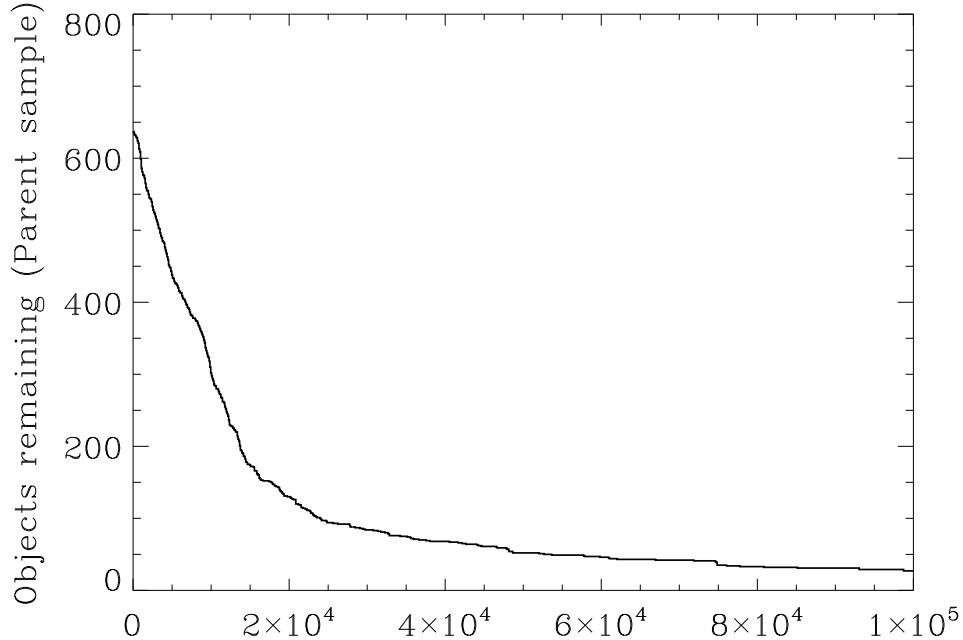


Fig. 4.27 The sample size as a function of XRT exposure cut-off in seconds. This parent sample does not include supplemental sample from Grupe et al. (2010). There are 637 objects if exposure cut-off is zero.

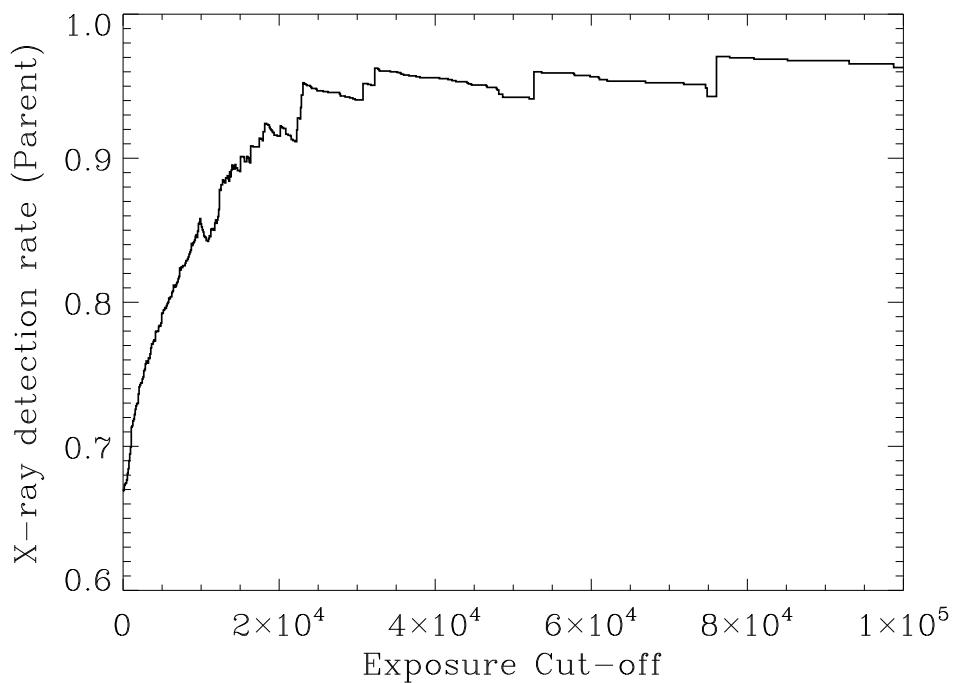


Fig. 4.28 The X-ray detection rate as a function of XRT exposure cut-off in seconds. This parent sample does not include supplemental sample from Grupe et al. (2010).

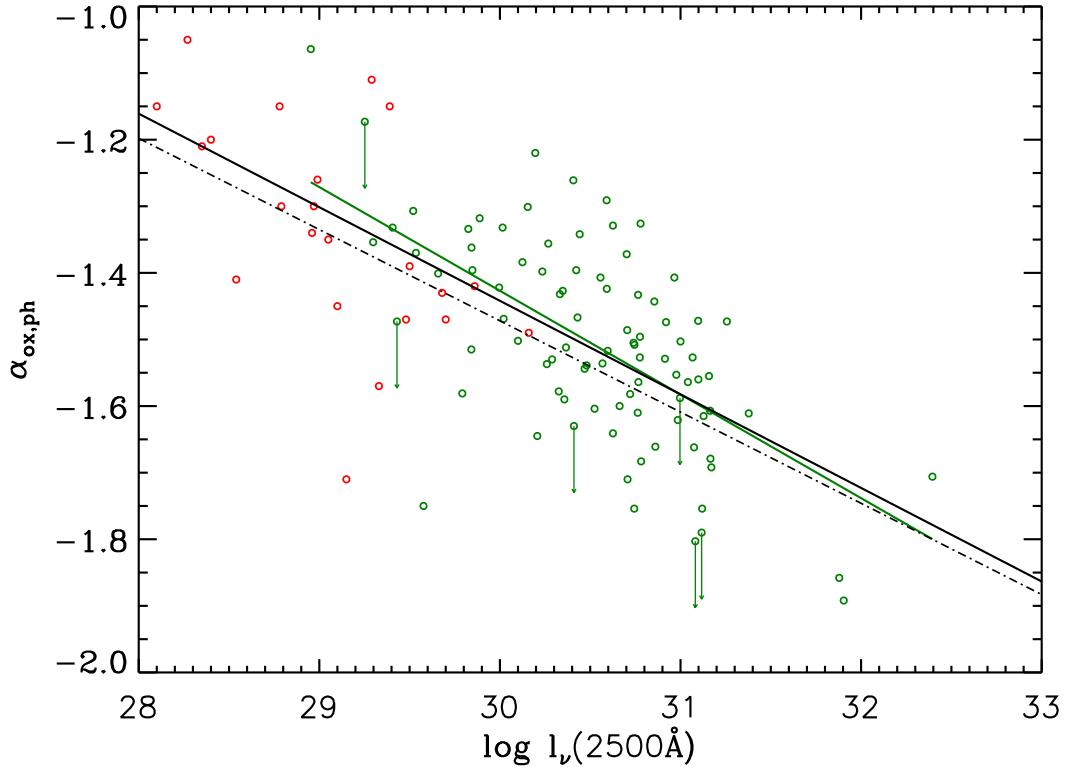


Fig. 4.29 The  $\alpha_{\text{ox}}-l_{\nu}(2500 \text{ \AA})$  relation for the *Swift* AGN sample, including the cleaned catalog sample (green), and the G10 sample (red). The solid black line is the best linear fit to the combined sample using EM method and the solid green line is the best linear fit to the cleaned catalog sample only. The dot dashed line is the best fit of Just et al. (2007).  $\alpha_{\text{OX,ph}}$  means the UV flux is obtained by fitting photometric data points rather than from spectra. The typical error bar is displayed at the upper right corner.

Table 4.1. XRT Data Binning Strategy.

$N_{\text{Xph}}$	Photon# per Bin	Statistics	Flag
$N_{\text{Xph}} \geq 200$	20	$\chi^2$	G(Good)
$100 \leq N_{\text{Xph}} < 200$	$N_{\text{Xph}}/10$	$\chi^2$	G(Good)
$10 \leq N_{\text{Xph}} < 100$	$N_{\text{Xph}}/10$	Cash	A(Acceptable)
$N_{\text{Xph}} < 10$	No Binning	Upper limit <sup>1</sup>	W(Weak)/O(Out of FOV)

<sup>1</sup>If the total number of X-ray photons is less than 10, we will not fit the X-ray spectrum but only calculate the flux limit.

Table 4.2. Models used to fit XRT spectra.

Model	$\Gamma$	$N_{\text{H,G}}$	$N_{\text{H,i}}$	#Object
A	Fixed = 2	Fixed	Fixed = 0	177(44%)
B	Free	Fixed	Fixed = 0	195(49%)
C	Free	Fixed	Free	15(4%)
D	Fixed = 2	Fixed	Free	12(3%)

Table 4.3.  $F_{\text{PSF,Xspec}}^0/\text{CR}$  vs. offset angle  $\theta$ .

$\theta [']$	$F_{\text{PSF,Xspec}}^0/\text{CR}$
1	$4.739 \times 10^{-11}$
3	$5.082 \times 10^{-11}$
5	$5.050 \times 10^{-11}$
7	$4.926 \times 10^{-11}$
9	$5.874 \times 10^{-11}$
11	$5.364 \times 10^{-11}$
13	$4.256 \times 10^{-11}$
21	$5.229 \times 10^{-11}$

Note. —  $F_{\text{PSF,Xspec}}^0$  is the real flux received with absorption. The count rate is measured within  $30''$  centered at the target object.

Table 4.4. UVOT Sky Image Flagging Strategy.

Flag	Description	Source Region	Background Region
-2	Bad Aspect	None	None
-1	Out of Image	None	None
1	Faint	Default	Default
0	$\delta \leq 0''.618$	Default <sup>a</sup>	Default <sup>a</sup>
2	$0''.618 < \delta \leq 3''.0$	Customized <sup>b</sup>	Default <sup>a</sup>
3	$\delta > 3''.0$	None	None
4	Close to the edge ( $d_E \leq 27''.5$ )	None	None
5	Out of FOV	None	None
10+	$27''.5 < d_E \leq 35''.0$	TBD <sup>c</sup>	TBD <sup>c</sup>
100+	$d_V \leq 100''$	TBD <sup>c</sup>	TBD <sup>c</sup>

<sup>a</sup>See Section 4.2.4 for a description of the default source and background region files.

<sup>b</sup>Source region circles are centered at the new Gaussian centroid.

<sup>c</sup>These images are passed to the visual inspection process (see Section 4.2.4)

Table 4.5. List of reasons to exclude unacceptable UVOT sky images.

#	Description
122	Diffuse emission from a nearby comet.
52	Unrecorded aspect problem
1556	Close to image edge. Subject to vignetting.
1	Contamination by emission from extended structure.
54	Contamination from spiral arms.
77	Contamination from galaxy halo.
405	Point sources become trails.
107	Contamination from nearby unseparatable sources.
140	Standard regions do not work.
318	Charge trails passing by source region.
351	Other
3183	Total

Table 4.6. Four types of objects in the raw catalog.

Type	UVOT	XRT	In Catalog	SED	Number
A	Y	Y	Y	Y	637
B	Y	N	Y	N	38
C	N	Y	Y	N	168
D	N	N	N	N	191
Total	675	805	843	637	1034

Table 4.7. Central wavelengths and FWHMs of UV/optical wavebands.

Band	Telescope /Instrument	$\lambda_0$ (Å)	FWHM (Å)
V	<i>Swift</i> /UVOT	5468	769
B	<i>Swift</i> /UVOT	4392	975
U	<i>Swift</i> /UVOT	3465	785
UVW1	<i>Swift</i> /UVOT	2600	693
UVM2	<i>Swift</i> /UVOT	2246	498
UVW2	<i>Swift</i> /UVOT	1928	657
<i>u</i>	SDSS	3541	560
<i>g</i>	SDSS	4653	1377
<i>r</i>	SDSS	6147	1377
<i>i</i>	SDSS	7461	1510
<i>z</i>	SDSS	8904	940
<i>J</i>	2MASS	12350	1620
<i>H</i>	2MASS	16620	2510
<i>Ks</i>	2MASS	21590	2620

Table 4.8. Photometric shift strategies.

Priority <sup>1</sup>	Condition UV Bands	Condition EUV Bands	Method <sup>2</sup> Method <sup>2</sup>
1	$g+u+U$		E
2		$g+u+U$	E
3	$g+B+u$		I
4		$g+B+u$	I
5	$r+g+B$		E
6	$r+V+g$		I
7		$r+V+g$	I
8	$V+u+g$		E
9	$i+r+V$	$i+r+V$	E
10	$g+u+UVW1$		E
11		$g+u+UVW1$	E
12	$g+u+UVM2$		E
13	$g+u+UVW2$		E
14		$g+u+UVW2$	E
15		$g+u+UVM2$	E

<sup>1</sup>The highest priority is 1.<sup>2</sup>E for extrapolation; I for interpolation

Table 4.9. Parameters for black hole mass calculation.

Redshift	FWHM	$\lambda_c^a$	$a$	$b$	BC	Reference
$z < 0.7$	FWHM(H $\beta$ )	5100 Å	0.660	0.53	9.26	McLure & Dunlop (2004)
$0.7 < z < 1.9$	FWHM(Mg II)	3000 Å	0.505	0.62	5.15	McLure & Jarvis (2002); McLure & Dunlop (2004)
$z > 1.9$	FWHM(C IV)	1350 Å	0.672	0.61	3.81	Vestergaard & Peterson (2006)

<sup>a</sup>Wavelength of continuum monochromatic luminosity.

Table 4.10. Catalog description.

Column	Format	Symbol	Description
1	A18	SDSSID	SDSS DR5 Designation hhmmss.ss+ddmmss.s(J2000)
2	F11.6	RA	SDSS right ascension in decimal degrees (J2000)
3	F11.6	DEC	SDSS declination in decimal degrees (J2000)
4	F7.4	$z$	Redshift from SDSS DR5 quasar catalog
5	F8.3	$M_i$	Absolute magnitude at $i$ band from SDSS DR5 quasar catalog
6	F7.3	$u$	BEST PSF $u$ magnitude (not corrected for Galactic extinction)
7	F6.3	$\Gamma$	Photon index between 0.3 and 10 keV
8	F6.3	$\delta^-(\Gamma)$	$1\sigma$ lower error bar of photon index
9	F6.3	$\delta^+(\Gamma)$	$1\sigma$ upper error bar of photon index
10	F6.3	HR	Hardness ratio
11	F7.3	$\log F_{\text{obs}}(0.3\text{--}10 \text{ keV})$	Observed flux between 0.3 and 10 keV
12	F7.3	$\log F_{\text{unobs}}(0.3\text{--}10 \text{ keV})$	Unobserved flux between 0.3 and 10 keV
13	F7.3	$\log l_{\nu}(2 \text{ keV})$	Monochromatic luminosity at 2 keV
14	F6.3	$\delta \log l_{\nu}(2 \text{ keV})$	$1\sigma$ uncertainty of monochromatic luminosity at 2 keV
15	F7.3	$\log L(0.3\text{--}10 \text{ keV})$	Integrated luminosity between 0.3 and 10 keV
16	F6.3	CR	Source count rate in $10^{-3} \text{ cnt s}^{-1}$
17	F8.1	$T_{\text{XRT}}$	Total XRT exposure time
18	F6.3	$N_{\text{H,G}}$	Galactic column density in $10^{20} \text{ cm}^{-2}$
19	F6.3	$N_{\text{H,i}}$	Intrinsic column density in $10^{22} \text{ cm}^{-2}$
20	F6.3	$\delta_-(N_{\text{H,i}})$	$1\sigma$ lower error bar in $10^{22} \text{ cm}^{-2}$ of $N_{\text{H,i}}$
21	F6.3	$\delta_+(N_{\text{H,i}})$	$1\sigma$ upper error bar in $10^{22} \text{ cm}^{-2}$ of $N_{\text{H,i}}$
22	F6.3	$f(V)$	Flux density at <i>Swift</i> V band <sup>1</sup>
23	F6.3	$\delta f(V)$	$1\sigma$ error bar of flux density at <i>Swift</i> V band
24	F8.1	$T_V$	Total exposure time (seconds) in V band
25	F6.3	$f(B)$	Flux density at <i>Swift</i> B band <sup>1</sup>
26	F6.3	$\delta f(B)$	$1\sigma$ error bar of flux density at <i>Swift</i> B band
27	F8.1	$T_B$	Total exposure time (seconds) in B band
28	F6.3	$f(U)$	Flux density at <i>Swift</i> U band <sup>1</sup>
29	F6.3	$\delta f(U)$	$1\sigma$ error bar of flux density at <i>Swift</i> U band
30	F8.1	$T_U$	Total exposure time (seconds) in U band
31	F6.3	$f(\text{UVM1})$	Flux density at <i>Swift</i> UVM1 band <sup>1</sup>
32	F6.3	$\delta f(\text{UVM1})$	$1\sigma$ error bar of flux density at <i>Swift</i> UVM1 band
33	F8.1	$T_{\text{UVM1}}$	Total exposure time (seconds) in UVM1 band
34	F6.3	$f(\text{UVW2})$	Flux density at <i>Swift</i> UVW2 band <sup>1</sup>
35	F6.3	$\delta f(\text{UVW2})$	$1\sigma$ error bar of flux density at <i>Swift</i> UVW2 band
36	F8.1	$T_{\text{UVW2}}$	Total exposure time (seconds) in UVW2 band
37	F6.3	$f(\text{UVM2})$	Flux density at <i>Swift</i> UVM2 band <sup>1</sup>
38	F6.3	$\delta f(\text{UVM2})$	$1\sigma$ error bar of flux density at <i>Swift</i> UVM2 band
39	F8.1	$T_{\text{UVM2}}$	Total exposure time (seconds) in UVM2 band
40	F6.3	$\alpha_{\text{UV,ph}}$	UV spectral index by fitting photometric data
41	F6.3	$\delta \alpha_{\text{UV,ph}}$	$1\sigma$ error bar of $\alpha_{\text{UV,ph}}$
42	F6.3	$\log l_{\nu,\text{ph}}(2500 \text{ \AA})$	$\log l_{\nu}(2500 \text{ \AA})$ by fitting UV photoemtric data
43	F6.3	$\delta \log l_{\nu,\text{ph}}(2500 \text{ \AA})$	$1\sigma$ error bar of $\log l_{\nu,\text{ph}}(2500 \text{ \AA})$
44	F6.3	$\alpha_{\text{OX,ph}}$	$\alpha_{\text{ox}}$ calculated using $\log l_{\nu,\text{ph}}(2500 \text{ \AA})$ and $\log l_{\nu}(2 \text{ keV})$
45	F6.3	$\delta \alpha_{\text{OX,ph}}$	$1\sigma$ error bar of $\alpha_{\text{OX,ph}}$
46	F6.3	$M_{\text{BH}}$	Black hole mass <sup>2</sup>
47	F6.3	$\log L_{\text{bol,exp}}$	Bolometric luminosity by exponential decay model
48	F6.3	$\log L_{\text{bol,tpl}}$	Bolometric luminosity by triple power-law model

Note. — Data entry is set to  $-9.999$  if unavailable.

<sup>1</sup>Fluxes are in  $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ , and are corrected for Galactic reddening.

<sup>2</sup>Black hole mass calculated using emission line FWHM and corresponding continuum flux calculated by power-law fitting of photometric data points.

Table 4.11. Classification for the entire SED sample.

Classification	Counts
BAL quasar	12
BL Lac	1
Lensed quasar	1
NLS1	7
Extended	1
RL	80
Sy 1.8	1
Dust reddened	17
RQ	225
Total	345

Table 4.12. Correlation and regression analysis.

Sample	Size	$\rho_s(P_0)^3$	EM			BJ		
			Slope	Intercept	Dispersion <sup>1</sup>	Slope	Intercept	Dispersion <sup>2</sup>
Cleaned catalog	89	-0.553	$-0.1557 \pm 0.0221$	$3.2441 \pm 0.6762$	0.1219	$-0.1568 \pm 0.0217$	3.2756	0.1151
Combined	111	-0.678	$-0.1405 \pm 0.0139$	$2.7731 \pm 0.4203$	0.1126	$-0.1424 \pm 0.0134$	2.8298	0.1072
J07	372	-0.762	$-0.1396 \pm 0.0070$	$2.7045 \pm 0.2119$	0.1504	$-0.1404 \pm 0.0063$	2.7230	0.1319
J07T <sup>4</sup>	289	-0.662	$-0.1341 \pm 0.0103$	$2.5406 \pm 0.3102$	0.1542	$-0.1367 \pm 0.0093$	2.6122	0.1320

Note. — EM: Expectation-Maximization algorithm; BJ: Buckley-James algorithm.

<sup>1</sup>Standard normal residual. See Lavalley et al. (1992).

<sup>2</sup>Kaplan-Meier residual. See Lavalley et al. (1992).

<sup>3</sup>Spearman rank correlation coefficient  $\rho_s$  with confidence level  $P_0$ .

<sup>4</sup>J07 sample after excluding high luminosity quasars with  $\log l_\nu(2500 \text{ \AA}) > 31.5$ .

## Chapter 5

# The Physical Conditions of the Intrinsic N<sub>V</sub> Narrow Absorption Line Systems of Three Quasars

### 5.1 Introduction

*Intrinsic* absorption lines seen in the spectra of quasars are thought to originate in ionized gas physically related with the quasar “central engine” (e.g., Hamann et al. 2004). They include resonance lines that appear in the rest-frame UV band, such as N<sub>V</sub>, C<sub>IV</sub> and/or high ionization species that have absorption lines and edges in the soft X-ray band, such as O<sub>VII</sub> and O<sub>VIII</sub> (e.g., Reynolds 1997; Giustini et al. 2010). The intrinsic UV absorption lines are usually blueshifted relative to the quasar emission lines and they are believed to arise from gas that is outflowing from the immediate vicinity of the central engine. Lines that are blueshifted by less than  $5000 \text{ km s}^{-1}$  are conventionally regarded as *associated* absorption lines (AALs). Many of these may indeed be connected to broad emission-line region (BELR) (Ganguly et al. 2001) and to X-ray “warm” absorbers (Mathur et al. 1995; Brandt et al. 2000). These absorption lines thus trace the physical conditions and kinematics surrounding the central engine over a wide range of redshifts, and thus provide a diagnostic of quasar evolution.

Narrow intrinsic absorption lines (NALs) typically have widths  $\lesssim 500 \text{ km s}^{-1}$  so that their doublets can be separated at high spectral resolution (see Hamann & Sabra 2004). This is in contrast to broad absorption lines (BALs) whose width can reach  $30,000 \text{ km s}^{-1}$  (Trump et al. 2006). Intrinsic NALs studies are complementary to studies of BALs, since they probe different regions in the quasar surroundings. Although the intrinsic NALs are more difficult to identify, several practical considerations make the derivation of their physical conditions more straightforward than for BALs. (Hamann & Ferland 1999). Because NALs are often unsaturated and resolved, we can measure directly the NAL coverage fractions and column densities of various ions. In BAL systems, doublet transitions are often self-blended, making such measurement more difficult.

Approximately 60% of all quasars show evidence of outflows (Misawa et al. 2007; Ganguly & Brotherton 2008). Although statistical methods can be used to determine this, partial coverage of doublets and multiplets and time variability analysis are the two most commonly used ways to determine decisively if a particular NAL system is intrinsic. These signatures are only seen in low ionization transitions of intervening absorbers, and only rarely in the case of small molecular clouds that could be smaller than the projected size of the background quasar broad emission line region. (e.g., Jones et al. 2010, Ivanchik et al. 2010). Partial coverage is described by the coverage fraction  $C_f$ , which is the fraction of photons from the background source that pass through the absorber (Barlow et al. 1997). It can be estimated using the residual flux ratio of resonance doublets (e.g., Barlow & Sargent 1997; Ganguly et al. 1999). Variability of

the absorption lines could be caused by transverse motion of the absorbing material or by changes in its ionization state (Hamann 1997; Misawa et al. 2005). Using UV spectra of  $z \leq 1.5$  quasars observed at two different epochs separated by 4–10 years with the *Hubble Space Telescope*, Wise et al. (2004) concluded that a minimum of 21% of the AALs are variable. A similar conclusion was reached for  $z \sim 2$  quasars by Narayanan et al. (2004).

Early work investigating metal abundances using narrow associated absorption lines indicated supersolar metallicities, *i.e.*,  $Z \geq Z_{\odot}$  (Petitjean et al. 1994; Tripp et al. 1996; Hamann 1997). More recent work measures metallicities consistent with the previously determined values, and in some cases, it has provided a more stringent constraint. Papovich et al. (2000) analyzed a  $z = 1.207$  AAL system towards Q0122+0388 with the help of photoionization models and found  $Z \sim 2Z_{\odot}$ . Fields et al. (2005) found that the metallicity must be at least  $5Z_{\odot}$  in order to produce the high-ionization absorption lines (O VI, C IV, and N V) of the narrow-line Seyfert galaxy Mrk 1044. Gabel et al. (2006) also found supersolar abundances by modelling the AALs of the quasar J2233–606, with  $[C/H]$ ,  $[O/H] \approx 0.5\text{--}0.9$ , and  $[N/H] \approx 1.1\text{--}1.3$ . High metallicity, up to  $3Z_{\odot}$ , is suggested in the study of quasar PG 0935+417 (Hamann et al. 2004). A study of the outflow from Mrk 279 by Arav et al. (2007) has found an overabundance carbon, nitrogen, and oxygen relative to the Sun by factors of  $2.2 \pm 0.7$ ,  $3.5 \pm 1.1$ , and  $1.6 \pm 0.8$ , respectively. Moreover, photoionization modeling of associated NALs systems find metallicities ranging between a few and 20 times the Solar value (e.g., Petitjean et al. 1994, Tripp, Lu & Savage 1996, D’Odorico et al. 2004). Such high metal abundances are also inferred from analysis of the broad emission lines. For example, Hamann et al. (2002) find that  $\alpha$ -elements, such as C and O, are overabundant in the broad-emission line regions of quasars by a factor of  $\sim 3$  relative to the Sun.

Previous studies indicate that the AALs and BALs are located within the scale of the host galaxy and suggest that the absorbers are physically associated with and may originate from the quasar outflow. For example, the strengths of the excited-state Si II\* AALs in the high resolution spectrum of 3C 191 indicate a distance of 28 kpc from the quasar assuming the gas is photonionized (Hamann et al. 2001). By studying the metastable excited levels of Si II and Si II\*, Dunn et al. (2010) determines the distance to the outflow for attenuated and unattenuated SEDs of  $\sim 6$  kpc or  $\sim 17$  kpc respectively. Variability analysis of a C IV NAL complex towards quasar HS 1603+3820 constrains an upper limit on the distance of  $r \leq 6$  kpc from the continuum source (Misawa et al. 2005).

Although it is generally agreed upon that intrinsic absorption lines are formed in an outflow from the central engine, there is no consensus on the geometry or the driving force of such an outflow. In the model of Murray et al. (1995), which was developed further by Proga et al. (2000), the wind is launched from the entire surface of the accretion disk, with many denser, absorbing filaments distributed throughout. In this picture, the BALs appear at low latitudes (small angles from the plane of the disk) when the line of sight passes directly through the dense, high velocity wind. The NALs could form at higher latitudes when the line of sight intercepts a dense filament in the hot, low-density flow. Elvis (2000), instead, proposed a funnel-shaped thin shell outflow model, in which the wind rises vertically from a narrow range of radii on a disk in the

BELR. The high-ionization UV NALs and the X-ray “warm absorbers” are seen when the disk is viewed at a low latitude, through the wind. When viewed along the fast outflow, the full range of velocity is seen in absorption with a large total column density, giving rise to a BAL. Recent work indicates that the column densities measured in the X-ray spectra of quasars hosting intrinsic NALs are considerably lower than what the Elvis (2000) scenario posits (Misawa et al. 2008; Chartas et al. 2009b), which suggests a preference on the model of Murray et al. (1995). There is also a suggestion that in some objects, the outflows are directed in the polar direction, parallel to the axis of the accretion disk (see, for example, Ghosh & Punsly 2007; Zhou et al. 2006a; Brotherton et al. 2006). In the context of such a picture, and by analogy with equatorial wind models, the absorbers giving rise to NALs could be detached from the dense part of the outflow and become visible along different lines of sight. Previous studies have shown that the absorbing clumps originate at similar physical locations and are driven by radiative acceleration. This is supported by the explanation of a double trough in the C IV BAL by the “line-locking” effect (e.g., Arav et al. 1995). A striking case can be seen in Srianand et al. (2002), which reports an observation of a highly structured flow in which distinct components have similar velocity separations.

Misawa et al. (2007) identified two families of intrinsic absorption lines among 39 intrinsic NAL systems. The “strong N V” family is characterized by strong, partially covered N V, relatively weak Ly $\alpha$  (less than twice the equivalent width of N V), occasionally detected C IV, and O VI lines (sometimes even stronger than the N V lines). The “strong C IV” family, on the other hand, is characterized by strong, partially covered C IV doublets, strong, usually black Ly $\alpha$  lines, and relatively weak or undetected N V. A possible third family, characterized by strong O VI lines is discussed by Ganguly (2011). They primarily used partial covering to diagnose the intrinsic nature of the systems, *i.e.*, they found that the members of the doublet could not be fit simultaneously assuming full coverage. Depending on the geometries and locations of the continuum source and broad emission line regions (and empirically on their positions relative to the emission lines) the absorber can partially cover either or both (Ganguly et al. 1999). There may be overlap between the C IV and O VI categories as some absorption systems can possess both strong C IV and O VI lines. Further observations covering both of these lines are needed to investigate this issue. These different types of NALs are of interest because they may allow us to probe different regions of the outflow. These three families of NALs may represent different lines of sight through the outflow, which is also suggested by the relations between the properties of UV NALs and the X-ray properties of the quasars that display them (e.g., Chartas et al. 2009b).

The question we address in this paper is the origin of the intrinsic N V NALs. To this end we construct photoionization models for the strong N V absorption systems in the spectra of three radio-quiet quasars from the HIRES/Keck sample studied by Misawa et al. (2007). This sample contains 37 optically bright quasars at  $z = 2 - 4$ . We choose these particular three quasars because their spectra exhibit characteristics of the “strong N V” family. In addition, these three systems offer many observational constraints because many ions are covered in their spectra, and the N V absorption lines have *relatively* simple profiles. Similar systems, containing multiple absorption components spread over

thousands of  $\text{km s}^{-1}$  have been reported in the spectra of RX J1230.8+0115 (Ganguly et al. 2003) and 3C 351 (Yuan et al. 2002).

In Section 5.2, we describe the absorption profiles of the three intrinsic N V NAL systems. In Section 5.3, we introduce our method for modelling the three systems using the Cloudy photoionization code (Ferland 2006). The modelling results, in the form of constraints on metallicity, ionization parameter, and volume number density are presented in Section 5.4. In Section 5.6, we discuss possible interpretations of our results on the location of the absorbers in the quasar winds. We present a summary and conclusion in Section 5.7. The cosmology we use in this paper is  $\Omega_\Lambda = 0.7$ ,  $\Omega_M = 0.3$ ,  $h = 0.7$ , which leads to the luminosity distances listed in Table 5.1.

## 5.2 Data

The Keck/HIRES spectra of our three intrinsic N V systems are described in Misawa et al. (2007). The spectral resolution is  $R = 37,500$ , or  $\sim 7 \text{ km s}^{-1}$ . Table 5.1 summarizes the basic properties of the three absorption systems and of the quasars that host them. The transitions listed are detected at a  $5\sigma$  confidence level at the NAL redshift. The sign of the velocity of a system is taken to be positive, if the line is blueshifted, *i.e.*, if the gas appears to be outflowing relative to the quasar; this is the opposite convention from that adopted in Misawa et al. (2007). In the same table, we also list the 4400 Å flux densities (taken from Misawa et al. 2007), as well as the bolometric luminosities and ionizing photon rates (obtained as described in §5.3).

Figures 5.1, 5.2, and 5.3 present, for each system, absorption profiles for transitions which are used as modelling constraints. An example of the best models we have found (discussed in later sections) is also shown in each figure. The velocity of an entire system is defined by the optical depth-weighted center of the strongest member of the N V doublet. We describe each system below.

*HE0130–4021* ( $z_{\text{abs}} = 2.973915$ ; Fig. 5.1): The system is found  $\sim 4000 \text{ km s}^{-1}$  blueward of the quasar emission redshift, just within the associated region. This system is kinematically more complex than the other two, with 9 components in each N V line, spreading over  $\sim 500 \text{ km s}^{-1}$ . All but one of these components have small coverage fractions ( $C_f < 0.5$ ; see Table 5.2). The strong features in the Ly $\alpha$  panel of Figure 5.1 at  $\sim -280$ , 50, 110, and 210  $\text{km s}^{-1}$  cannot be Ly $\alpha$  components because the corresponding Ly $\beta$  components are not detected. Thus only portions of the Ly $\alpha$  profile can be used as constraints, but it is clear that the absorption in Ly $\alpha$  is quite weak. The O VI doublet appears to be detected in many of the same components as N V, though inconsistencies between O VI  $\lambda 1032$  and OVI  $\lambda 1038$  betray a number of blends. For this system the alignment of kinematic components of different transitions is convincing and it seems at least some of the OVI absorption is real. We can use the observed absorption at the positions of O VI  $\lambda\lambda 1032, 1038$  as an upper limit.

*Q1009+2956* ( $z_{\text{abs}} = 2.649510$ ; Fig. 5.2): This system is  $\sim 450 \text{ km s}^{-1}$  redward of the quasar emission redshift. The N V  $\lambda\lambda 1239, 1242$  doublet of this system exhibits partial coverage in each of the three blended kinematic (Voigt) components required

to fit its asymmetric profile, with the strongest component having  $C_f = 0.70 \pm 0.06$  (Fig. 5.2 and Table 5.2). Ly $\alpha$  and Ly $\beta$  are detected, but are unusually weak compared to Nv. The ratio of Ly $\alpha$  to Ly $\beta$  is reasonable, but differences in the shapes of these lines, particularly the presence of a wing on the red side of the Ly $\beta$  profile, suggest a possible blend with another line. We will thus weight the Ly $\alpha$  line more heavily as a constraint. Narrow CIII  $\lambda 977$  and SVI  $\lambda 933$  lines appear to be detected, albeit with some uncertainty because of their location in the Ly $\alpha$  forest. The red member of S VI  $\lambda\lambda 933, 945$  doublet is heavily blended and thus cannot be used to judge whether this doublet is truly detected or not. The feature at the position of Si III  $\lambda 1207$  is not aligned with the C III  $\lambda 977$  and N V  $\lambda 1239$  profiles and is too broad, so we believe that we only have an upper limit on Si III  $\lambda 1207$  absorption from this system. There is absorption at the positions of both members of the O VI  $\lambda\lambda 1032, 1038$  doublet, however the minima are not aligned with each other or with the Nv profile. The O VI  $\lambda 1032$  is too broad relative to OVI  $\lambda 1038$ , and both transitions are clearly affected by blends. The limits on O VI for this system are not restrictive constraints for our models.

*HS1700+6416* ( $z_{\text{abs}} = 2.7125$ ; Fig. 5.3): This system is  $\sim 750$  km s $^{-1}$  blueward of the quasar emission redshift. The N v  $\lambda\lambda 1239, 1242$  and C IV  $\lambda\lambda 1548, 1550$  doublets are detected in this system, as well as Ly $\alpha$  and Ly $\beta$  (Fig. 5.3). The same system was presented in Barlow et al. (1997), who found it to vary significantly on a time scale of  $\sim 6.5$  months in its rest-frame. The central absorption is detected at  $z \sim 2.7125$  ( $v_{\text{rel}} \sim 0$  km s $^{-1}$ ) and a weaker component is found at  $\sim 325$  km s $^{-1}$ . The central component shows an asymmetry, indicating the need for two Voigt components for an adequate fit. The feature at  $\sim 70$  km s $^{-1}$  in the Ly $\alpha$  panel cannot be Ly $\alpha$  belonging to this absorption system, since no counterpart is detected in the Ly $\beta$  panel. All three components of the N v profile show partial coverage, with coverage fractions between 0.33 and 0.43 (Table 5.2). The Ly $\beta$  component at 0 km s $^{-1}$  is too strong relative to Ly $\alpha$ , indicating a blend that affects the Ly $\beta$  so that it can only be used as an upper limit. If saturated (and affected by partial coverage), Ly $\beta$  could have an equivalent width equal to that of Ly $\alpha$ , but it can never have a larger equivalent width. The O VI doublet lies in the Ly $\alpha$  forest, such that it is hard to evaluate whether it is detected for this system or not. If all of the detected absorption, either at 0 km s $^{-1}$  or at 325 km s $^{-1}$ , is due to O VI, the absorber would have a much larger coverage fraction than the N v, C IV, or Ly $\alpha$  absorbers at these same velocities. In our modelling we will take the absorption lines coincident with O VI as an upper limit, but we note that the minima of the absorption at  $\sim 325$  km s $^{-1}$  in both members of the O VI doublet are at the same velocity as those in the N v. Because of this, we will later consider whether viable single phase models can also explain this possible O VI absorption.

Transitions appearing in Figures 5.1, 5.2, and 5.3, but not described above, provide limits for model constraints, either because they are not detected or because they thought to be affected by blends.

### 5.3 Photoionization Modeling

Our calculations were performed with version 07.02.01 of Cloudy, last described by Ferland et al. (1998). Our modeling assumes a series of plane-parallel slabs of gas (clouds) exposed to the ionizing continuum from the central engine. The gas within each absorbing cloud is assumed to have a uniform density, metallicity and abundance pattern. In most cases, we assume a solar abundance pattern (Holweger 2001), so that for *every* element  $i$ ,  $A_i/A_{i,\odot} = Z/Z_\odot$  where  $A_i = n_i/n_H$ . When an adequate fit cannot be achieved using the solar abundance pattern, we consider deviations. The intensity of the ionizing continuum is parameterized by the ionization parameter,

$$U = \frac{n_\gamma}{n_H} = \frac{1}{4\pi r^2 c n_H} \int_{\epsilon_1}^{\epsilon_2} \frac{L_\epsilon}{\epsilon} d\epsilon, \quad (5.1)$$

where  $r$  is the distance of the illuminated face of the cloud from the continuum source,  $\epsilon_1$  is the lowest energy required to photoionize the gas, *i.e.*,  $\epsilon_1 = 1$  Ry, and  $\epsilon_2$  is the high energy cutoff of the SED. In our model, we take  $\epsilon_2 = 7.354 \times 10^6$  Ry. Each of our models are specified by  $U$  and  $n_H$ , which (given the quasar luminosity and spectral shape) corresponds to a given  $r$ . We also assume that all the ions within a cloud have the same coverage fraction, and that the gas cloud is in a state of thermal equilibrium, so it is parameterized by a single electron temperature,  $T_e$ . The assumption of the same coverage fraction for different lines is likely to be valid for those from ions with similar ionization states, particularly C IV and N V which are the most important constraints for our models. Ly $\alpha$  absorption may also arise from additional regions, but in any case will provide a lower limit on metallicity. In our favored models, the clouds are optically thin to the incident continuum so that the incident SED does not change after the ionizing photons pass through an absorbing cloud.

We have measured the Doppler  $b$  parameters, column densities and coverage fractions of the N V doublets, as we describe below, and optimized on these values (*i.e.*, we required that Cloudy models produce them). We choose N V as the transition on which we optimize because it is located in a relatively “clean” spectral region where there are few blends. This allows us to determine its column density, Doppler parameter, and coverage fraction via Voigt profile fitting. The code AUTOVP (Dave et al. 1997) is used to derive an initial solution, and then MINFIT is used to determine the minimum number of components that produce an adequate fit. The goal of the modeling exercise is to reproduce the observed absorption profiles for all other ions by adjusting three physical parameters: metallicity,  $Z/Z_\odot$ , ionization parameter,  $U$ , and hydrogen number density,  $n_H$ .

#### 5.3.1 Voigt Profile Fitting of the N V Doublets

The approach of the MINFIT code is to first “overfit” the system using many Voigt components and then to reject components that do not improve the fits at a confidence level above 95%. This fitting technique has been used extensively in studies of intervening Mg II systems (e.g., Ding et al. 2003; Ding, Charlton & Churchill 2005; Lynch, Charlton & Kim 2006). Table 5.2 lists the resulting fitting parameters. It is

essential to include the coverage fractions as free parameters in the fitting process for these systems since  $C_f = 1$  does not provide acceptable fits to these doublets within the observed uncertainties (Fig. 5.4; Misawa et al. 2007).

### 5.3.2 Specification of the Ionizing Continuum

Although multi-band photometry and/or spectra of these three quasars are not available over the entire spectral range, some observational data are available to place constraints on the input SED. We consider three possible quasar SEDs.

1. A typical quasar SED from Elvis et al. (1994). This is a mean SED determined from 29 moderately luminous quasars.
2. A broken power-law SED from Mathews & Ferland (1987). This is a mean continuum determined by combining direct observations, without distinguishing between radio quiet and radio loud quasars.
3. A multi-component continuum with either a small, a medium, or a big blue bump (BB) (see the Cloudy document, *Hazy* § 6.2, for details), parameterized by  $kT_{\text{BB}} = 10 \text{ eV}$ ,  $100 \text{ eV}$  and  $300 \text{ eV}$ , respectively (Casebeer et al. 2006). By adjusting the parameters of this prescription we produce three different strengths of the BB, which we list in Table 5.3 (“small”, “medium”, and “big”). For each BB strength we examine three different values of  $\alpha_{\text{ox}}$  (denoted as I, II, and III in Table 5.3), which gives us nine different SED shapes for this model.

We utilize three parameters to constrain the SED:  $\alpha_o$ ,  $\alpha_x$ , and  $\alpha_{\text{ox}}$ , which are the UV/optical spectral slope, X-ray band spectral slope and optical-X-ray spectral slope, respectively (assuming that  $f_\nu \propto \nu^{+\alpha}$ ). The last is defined as  $\alpha_{\text{ox}} = 0.3838 \log [\ell_\nu(2 \text{ keV})/\ell_\nu(2500 \text{ \AA})]$ , where  $\ell_\nu(2 \text{ keV})$  and  $\ell_\nu(2500 \text{ \AA})$  are the monochromatic luminosities at 2 keV and at 2500 Å (Tananbaum et al. 1979).

The low resolution ( $R \sim 2000$ ) UV/optical spectrum of HS1700+6416 is available from the Data Release 5 (DR5) quasar catalog (Schneider et al. 2007) of the Sloan Digital Sky Survey (SDSS; York et al. 2000). We fitted the underlying continuum using a power-law, with a small Balmer bump and the iron emission forest superimposed (Vanden Berk 2011), and obtained  $\alpha_o = -0.47$  (see Table 5.3). This object was also observed by *Chandra* and *XMM-Newton*. The X-ray spectrum was adequately fit with a power-law index  $\sim -1$  ( $\alpha_x = -1.2 \pm 0.2$  for the *Chandra* spectrum and  $\alpha_x = -1.1 \pm 0.2$  for the *XMM-Newton* spectrum; Misawa et al. 2008). Thus we adopted  $\alpha_x = -1$ .

To find  $\alpha_{\text{ox}}$  we used the empirical correlation between it and the 2500 Å quasar luminosity (Vignali et al. 2003; Strateva et al. 2005; Steffen et al. 2006). We adopt the  $\alpha_{\text{ox}}-\ell_\nu(2500 \text{ \AA})$  relation by Steffen et al. (2006):

$$\alpha_{\text{ox}} = (-0.137 \pm 0.008) \log \ell_\nu(2500 \text{ \AA}) + (2.638 \pm 0.240).$$

The 2500 Å luminosity is calculated as  $\ell_\nu(2500 \text{ \AA}) = 4\pi d_L^2 f_\nu(2500 \text{ \AA})$ , in which  $d_L$  is the luminosity distance and  $f_\nu(2500 \text{ \AA})$  is the rest-frame flux, derived from an extrapolation

of the flux density at 4400 Å (Misawa et al. 2007) assuming  $f_\nu \propto \nu^{-0.44}$  (Vanden Berk et al. 2001). Thus, for HS1700+6416 we find  $\alpha_{\text{ox}} = -1.87$

The three spectral indices for HS1700+6416, from spectroscopic observations, are compared to those for possible quasar SEDs (Table 5.3). We find that the Medium BB III SED best matches the observational constraints thus, we use this SED as the input ionizing flux for HS1700+6416. This SED is plotted as a solid line in Fig. 5.5.

We measured  $\alpha_o$  of the other two quasars from the low-resolution spectra of Q1009+2956 (Burles & Tytler 1998) and HE0130–4021 (Osmer et al. 1994), respectively. These UV/optical power-law indices are determined by connecting two empirical “line-free” regions redward of the Ly $\alpha$  emission line and are  $\sim -0.25$  for both quasars. The 2500 Å luminosities for these two quasars are determined in the same manner as for HS1700+6416 and their  $\alpha_{\text{ox}}$  values are both  $\sim -1.8$ . There have not been any X-ray spectral observations of Q1009+2956 or HE0130–4021 so, we assume they have the same  $\alpha_x$  as HS1700+6416. Generally, the photoionization cross section decreases rapidly with photon energy so the ionization state is not sensitive to the X-ray spectral slope. We then also use the Medium BB III multi-component SED for both of these quasars, but we adjust the optical slope to  $-0.25$  to match the observations. This SED is shown as a dotted line in Fig. 5.5.

We have integrated the above SEDs to compute the bolometric luminosities and ionizing ( $E > 1$  Ry) photon rates for the three quasars ( $L_{\text{bol}}$  and  $Q$ , respectively). We first normalized them to match the 4400 Å luminosity density of each quasar (using the flux density and luminosity distance from Table 5.1). The resulting values are listed in Table 5.1; they will be useful for our estimates of the properties of the absorbing gas in §5.6.2.

### 5.3.3 Comparison of Cloudy Model Results to the Data and the Role of the Coverage Fraction

Cloudy calculates the electron temperature based upon thermal and ionization balance. Using this temperature, we calculate the microturbulence velocity for Nv based on its observed total Doppler  $b$  parameter assuming that both of them follow Gaussian distribution. We then calculate the  $b$  parameters for all other elements by combining this turbulent  $b$  with the thermal  $b$ , calculated from the atomic weight and the electron temperature. Using these  $b$  parameters and the column densities of each ion, output by Cloudy, we synthesize noiseless model spectra, convolving with a Gaussian function which represents the instrumental profile. We compare the theoretical profiles with the observed ones using a  $\chi^2$  test in combination with visual inspection. The wavelength range within which  $\chi^2$  is evaluated is carefully chosen because even a single pixel affected by a blend or an instrumental artifact can dominate the  $\chi^2$  values.

We apply a grid method in  $(Z/Z_\odot)$ – $U$ – $n_{\text{H}}$  space to search for acceptable solutions by comparing model profiles to the observed absorption lines. In order to better visualize the models and compare them to the observed spectra, we take constant- $n_{\text{H}}$  slices and examine combinations of  $(Z/Z_\odot)$  and  $U$  in this 2-dimensional parameter space. The ranges of  $(Z/Z_\odot)$ ,  $U$ , and  $n_{\text{H}}$  we explore vary depending on the specific system, but

typically they span  $1 \lesssim \log(Z/Z_{\odot}) \lesssim 3^1$ ,  $-1.5 \lesssim \log U \lesssim 1.5$ , and  $2 \lesssim \log(n_{\text{H}}/\text{cm}^{-3}) \lesssim 14$ . The initial increments are  $\Delta \log(Z/Z_{\odot}) = 0.1$ ,  $\Delta \log U = 0.1$ , and  $\Delta \log(n_{\text{H}}/\text{cm}^{-3}) = 2$ . Finer grids are applied only if necessary. We assume that different kinematic (Voigt) components represent different parcels of gas, which we model separately. However, if kinematic components are blended together, we start by assuming that they are all described by the same model parameters. We refer to such models as the “initial” or “preliminary” models. After finding the best preliminary solution, we vary  $(Z/Z_{\odot})$  and  $U$  for each kinematic component within 0.2 dex of their preliminary values to seek better solutions with smaller  $\chi^2$  values.

In the process of producing synthetic line profiles for comparison to the observed spectra, we make use of the coverage fraction inferred from the relative strengths of the lines in the N V doublet (see Table 5.2). Since  $C_f$  represents the fraction of photons from the background source that pass through the absorber, we effectively dilute all the absorption lines of a given system by a factor equal to  $C_f$ . We have to assume that the same value of  $C_f$  applies to all transitions for lack of additional information. This need not be the case in general, however; Misawa et al. (2007) do find cases where different resonance doublets in the same system yield different values of  $C_f$ , for example. In one of our three quasars, HS1700+6416, we do have separate measurements of  $C_f$  for the N V and C IV doublets but they agree with each other within uncertainties. We return to the issue of the covering factor in Section 5.5 where we discuss the possibility of the absorbers covering different fractions of the continuum and broad-emission line sources.

## 5.4 Photoionization Model Results

### 5.4.1 HE0130–4021

This absorption complex can be described by nine kinematic components (Fig. 5.1 and Table 5.2). These components are blended so we model them by first assuming that they are described by the same model parameters. Si IV  $\lambda 1396$  is not detected, which yields a lower limit on  $\log U$ . Although O VI  $\lambda\lambda 1032, 1038$  and C III  $\lambda 977$  do not provide strong constraints, we require that they are not over-produced by our models. The Ly $\alpha$  line provides a strong constraint for this system. Because the saturated absorption troughs around  $50 \text{ km s}^{-1}$  and  $100 \text{ km s}^{-1}$  are not seen in Ly $\beta$ , they cannot be due to Lyman absorption from this system. These two absorption features actually divide the whole Ly $\alpha$  absorption system into a red part (positive velocity) and a blue part (negative velocity), which we can model separately. The blue part comprises Components 1–4 and 9 while the red part consists of Components 5–8. Although this system comprises many kinematics components, we can still obtain robust results for the coverage fraction of each component. This is because we fit the blue and the red members doublet simultaneously so that only features appearing in both of doublets are contribute to the determination of  $C_f$ . The resulting relative uncertainties (listed in Table 5.2) are typically of order 20%, which includes errors in profile fitting as well as errors in continuum fitting.

---

<sup>1</sup>In our initial simulations we tried to reproduce the absorption line strengths with  $\log(Z/Z_{\odot}) \lesssim 1$ . However, we found that acceptable models require relatively high metallicities, which led us to adopt this range of values in our grid search.

We start by seeking models that can reproduce the observed Ly $\alpha$  absorption complex. For either the red or the blue part of the profile, we can find a series of solutions for  $\log(Z/Z_{\odot})$  and  $\log U$  on a constant- $n_{\text{H}}$  slice. Fig. 5.6 illustrates the methodology we apply to all systems: it shows the acceptable models for Ly $\alpha$  for  $\log(n_{\text{H}}/\text{cm}^{-3}) = 8$  and a range of  $\log(Z/Z_{\odot})$  and  $\log U$ . The highest and lowest possible values of  $\log(Z/Z_{\odot})$  and  $\log U$  for different values of  $\log(n_{\text{H}}/\text{cm}^{-3})$  are tabulated in Table 5.4. The blue and red parts of the profile yield different sets of acceptable model parameters. Fig. 5.6 presents constraints on models based on the observed O VI  $\lambda 1032$  and Si IV  $\lambda 1396$  lines. The effect of changing  $\log U$  and  $\log(Z/Z_{\odot})$  on the model profiles of Si IV  $\lambda 1396$  and Ly $\alpha$  is illustrated in Figures 7a, and 7b. The O VI  $\lambda 1032$  profile is very insensitive to metallicity as well as to  $\log U$  when  $\log U \gtrsim -1.5$ . Below this value, the O VI  $\lambda 1032$  absorption decreases gradually. Because the O VI doublet is in the Ly $\alpha$  forest, we use it as an upper limit on the O VI absorption that a model should produce. However, at  $\log U \sim -1.5$ , the red parts of O VI 1032,1038 can be adequately fit. Therefore, the intersection between the acceptable models for the red part of Ly $\alpha$  and the acceptable models for O VI 1032 happens at  $\log U \sim -1.5$  and  $\log(Z/Z_{\odot}) \sim 1.7$ . This model also fits the observed C III  $\lambda 977$  absorption and it does not overproduce Si IV  $\lambda 1396$ . For the blue part of the system, within the range of models that are consistent with the Ly $\alpha$  and Si IV 1394, the O VI  $\lambda 1032$  is never overproduced. Since the O VI is taken as an upper limit, these models are all considered acceptable, as listed in Table 5.4.

The synthesized line profiles are consistent with the data for values of  $\log(n_{\text{H}}/\text{cm}^{-3})$  between 2 and 14. The ranges of  $\log(Z/Z_{\odot})$  and  $\log U$  are slightly different for different  $\log n_{\text{H}}$  values (Table 5.4). An example of an acceptable model is presented in Fig. 5.1.

There are still discrepancies between the model and the observed spectrum of Ly $\alpha$ . Perhaps the Ly $\alpha$  line is blended with another transition because the corresponding absorption trough is not seen in Ly $\beta$ . The  $v \sim 300 \text{ km s}^{-1}$  absorption trough in the Ly $\alpha$  window is also underproduced, which seems to suggest an additional weak component with a small Doppler parameter. Most of the absorption in O VI is underproduced by this model, with the exception of that at  $v \sim 0 \text{ km s}^{-1}$ . The absorption at this velocity is mainly contributed by Component 9, which is the only component with a coverage fraction of unity. We have assumed that the O VI absorption is only an upper limit, but if it is real, a separate phase would be required that has a larger coverage fraction. In the other two systems O VI is similarly underproduced; we discuss this further in Section 5.4.4.

We have attempted to optimise the model parameters for each kinematic component separately but we found that this exercise did not yield useful results. We found a large number of “pockets” in parameter space, spanning a wide range of parameter values, where an acceptable fit can be achieved but there is no strong reason to prefer one particular “pocket” over the others. We attribute this outcome to the fact that the Lyman series lines that we use as constraints are blended with other lines and the profile of C III has a relatively low signal-to-noise ratio. Therefore, Table 5.4 provides the final summary of acceptable model parameters.

### 5.4.2 Q1009+2956

The absorption-line profile of this system can be reproduced using 3 kinematic components (Fig. 5.2 and Table 5.2), which are blended together. In addition to Nv, there are three main constraints: Ly $\alpha$ , C III  $\lambda$ 977, and S VI  $\lambda$ 933. We initially assume that the three components have the same parameters. We find that using a solar abundance pattern, we can always find overlapping regions in parameter space between the best fits of Ly $\alpha$  and C III  $\lambda$ 977 if  $\log(n_{\text{H}}/\text{cm}^{-3}) \lesssim 13$ . These solutions are at  $2.3 \lesssim \log(Z/Z_{\odot}) \lesssim 2.5$  and  $0.2 \lesssim \log U \lesssim 0.4$ , (Figs. 7c, and 7d). However, we cannot find any overlapping parameter combinations between suitable fits for S VI  $\lambda$ 933 and C III  $\lambda$ 977. The S VI  $\lambda$ 933 solutions also do not intersect with acceptable Ly $\alpha$  solutions, except when  $\log(n_{\text{H}}/\text{cm}^{-3}) \approx 12$ . For most densities, only if we lower the sulphur abundance by about 1 dex (relative to the solar abundance pattern) can we obtain acceptable model profiles for all three of the constraints (Table 5.4). In the case that the absorption at the position of S VI  $\lambda$ 933 is contaminated by blends with unrelated lines, an even larger sulphur abundance reduction would be required. In principle, it is possible to produce less S VI  $\lambda$ 933 by increasing the nitrogen abundance, since Nv is the optimized transition, however we find that such an adjustment is not sufficient to fully resolve the discrepancy. For the special case of  $\log(n_{\text{H}}/\text{cm}^{-3}) \approx 12$ , the S VI  $\lambda$ 933 and Ly $\alpha$  profiles can both be reproduced for  $\log(Z/Z_{\odot}) \sim 2.3$  and  $\log U \sim 0.7$ . For that case, with a higher ionization parameter, an alternative acceptable model is found if we raise the carbon abundance by 0.5 dex, leaving the sulfur abundance unchanged from the solar pattern. These results are summarized in Table 5.4.

We next tune separately the metallicity and ionization parameter of each component, assuming that all three components have the same volume density, but allowing this common value of the density to vary. We select the best model parameters at each volume density value using  $\chi^2$  technique and visual inspection (Table 5.5). An example of acceptable models is presented in Fig. 5.2. To obtain this improved fit we vary the metallicities and ionization parameters of different components over a range about 0.2 dex. Through this exercise, we require that S VI  $\lambda$ 933 is not overproduced. The left wing of Ly $\alpha$  cannot be well modelled and probably requires another component which may not be detected in metal-line absorption or may be related to a region to the blue of the Nv  $\lambda$ 1239 profile. We had neglected that in our fit, since we could not consider whether it was consistent with Nv  $\lambda$ 1243. Nevertheless, the modelling results for this system indicate metallicities considerably higher than the solar values.

### 5.4.3 HS1700+6416

This absorption system can be described by 3 kinematic components (Fig. 5.3, Table 5.2). Because the third component, at  $v \sim 315 \text{ km s}^{-1}$ , is well separated from the central two components, we determine its parameters independently. Assuming the central two clouds to have the same parameters, the preliminary best fits, consistent with both Ly $\alpha$  and C IV  $\lambda$ 1549 as well as with undetected transitions, are tabulated in Table 5.4. In Figures 7e and 7f we plot the profiles of the two central components of Ly $\alpha$  and C IV in the  $\log(Z/Z_{\odot})$ - $\log U$  parameter slice around this preliminary best fit for  $\log(n_{\text{H}}/\text{cm}^{-3}) = 2, 4, 6, 8, 10$ , and 12. For  $2 \lesssim \log(n_{\text{H}}/\text{cm}^{-3}) \lesssim 8$ , the line strengths are

almost independent of density, but for higher densities we see weaker Ly $\alpha$  and slightly stronger C IV absorption. When  $\log(n_{\text{H}}/\text{cm}^{-3}) \gtrsim 11$ , the acceptable fits to C IV  $\lambda 1549$  occur at  $\log U \sim 0.6$ , but Ly $\alpha$  is underproduced by the models so that it is impossible to obtain an acceptable fit at these very high densities. Observationally, variability on a timescale of 6.5 months in the quasar's rest-frame implies  $\log(n_{\text{H}}/\text{cm}^{-3}) > 3$  (based on considerations of the recombination time; see §5.2).

To refine our preliminary models, we next allow the central two components to have different parameters. Thus, we adjust  $\log(Z/Z_{\odot})$  and  $\log U$  of each cloud separately, within  $\pm 0.2$  dex of the preliminary best fit. A complete tabulation of the best model parameters at each value of the volume density can be found in Table 5.6. Take  $\log(n_{\text{H}}/\text{cm}^{-3}) = 8$  for example, a better fit is found at  $\log(Z/Z_{\odot}) = 1.3$ ,  $\log U = 0.5$  for the first component and  $\log(Z/Z_{\odot}) = 1.6$ ,  $\log U = 0.6$  for the second component. Even for this example, shown in Fig. 5.3, the Ly $\alpha$  components at  $v \sim 0 \text{ km s}^{-1}$  are slightly underproduced. This may imply that an additional component would be needed, or that the Ly $\alpha$  is blended with a line from a system at another redshift. The latter is not unlikely given the unidentified component present at  $v \sim 70 \text{ km s}^{-1}$ .

The best fit for the third kinematic component at  $v \sim 315 \text{ km s}^{-1}$  is achieved at  $\log(Z/Z_{\odot}) = 1.7$ ,  $\log U = 1$  at  $\log(n_{\text{H}}/\text{cm}^{-3}) = 2\text{--}8$ . We cannot obtain any acceptable fits for  $\log(n_{\text{H}}/\text{cm}^{-3}) > 8$  because model Ly $\alpha$  profiles become broader than the observed profiles for cases that match the observed C IV (Figures. 7g and 7h). Combining all three components, Fig. 5.3 shows one of the acceptable fits we obtain for this system.

#### 5.4.4 The O VI Doublet

The energy required to produce O VI is much higher (8.371 Ry) than C IV (3.520 Ry) and N V (5.694 Ry). Thus, the O VI doublet probes highly-ionized gas and could, in fact, arise in a more tenuous gas that is not co-spatial with the gas responsible for the lower-ionization absorption lines and also has a different coverage fraction of the background source(s). With the above in mind, we note that in the case of HS1700+6416 (Fig. 5.3), the best model cannot fully reproduce the O VI absorption in both the central and redward components. This could signal that the O VI absorber has a larger coverage fraction. Another possibility is that the O VI is contaminated by lines from the Lyman forest, which is supported by the large widths of the “black” troughs around the O VI  $\lambda 1038$  line of HS1700+6416 and Q1009+2956 (Figures. 5.2 and 5.3). Almost all spectral regions around O VI for the three quasars have some indication of blending, which makes the comparison of the synthesized O VI doublet profile to the data ambiguous. Therefore, we can only use the observed O VI absorption troughs to set an upper limit on the absorption produced by our models.

### 5.5 Partial Coverage of Continuum and Broad-Emission Line Regions

In principle, the intrinsic absorbers could cover different fractions of the continuum source and the BELR (e.g., Ganguly et al. 1999). Thus an absorption line superposed on a broad emission line represents the absorption of different fractions of emission-line and continuum photons. If this is true, the coverage fraction one obtains

from fitting a doublet is an *effective* coverage fraction, which represents the fraction of the photons from all the background sources that pass through the absorber. Thus, for an absorption line that is superposed on a broad emission line and the continuum, a more general expression for the normalized residual flux at a certain velocity is given by (see Ganguly et al. 1999)

$$R(v) = 1 - \frac{\left[1 - e^{-\tau(v)}\right] [C_c(v) + W(v) C_e(v)]}{1 + W(v)} \quad (5.2)$$

(see also the discussion and more general formalism in Gabel et al. 2005). In the above expression,  $C_c$  and  $C_e$  represent the coverage fractions of the continuum source and of the BELR and  $W = f_e/f_c$  is the ratio of the flux contribution of the emission line (without continuum) and the continuum only at the position of the absorption line. In principle  $C$  and  $W$  are functions of velocity across the absorption-line profile. However, since the absorption lines of interest here are fairly narrow,  $W$  does not change appreciably over the profile of the absorption line and we can regard it as independent of velocity. By rearranging equation (5.2) one can show that the effective coverage fraction,  $C_f$ , can be expressed as the weighted average of  $C_e$  and  $C_c$  as

$$C_f = \frac{C_c + WC_e}{1 + W}. \quad (5.3)$$

In practice, we can determine  $C_f$  by fitting the profiles of the UV resonance doublets and  $W$  by measuring the strength of the broad emission line relative to the continuum at the location of the absorption doublet. This leads to a relation between  $C_e$  and  $C_c$  but it does not allow us to determine  $C_e$  and  $C_c$  separately. An example of a graphical depiction of this relation can be found in Fig. 10 of Ganguly et al. (1999).

The N V absorption doublet presents a further complication because it is likely to be superposed on the *blend* of the Ly $\alpha$  and N V emission lines. The regions emitting Ly $\alpha$  and N V are likely to have a different spatial extent and their coverage fractions should also be different. This is suggested by the results of reverberation mapping studies (see, for example Peterson & Wandel 2000; Onken & Peterson 2002; Kollatschny 2003), which indicate an ionization stratification in the BELR. Under these circumstances, in the numerator of equations (5.2) and (5.3), we must replace  $WC_e$  by  $W_{\text{Ly}\alpha}C_{\text{Ly}\alpha} + W_{\text{N V}}C_{\text{N V}}$  and in the denominator, we must replace  $W$  by  $W_{\text{Ly}\alpha} + W_{\text{N V}}$ . In this new notation  $C_{\text{Ly}\alpha}$  and  $C_{\text{N V}}$  are the coverage fractions of the Ly $\alpha$  and N V emitting regions and  $W_{\text{Ly}\alpha}$  and  $W_{\text{N V}}$  are defined in a manner analogous to the definition of  $W$ .

Putting aside the above complication, we can invert equation (5.3) to obtain an expression for  $C_c$ , in terms of  $C_f$ ,  $W$ , and  $C_e$ . However, since  $C_e$  is unknown in practice, this relation can only lead to upper and lower bounds on  $C_c$ , corresponding to  $C_e = 0$  and 1, respectively. Even if the absorption doublet is superposed on a blend of emission lines, the expression for  $C_f$  reduces to equation (5.3) if the BELR is fully covered or not covered at all. Taking the error bars on the measured quantities into account we obtain the following expressions for the limits on  $C_c$ :

$$\begin{aligned} C_c^{\max} &= \min \left\{ 1, [1 + (W + \delta_W)](C_f + \delta_{C_f}) \right\} \quad \text{and} \\ C_c^{\min} &= \max \left\{ 0, [1 + (W - \delta_W)](C_f - \delta_{C_f}) - (W + \delta_W) \right\}, \end{aligned} \quad (5.4)$$

where  $\delta_X$  denotes the error bar on  $X$ .

In an analogous manner, we can use equation (5.3) to derive limits on  $C_e$ , using the fact that  $C_e$  must lie between 0 and 1. The corresponding expressions, including uncertainties on the measured quantities, are:

$$C_e^{\max} = \min \left\{ 1, \frac{[1 + (W + \delta_W)](C_f + \delta_{C_f})}{(W - \delta_W)} \right\} \quad \text{and} \quad (5.5)$$

$$C_e^{\min} = \max \left\{ 0, \frac{[1 + (W - \delta_W)](C_f - \delta_{C_f})}{(W + \delta_W) - 1} \right\}. \quad (5.6)$$

We have used flux-calibrated, low-resolution spectra of the three quasars to measure the values of  $W$  for the Ly $\alpha$ , and Nv absorption lines (as well as C iv absorption line in the case of HS1700+6416). Using these values of  $W$  and the values of  $C_f$  from Table 5.2 we have estimated the limits on the coverage fractions of the continuum source and the BELR,  $C_c$  and  $C_e$ , which we list in Table 5.7. Limits derived from the Ly $\alpha$  absorption line involve the assumption that the effective coverage fraction of the Ly $\alpha$  absorber is the average of the values measured for the Nv absorber, which need not be correct. Nevertheless, these limits appear to be consistent with those derived from the Nv and C iv doublets. The limits on the continuum coverage fraction derived from different Nv kinematic components in the same quasar appear to be consistent with each other with one exception: component 9 of HE0130–4021, for which the effective coverage fraction is measured to be  $C_f = 1$ . The only reliable limit on the BELR coverage fraction comes from the C iv doublet of HS1700+6416.

A closely related issue is whether the superposition of absorption lines on strong emission lines dilutes their strengths by different amounts and leads to large errors in our estimated abundances. The relative strengths of the Ly $\alpha$  and Nv absorption lines are particularly sensitive to this effect, especially when the blueshift of the absorption lines is small. In a specific scenario where the absorber covers only the continuum source, the Ly $\alpha$  absorption line is superposed on the peak of the Ly $\alpha$  emission line and can be diluted much more than the Nv line which is superposed on the red wing of the Ly $\alpha$  emission line (the velocity difference between the Ly $\alpha$  and Nv absorption lines is approximately 5925 km s $^{-1}$ ). In such a case, the procedure we follow here to determine column densities, will lead us to underestimate the column density of Ly $\alpha$  relative to that of Nv and our models would yield a nitrogen abundance that is higher than the true one.

Examining the values of  $W$  listed in Table 5.7, we see that if such a scenario is true, the Nv absorption line is actually diluted more than the Ly $\alpha$  absorption line in HE0130–4021 and HS1700+6416. In other words, in these two quasars we would

have underestimated the nitrogen abundance rather than overestimated it. To verify our expectation quantitatively, we explored a new set of models for HS1700+6416 in which we assumed that the absorber covers the continuum source but not the BELR. Thus, we started by adopting the  $C_f$  value of Cloud 2 because it is the dominant component and set  $C_e = 0$  and  $C_c = 0.64$  for all transitions. As a result, the *effective* coverage fraction for each transition depends on its  $W$  value, *e.g.*,  $C_f = 0.33$  for Nv,  $C_f = 0.45$  for C IV, and  $C_f = 0.34$  for Ly $\alpha$ . By setting  $\log(n_{\text{H}}/\text{cm}^{-3}) = 8$ , we find an acceptable fit at  $(\log(Z/Z_{\odot}), \log U) = (1.6, 0.8)$ . Compared to our best fit with  $C_c = C_e$ , this metallicity is 0.2 dex higher and the ionization parameter is 0.4 dex higher.

In Q1009+2956 we have the opposite situation, in which  $W_{\text{Ly}\alpha}$  is larger than  $W_{\text{Nv}}$ , *i.e.*, Ly $\alpha$  is diluted more by emission line than N v, suggesting that metallicity may have been overestimated. To assess whether this is indeed the case, we explored a new set of models in which the  $C_e$  has the minimum value possible and  $C_c$  has the maximum value possible. According to Eq. (5.4) and (5.6), these values are  $C_c = 1$  and  $C_e = 0.245$ . Thus, the effective covering factors are  $C_f = 0.593$  for Ly $\alpha$ , and  $C_f = 0.895$  for C III. By setting  $\log(n_{\text{H}}/\text{cm}^{-3}) = 8$ , we find an acceptable model at  $(\log(Z/Z_{\odot}), \log U) = (2.5, 0.4)$ , very similar to our previous best-fitting model which had  $C_c = C_e$ .

In the calculation above, we assumed the same *average* value of  $W$  of two members of the N v doublets. However, the value of  $W$  could differ between the blue and red members of the N v doublet, which could affect not only the metallicity estimate but also values of the coverage fractions. In this particular case, however, the value of  $W$  is the same for the two members of the double, within uncertainties.

## 5.6 Discussion

### 5.6.1 High Metallicities

One of the main results of this work is that the metallicities in intrinsic NAL systems are quite high,  $\log(Z/Z_{\odot}) \gtrsim 1$ . Several possible scenarios have been proposed to produce super-solar metallicities in the vicinity of the central massive black holes of quasars (Hamann & Ferland 1999). The most natural scenario involves normal galactic chemical evolution (see Hamann & Ferland 1992, 1993). By considering spectral synthesis and chemical enrichment models of the Nv/CIV and Nv/He II emission line ratios, it was suggested that nitrogen is over-abundant by factors of 2–9 in high redshift ( $z > 2$ ) quasars (Hamann & Ferland 1999). Highly evolved gas with  $\log(Z/Z_{\odot}) > 0.5$ –1.1 is needed to produce such a large over-abundance of nitrogen. This evolutionary model requires a large number of high mass stars and implies a flat initial mass function (IMF) and rapid star formation. The power-law index of the IMF ( $\Phi \propto M^{\gamma}$ ) can be in the range  $\gamma = 0.9$ –1.4 and star formation must occur in less than 0.5 Gyr for  $z > 4$  objects (Hamann & Ferland 1992). This vigorous star formation is consistent with that found in the environment of active galactic nuclei (AGNs). The required time scales, metallicities, and IMFs are similar to those of elliptical galaxies and bulges of disk galaxies (Romano et al. 2002). Quasars with more massive host galaxies tend to be more metal rich, based on observations of massive ellipticals (Clemens et al. 2009; De Lucia et al. 2006). Therefore, the high abundances derived from our absorbing systems are not surprising,

assuming this scenario. The quasars could undergo extensive chemical evolution before they are observable. Other possible processes within the accretion disk that can lead to supersolar metallicities include metal enrichment by star trapping (Artymowicz et al. 1993; Shields 1996), star formation (Collin & Zahn 1999; Goodman & Tan 2004), and non-stellar nucleosynthesis (Kundt 1996). However, it is unlikely that these mechanisms can produce metallicities as high as  $\log(Z/Z_{\odot}) \sim 1$ .

Early measurements of extremely high quasar metallicities based on BALs (Turnshek 1986) could have been overestimated because partial coverage of the background emission was not taken into account, leading to underestimation of column densities (Hamann & Ferland 1999). However, since our analysis of NALs takes partial coverage into account, our measurements do not suffer from this bias. Actually, super-solar metallicities have also been derived for associated absorption systems in low redshift AGNs (Papovich et al. 2000; Fields et al. 2005; Gabel et al. 2006). However, those metallicities are not higher than  $\log(Z/Z_{\odot}) \sim 1$ , *i.e.*, somewhat lower than our results for quasar intrinsic NAL systems. This difference could be due to the higher redshifts of our NAL systems and/or due to their presence in more luminous AGNs. Rapid and vigorous stellar evolution is likely to accompany quasar activity, as suggested by a correlation between the Eddington ratio and the BELR metal abundance spanning 3 orders of magnitude in both quantities (e.g., Shemmer et al. 2004). The resulting high-metallicity gas could then be diluted with a low-metallicity interstellar medium as it moves outward in the host galaxy. Eventually, it can escape the host-galaxy potential, enriching the interstellar and intergalactic medium.

It is also possible that we are looking only at the tip of the iceberg, *i.e.*, at the most highly enriched gas. A diversity of absorber properties, including metallicity, has been reported by Petitjean et al. (1996). Not only can absorbers have different coverage fractions for the continuum and BELR but also different ions/species can have different coverage fractions due to the multi-phase structure of the gas. Different coverage fractions for different ions/species are also inferred from the general properties of the strong C IV family of intrinsic NALs (Misawa et al. 2007), which usually have a saturated Ly $\alpha$  line, but no detected N V, implying a lower ionization state for the gas. In the strong-N V NAL systems studied in this paper, the O VI doublet may sometimes be under-produced by the models, which may be the result of a greater coverage fraction of the O VI absorber. Thus the high metallicities that we derived for the NAL gas in which N V is detected need not apply to all of the gas surrounding the central engine.

### 5.6.2 Properties of the Absorbing Gas

The ionization parameters that we obtained are similar to those for BALs. A typical value of a BAL ionization parameter is  $\log U \sim 0$  (with a range of  $-1.0$  to  $+0.6$ ; e.g., Crenshaw et al. 2003; Wampler et al. 1995; Telfer et al. 1998). Both an intrinsic NAL and a BAL are observed in the spectrum of Q0059–2735 by Wampler et al. (1995), who show that both can be fitted with  $\log U = -0.7$ . Hamann (1998) finds that  $\log U \gtrsim -0.6$  is required to model the BAL of PG1254+047, while Telfer et al. (1998), using a two-slab geometry to model the BAL system of QSO SBS 1542+541, find  $-1.1 < \log U < 0.6$  for the lower-ionization zone and  $\log U \gtrsim 0.3$  for the higher-ionization zone.

If we assume that the outflow includes thin filaments (justified below) that are responsible for the UV absorption lines that we observe, we can estimate the outflow rate in these filaments as

$$\dot{M} = 4\pi r^2 \left( \frac{\delta r}{r} \right) \rho v \left( \frac{\delta\Omega}{4\pi} \right), \quad (5.7)$$

where  $\rho$  is the typical mass density of absorbing filaments,  $r$  is their typical distance from the center,  $\delta r$  is their typical thickness in the radial direction and  $\delta\Omega$  is the solid angle they subtend to the center of the flow. We can re-cast the above expression by using the definition of the ionization parameter,  $U \equiv Q/4\pi r^2 n_{\text{H}} c$ , to replace  $4\pi r^2 \rho$  ( $Q$  is the rate of emission of ionizing photons at  $E > 1$  Ry, and  $n_{\text{H}}$  is the hydrogen number density). We also note that the thickness of a filament is related to the hydrogen density, and radial column density,  $N_{\text{H}}$ , via  $N_{\text{H}} = n \delta r$ . Thus we can write

$$\dot{M} = 5.2 \times 10^{-5} N_{18} v_2 f_{-1} \left( \frac{Q_{58}}{n_6 U} \right)^{1/2} M_{\odot} \text{ yr}^{-1}, \quad (5.8)$$

where  $N_{\text{H}} = 10^{18} N_{18} \text{ cm}^{-2}$ ,  $v = 10^2 v_2 \text{ km s}^{-1}$ ,  $(\delta\Omega/4\pi) = 10^{-1} f_{-1}$ ,  $Q = 10^{58} Q_{58} \text{ s}^{-1}$ , and  $n_{\text{H}} = 10^6 n_6 \text{ cm}^{-3}$ . We also note that from the definition of the ionization parameter, we can obtain an estimate of the distance of the absorbing filament from the ionizing source.

To justify the assumption that the absorbers are thin, we note that using the definition of  $U$  and  $N_{\text{H}} = n \delta r$ , we obtain  $\delta r/r = N_{\text{H}}^2 4\pi c U / n_{\text{H}} Q$ . The values of  $N_{\text{H}}$  (neutral plus ionized) inferred from our photoionization models are in the following range:  $\log(N_{\text{H}}/\text{cm}^{-2}) \sim 16\text{--}19$  for HE0130–4021,  $\log(N_{\text{H}}/\text{cm}^{-2}) \sim 17\text{--}18$  for Q1009+2956, and  $\log(N_{\text{H}}/\text{cm}^{-2}) \sim 19$  for HS1700+6416, the values of  $U$  are given in Tables 5.4–5.6 and the values of  $Q$  are given in Table 5.1. Thus, any combination of values of the above quantities yields  $\delta r/r < 0.01$  which indicates that the absorbers are geometrically thin along the line of sight.

We can now use equation 5.8 to obtain constraints on the mass outflow rate. The necessary values of  $N_{\text{H}}$ ,  $v$ ,  $U$  and  $Q$  are given in Tables 5.4–5.6. We also use a value of  $f_{-1} \sim 5$  based on the results of Misawa et al. (2007). Unfortunately, we have very poor constraints on the hydrogen volume density, which lead to a very wide range of values for the mass outflow rate, i.e.  $\dot{M} \sim \text{few } M_{\odot} \text{ yr}^{-1}$  for the lowest densities and  $\dot{M} < 10^{-4} M_{\odot} \text{ yr}^{-1}$  for  $n_{\text{H}} \geq 10^{10} \text{ cm}^{-3}$ . The kinetic power of these filaments is negligible compared to the electromagnetic luminosity ( $\varepsilon_{\text{k}} \equiv \frac{1}{2} \dot{M} v^2 / L_{\text{bol}} \lesssim 10^{-5}$ ), however they may be embedded in a hotter, more massive outflow or they may be accompanied by a more massive outflow along a different line of sight.

Finally, we note that the poor constraints on the density also lead to poor constraints on the distance of the filament from the ionizing source,  $r$ . In Tables 5.5–5.6 we list the constraints on the radial distance of the filaments as a function of density for Q1009+2956 and HS1700+6416 (the results for HE0130–4021 are very similar). For  $n_{\text{H}} \geq 10^9 \text{ cm}^{-3}$ , the absorbers turn out to be very close to the ionizing source, at distances comparable to the size of the BELR, while at the lowest densities, the filaments are located in the outskirts of the host galaxy.

## 5.7 Summary and Conclusions

In this paper, we modelled the intrinsic NAL systems in the spectra of the quasars HE0130–4021, Q1009+2956, and HS1700+6416. We identify these systems as intrinsic because they exhibit partial coverage in the N V  $\lambda\lambda 1239, 1242$  doublet. Using the photoionization code Cloudy, we simulate and reconstruct each absorption feature observed in a system by adjusting the metallicity (relative to the Sun),  $Z/Z_{\odot}$ , ionization parameter,  $U$  and hydrogen volume density,  $n_{\text{H}}$ . Our main conclusions are:

1. All three intrinsic systems have metallicities of  $\log(Z/Z_{\odot}) \gtrsim 1\text{--}2$ , regardless of constraints on  $U$  or  $n_{\text{H}}$ . We find the high metallicities that we infer cannot alternatively be explained by different coverage fractions of the continuum source and BELR. The origin of these supersolar metallicities is unknown, but we speculate that they may result from intensive star formation before the quasar becomes observable. They could also be related to an inhomogeneous metallicity distribution in the vicinity of the quasar central engine.
2. All three intrinsic systems require high ionization parameters,  $\log U \sim 0$ , which are similar to those derived for BALs.
3. We can constrain the coverage fractions of the continuum source and the BELR separately, and find that the continuum source is not fully covered by the absorbers.
4. Although we cannot constrain the hydrogen volume density very well, we do find that the widths of the synthesized line profiles become larger than the observed ones at very high densities. This is a result of a sharp increase in temperature at high densities, produced by free-free absorption of infrared photons by free electrons. This leads us to conclude that the hydrogen volume density is in the range  $2 \lesssim \log(n_{\text{H}}/\text{cm}^{-3}) \lesssim 9$ .
5. As a result of the large range of possible densities we cannot obtain robust estimates of the mass outflow rate in the filaments responsible for the N V absorption. We can conclude, nevertheless, that  $\dot{M} \lesssim \text{few } M_{\odot} \text{ yr}^{-1}$ , and the corresponding power is a negligible fraction of the electromagnetic luminosity.
6. The strength of the O VI  $\lambda\lambda 1032, 1038$  doublet is underproduced by the models that reproduce the other absorption lines. This suggests that the O VI lines are either contaminated by lines in the Lyman forest or that they arise in a separate gas phase with a larger coverage fraction.

The strong-N V absorbers studied in this paper make up one of two, apparently distinct, families of intrinsic absorber identified by (Misawa et al. 2007). Our next goal is to study the physical properties of the other family of absorbers, the strong-C IV family and probe the origin of the two families.

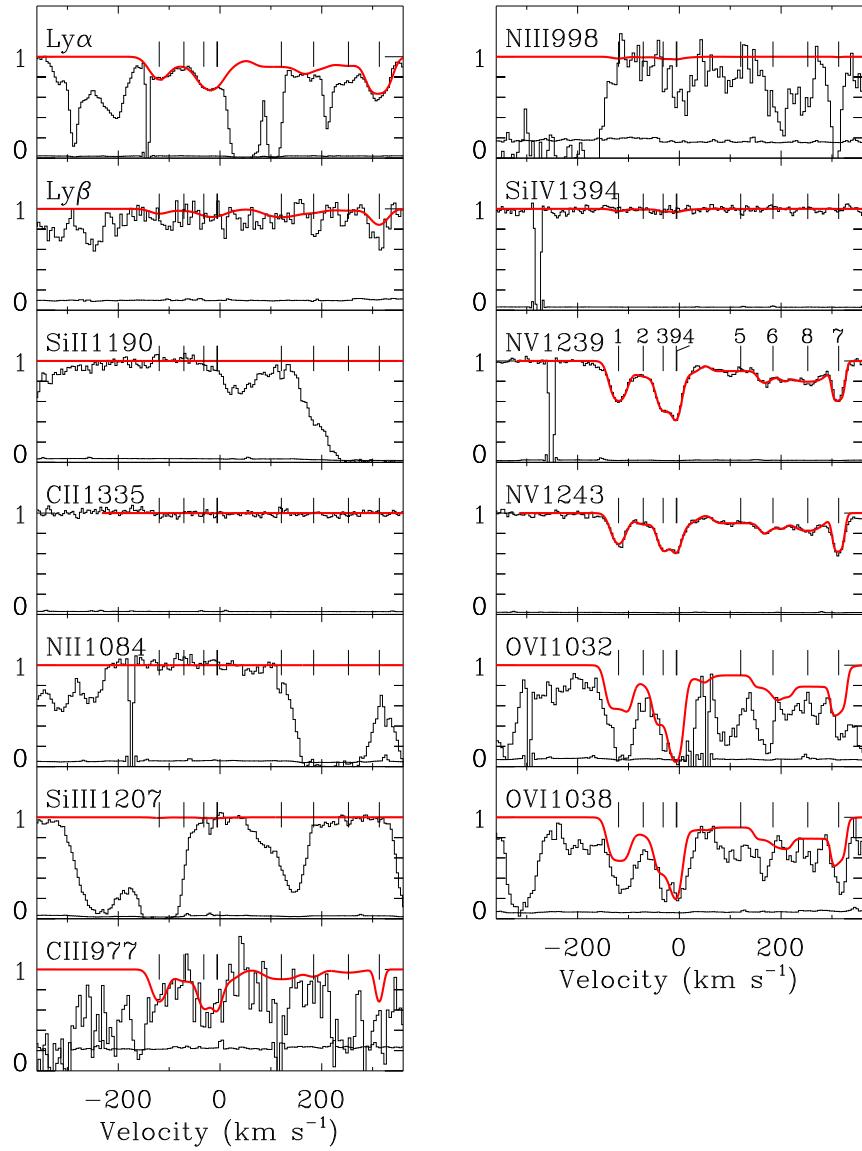


Fig. 5.1 Plots of absorption lines from all the transitions used as constraints for the  $v_{\text{ej}} = 4204 \text{ km s}^{-1}$  ( $z_{\text{abs}} = 3.973915$ ) absorption system in the spectrum of HE0130–4021. Transitions are ordered by increasing ionization potential, except that Ly $\alpha$  and Ly $\beta$  (H $\alpha$ 1026) are always the first. The zero-point of the velocity scale is set by the bisector of the strongest blend of the N V line (the bisector divides the optical depth of the blend in half). Over-plotted (red) in each panel is the best theoretical profile we obtained by adjusting the photoionization model parameters (see Section 5.4 for details). We mark each kinematic (Voigt) component using a tick mark; the component numbers are printed in the N V  $\lambda 1239$  panel. The corresponding components for different transitions have the same partial coverage factor (see Section 5.3). The corresponding components for different transitions have the same partial coverage factor (see Section 5.3), which are listed in Table 5.2.

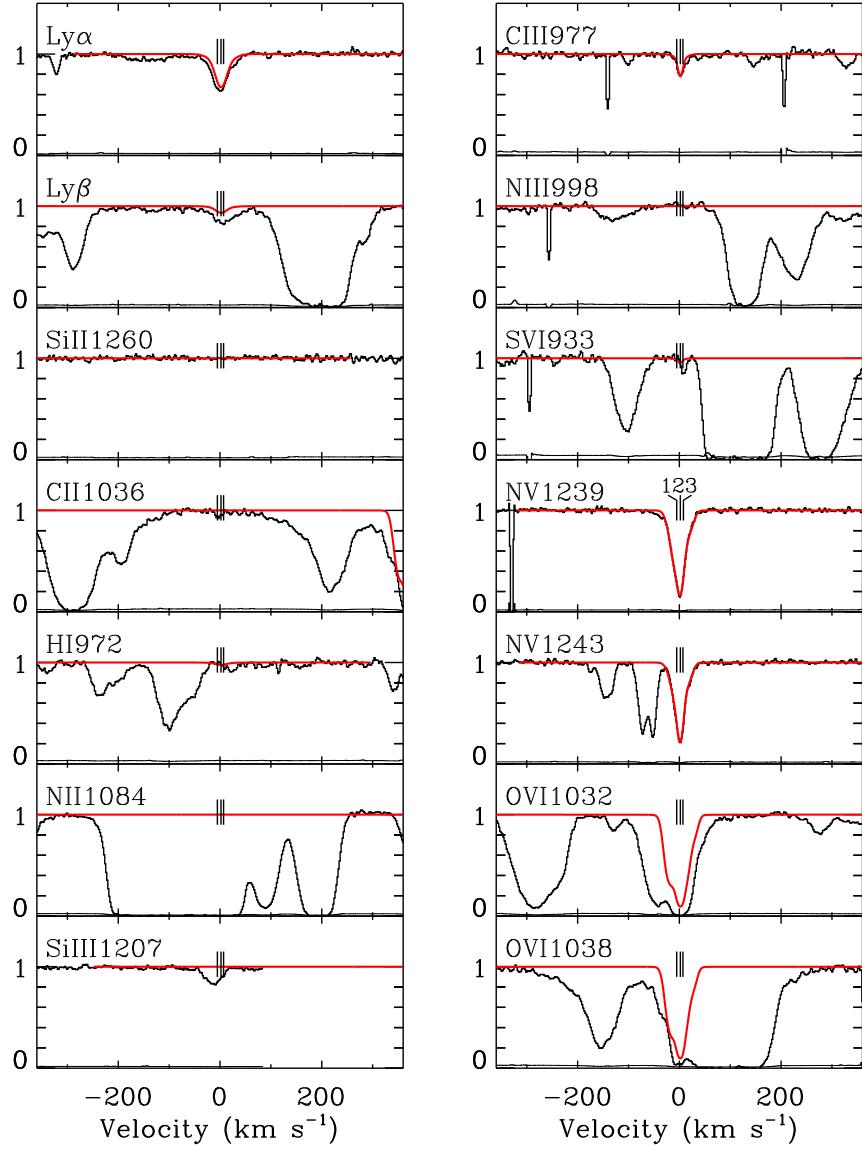


Fig. 5.2 Same as Fig. 5.1 but for the  $v_{\text{ej}} = -453 \text{ km s}^{-1}$  ( $z_{\text{abs}} = 2.649510$ ) absorption system in the spectrum of Q1009+2956.

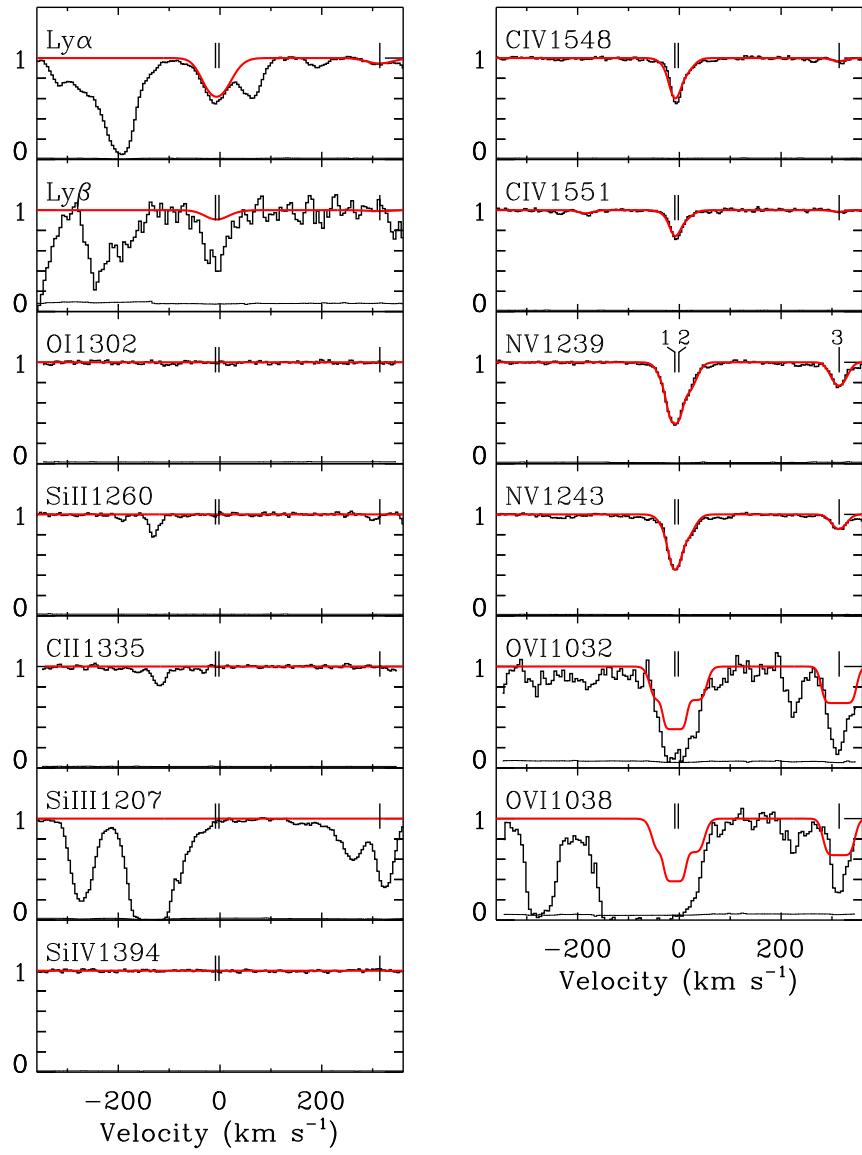


Fig. 5.3 Same as Fig. 5.1 but for the  $v_{\text{ej}} = 764 \text{ km s}^{-1}$  ( $z_{\text{abs}} = 2.7125$ ) absorption system in the spectrum of HS1700+6416.

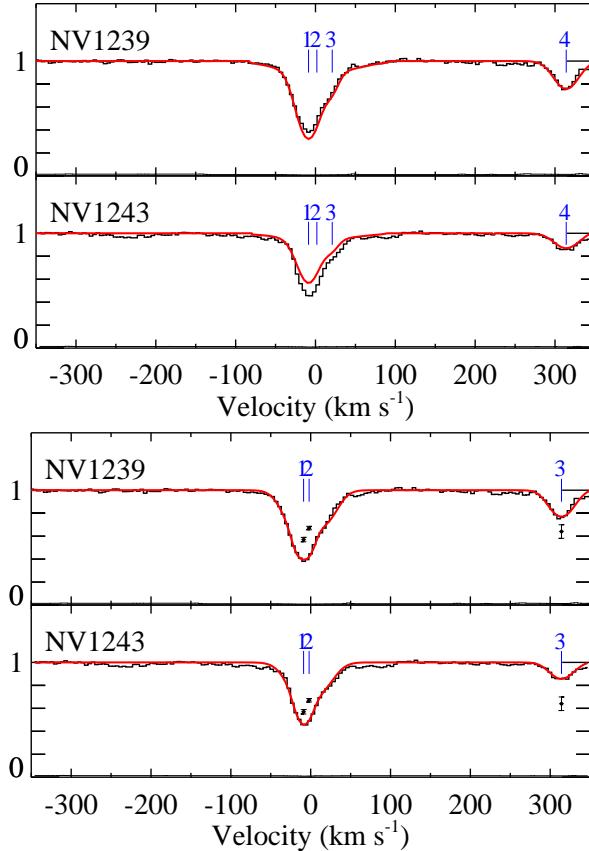


Fig. 5.4 Voigt profile fits to the Nv  $\lambda\lambda 1239, 1242$  doublet in the  $v_{\text{ej}} = 764 \text{ km s}^{-1}$  ( $z_{\text{abs}} = 2.7125$ ) system in the spectrum of HS1700+6416. The velocity zero point is defined as in Fig. 5.1. The fitting process minimizes the number of kinematic components needed for an adequate representation of the profile. Vertical ticks mark the component centers. In the upper pair of panels the coverage factors of all components are fixed to be unity; in the lower pair of panels, the coverage factors of all components are treated as free parameters. Black dots with error bars represent the value of  $(1 - C_f)$  for each component. Clearly,  $C_f = 1$  does not provide an adequate fit, so that partial coverage is required.

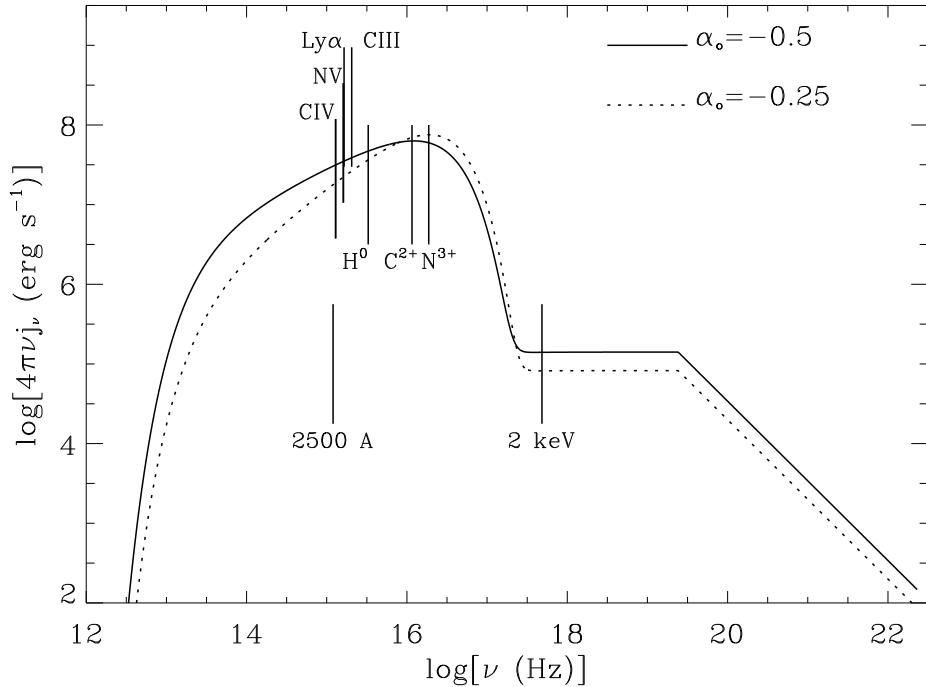


Fig. 5.5 Shapes of two SEDs we used as input to our photoionization models. The solid line shows the SED used for HS1700+6416 while the dotted line shows the SED used for the other two quasars. These SEDs are normalized to an ionization parameter  $\log U = 0.1$  for a hydrogen density of  $\log(n_{\text{H}}/\text{cm}^{-3}) = 8$ . Thus, they have the same photon luminosity between 1 and  $7.354 \times 10^6$  Ry. Also labelled are positions of important transitions – Ly $\alpha$ , N V  $\lambda\lambda 1239, 1242$  and C IV  $\lambda\lambda 1548, 1550$ . We also mark the photon frequencies needed to ionize N $^{+3}$  to N $^{+4}$ , C $^{+2}$  to C $^{+3}$ , and H $^0$  to H $^+$ . We also mark the frequencies corresponding to 2500 Å and 2 keV since flux densities at these two frequencies are used to define  $\alpha_{\text{ox}}$ .

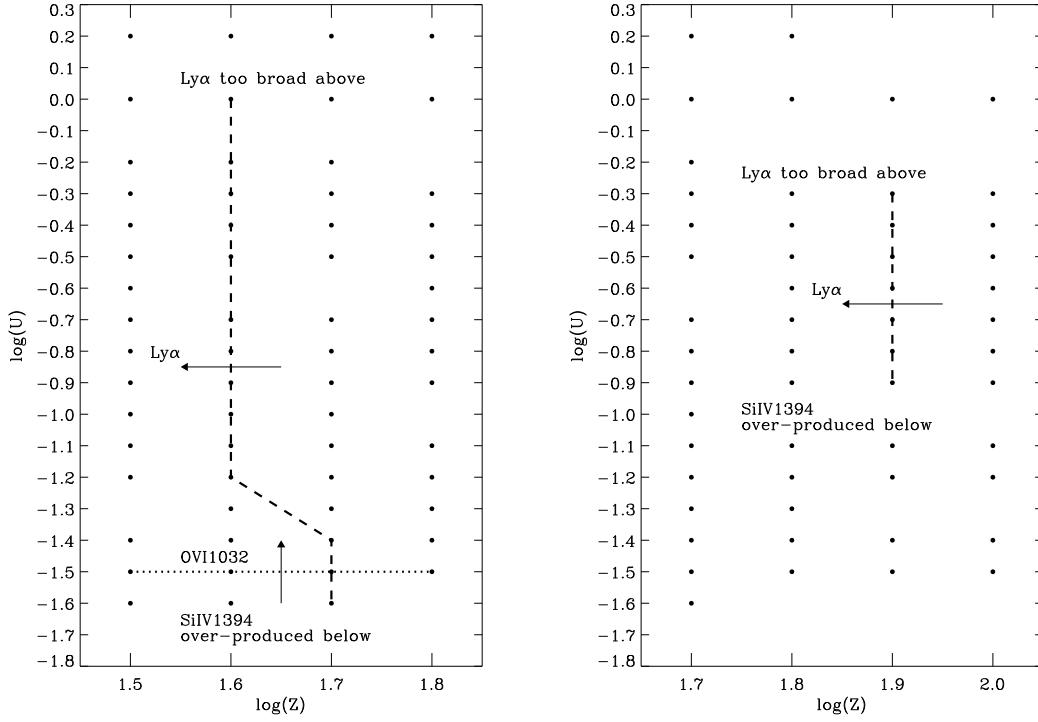


Fig. 5.6 The  $\log(Z/Z_{\odot})$ - $\log U$  parameter slice at  $\log n_{\text{H}} = 8$ , showing the tracks of acceptable preliminary models for the absorption system in the spectrum of HE0130–4021. The left panel shows the tracks for the blue group (kinematic components 1–4, and 9). The right panel shows the tracks for the red group (kinematic components 5–8). Black dots are the cases we covered in our modeling. The dashed line represents the best-fit track for  $\text{Ly}\alpha$  while the dotted line represents the best-fit track for O VI. Arrows indicate the directions in which a specific theoretical absorption profile becomes stronger, so on the “tail” side, the absorption line is under-produced while on the “head” side, it is over-produced.

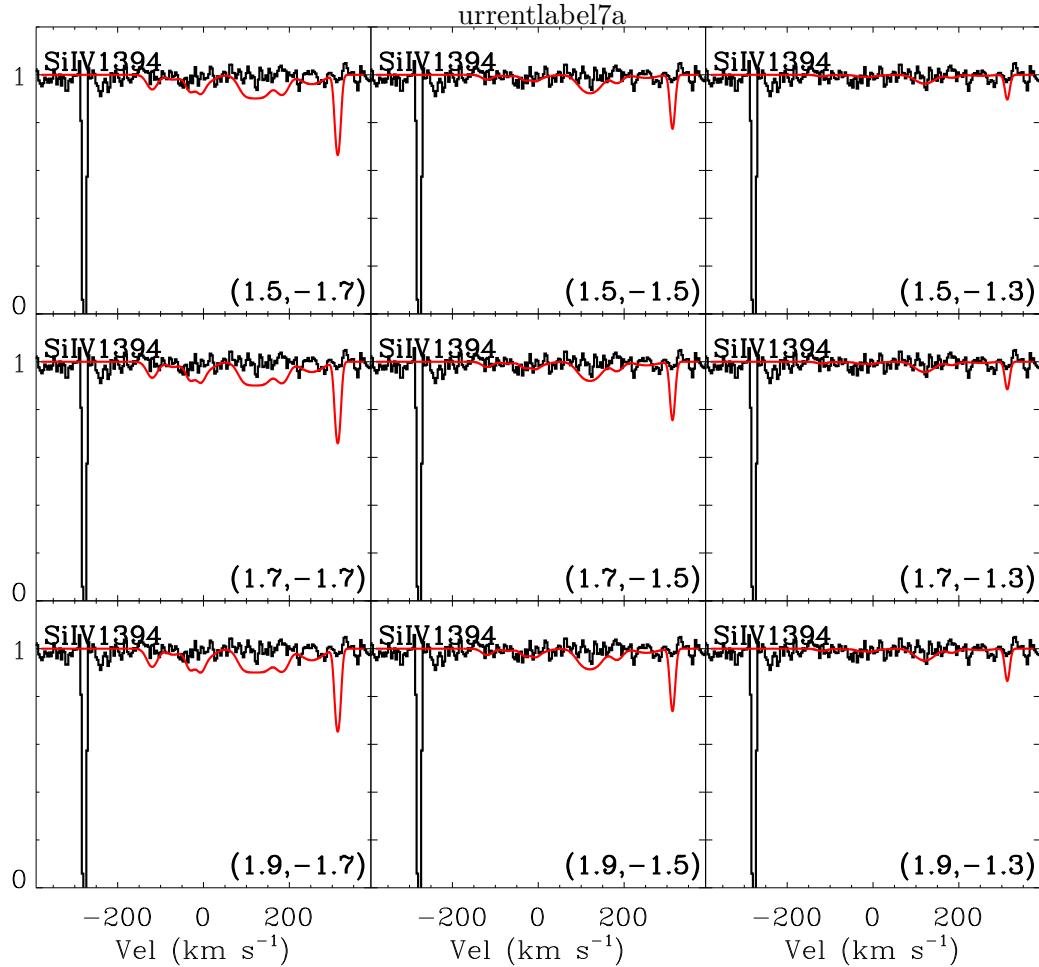


Fig. 7a Plots of observed and synthesized profiles of all kinematic components of SiIV  $\lambda 1396$  for the absorption system in the spectrum of HE0130–4021. We set  $\log(n_{\mathrm{H}}/\mathrm{cm}^{-3}) = 8$  and illustrate how the absorption profiles change as  $(\log(Z/Z_{\odot}), \log U)$  vary about  $(1.7, -1.5)$ .

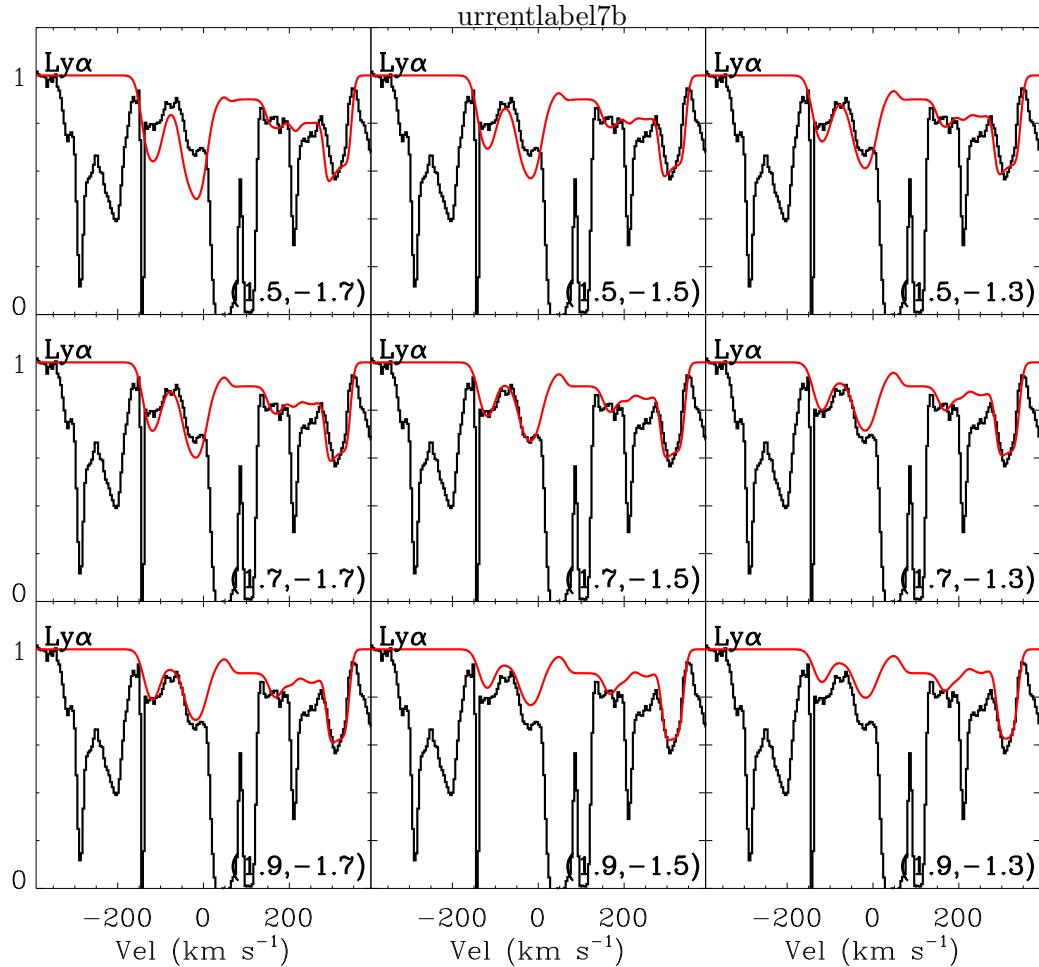


Fig. 7b Plots of observed and synthesized profiles of all the kinematic components of Ly $\alpha$  for the absorption system in the spectrum of HE0130–4021. We set  $\log(n_{\text{H}}/\text{cm}^{-3}) = 8$  and illustrate how the absorption profiles change as  $(\log(Z/Z_{\odot}), \log U)$  vary about  $(1.7, -1.5)$ .

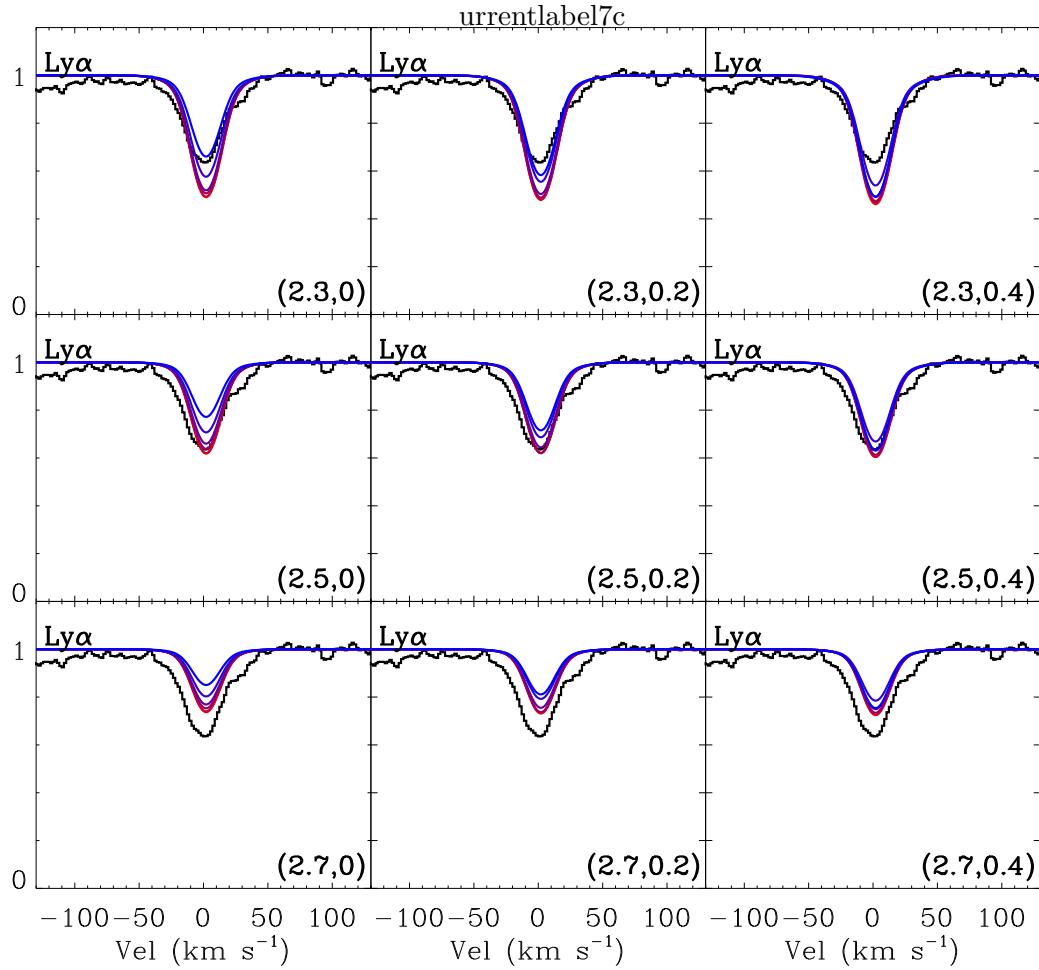


Fig. 7c Plots of observed and synthesized profiles of all the kinematic components of Ly $\alpha$  for the absorption system in the spectrum of Q1009+2956. The different panels show how the models change as  $(\log(Z/Z_{\odot}), \log U)$  vary about (2.5,0.2). In each panel, the model color changes gradually from red to blue for  $\log(n_{\text{H}}/\text{cm}^{-3}) = 2, 4, 6, 8, 10$ , and 12.

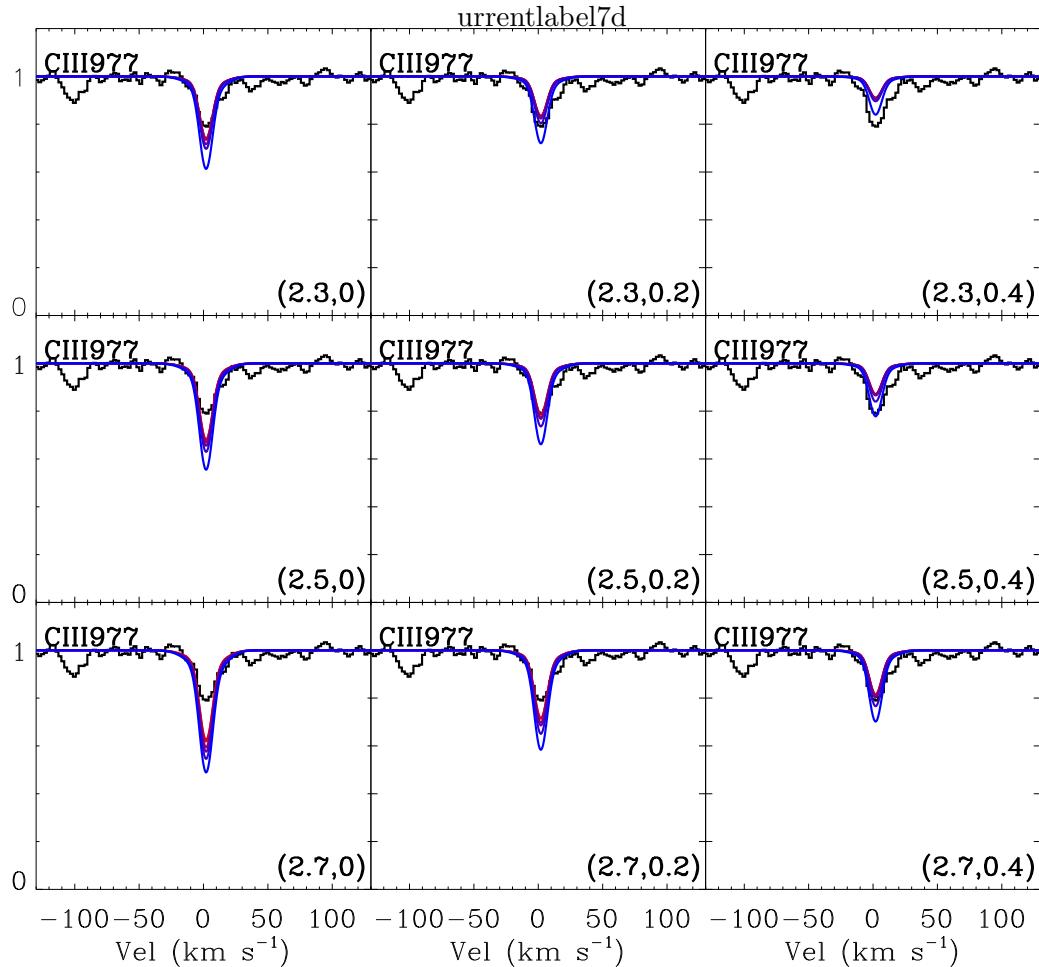


Fig. 7d Plots of observed and synthesized profiles of all the kinematic components of C III  $\lambda 977$  for the absorption system in the spectrum of Q1009+2956. The different panels show how the models change as  $(\log (Z/Z_{\odot}), \log U)$  vary about  $(2.5,0.2)$ . In each panel, the profiles are color-coded according to density in the same way as Fig. 7c.

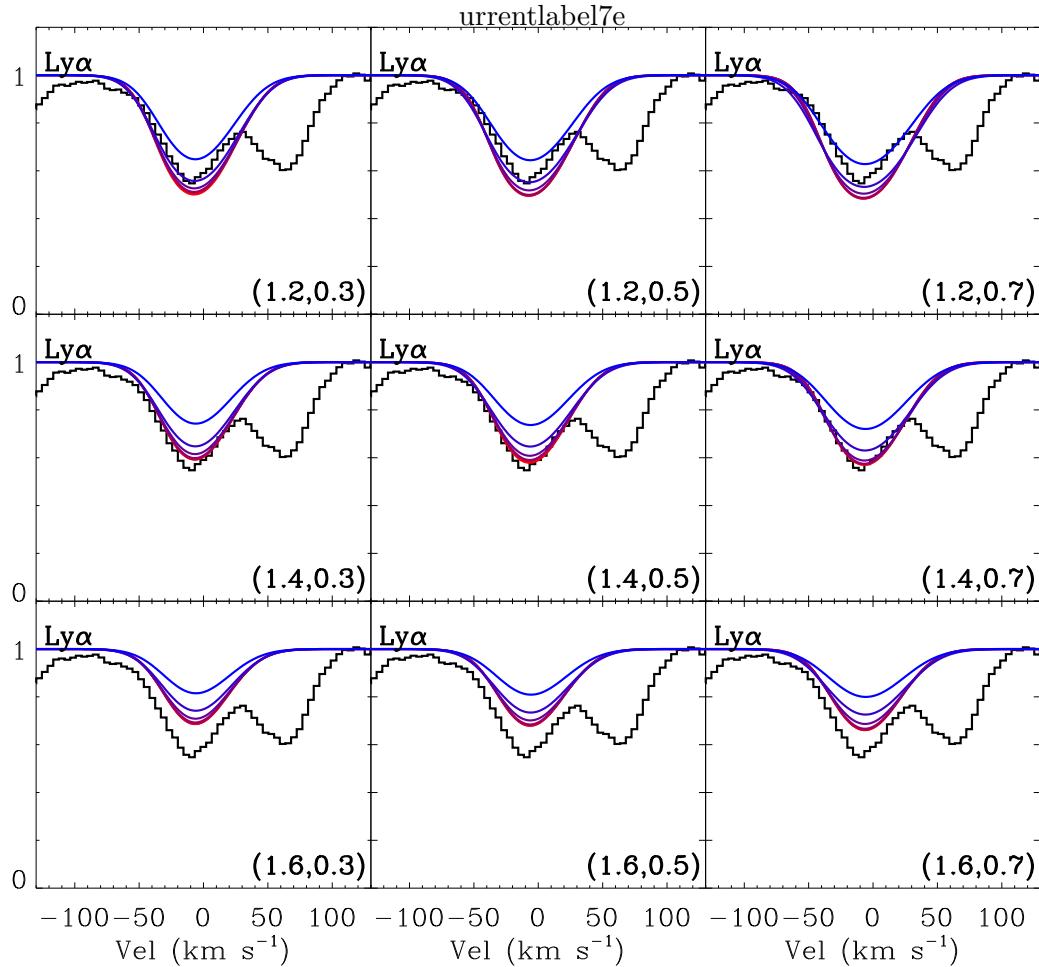


Fig. 7e Plots of observed and synthesized profiles of the two central kinematic components of Ly $\alpha$  for the absorption system in the spectrum of HS1700+6416. The different panels show how the models change as  $(\log(Z/Z_{\odot}), \log U)$  vary about  $(1.4, 0.5)$ . In each panel, the profiles are color-coded according to density in the same way as Fig. 7c. Figures 7a–7d, 7g, and 7h are available in the online version of the Journal.

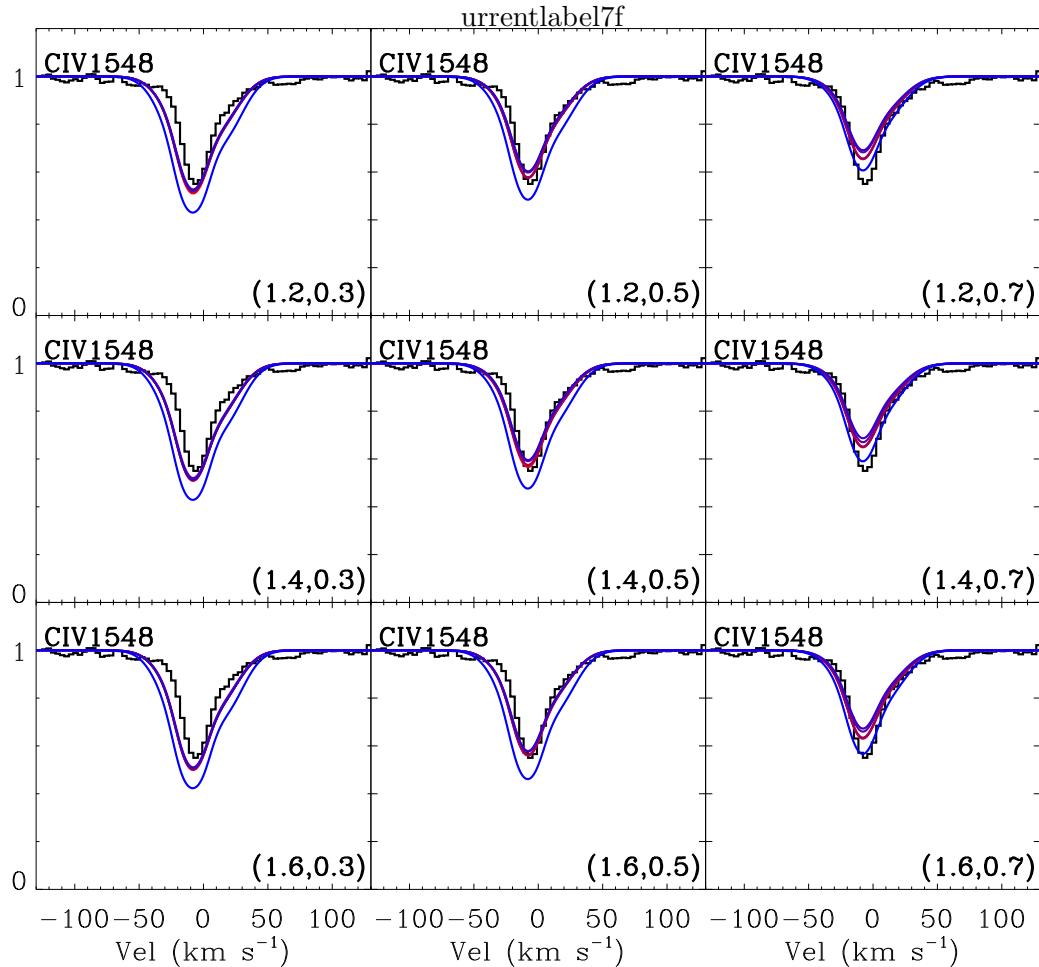


Fig. 7f Plots of observed and synthesized profiles of the two central kinematic components of C IV  $\lambda 1549$  for the absorption system in the spectrum of HS1700+6416. The different panels show how the models change as  $(\log (Z/Z_{\odot}), \log U)$  vary about  $(1.4, 0.5)$ . In each panel, the profiles are color-coded according to density in the same way as Fig. 7c.

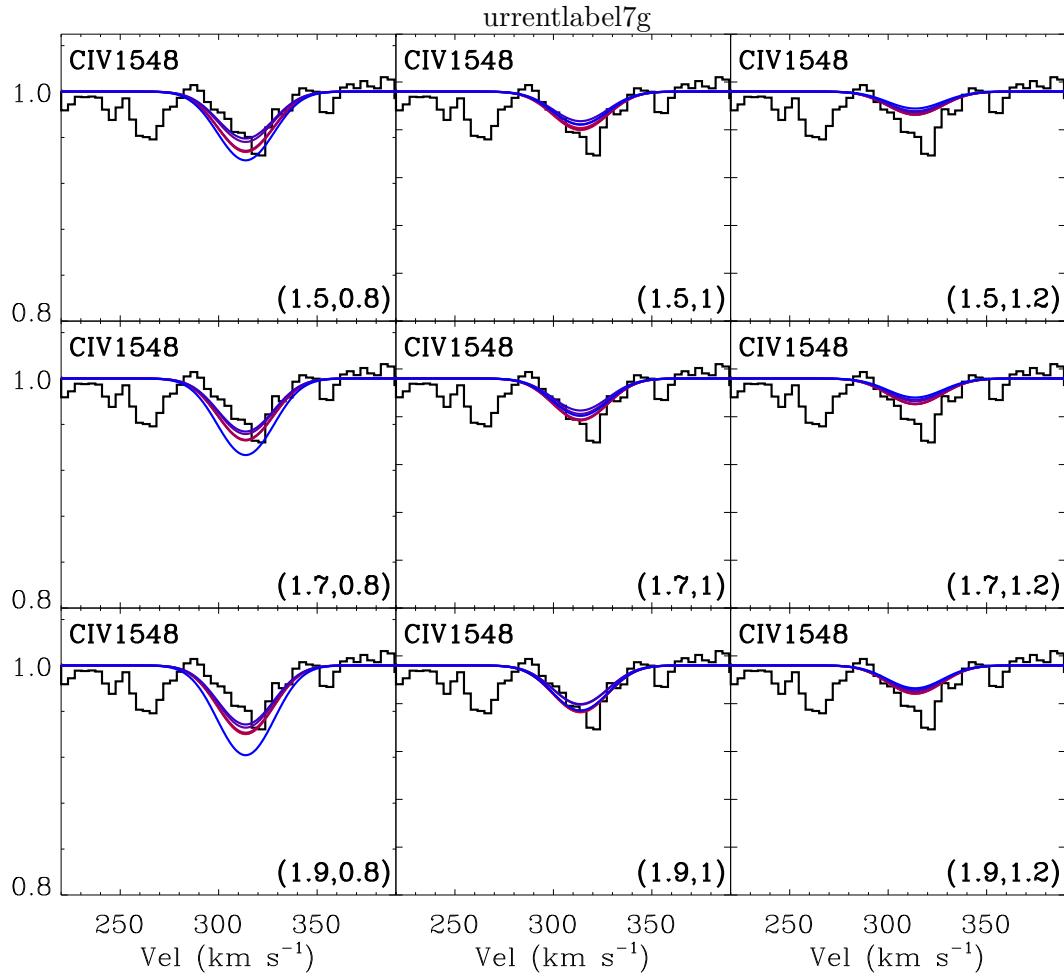


Fig. 7g Plots of observed and synthesized profiles of the kinematic component at  $315 \text{ km s}^{-1}$  of C IV  $\lambda 1549$  for the absorption system in the spectrum of HS1700+6416. The different panels show how the models change as  $(\log(Z/Z_{\odot}), \log U)$  vary about  $(1.7, 1.0)$ . In each panel, the profiles are color-coded according to density in the same way as Fig. 7c. The vertical axis scale is expanded in order to display this weak feature.

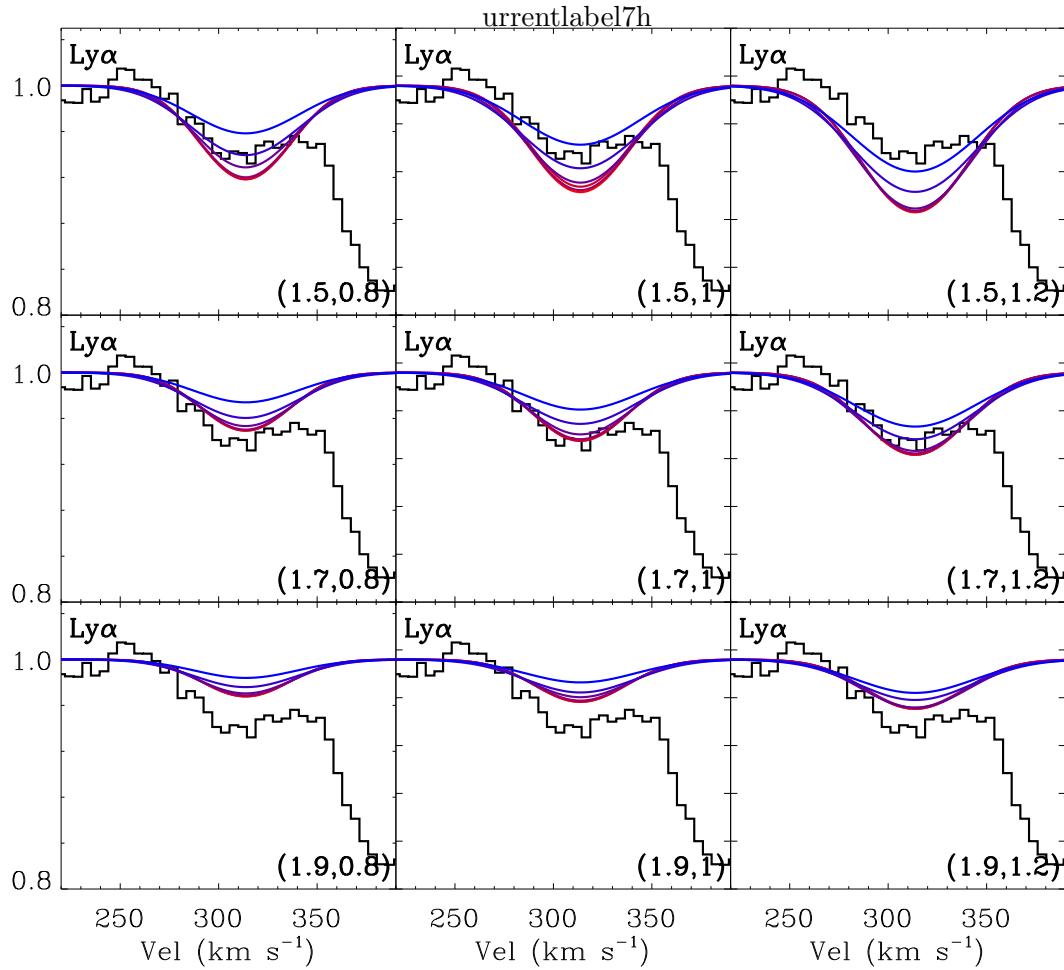


Fig. 7h Plots of observed synthesized profiles of the kinematic component at  $315 \text{ km s}^{-1}$  of Ly $\alpha$  for the absorption system in the spectrum of HS1700+6416. The velocity zero point is defined as in Fig. 5.1. In each panel, the profiles are color-coded according to density in the same way as Fig. 7c. The vertical axis scale is expanded in order to display this weak feature.

Table 5.1. Summary of Properties of Intrinsic NAL Systems <sup>a</sup>

Quasar	$z_{\text{em}}$	$d_L^{\text{b}}$ (Gpc)	$\log L_{\text{bol}}^{\text{c}}$	$\log Q_i^{\text{d}}$	$z_{\text{abs}}^{\text{e}}$	$v_{\text{ej}}^{\text{f}}$ (km s $^{-1}$ )	Spectral Coverage $^g$ (Å)	Transitions Detected
HE0130–4021	3.030	25.7	48.36	58.40	2.9739	4204	3612–6084	$\text{Ly}\alpha, \text{Ly}\beta, \text{C III}] \lambda 1908, \text{N V}, \text{O VI}$
Q1009+2956	2.644	21.8	48.49	58.53	2.6495	−453	3093–4623	$\text{Ly}\alpha, \text{Ly}\beta, \text{C III}] \lambda 1908, \text{S VI}, \text{N V}, \text{O VI}$
HS1700+6416	2.722	22.6	48.98	58.97	2.7125	764	3726–6190	$\text{Ly}\alpha, \text{Ly}\beta, \text{C IV}, \text{N V}, \text{O VI}$

<sup>a</sup>Adopted from Misawa et al. (2007).

<sup>b</sup>The luminosity distance, assuming  $\Omega_\Lambda = 0.7, \Omega_M = 0.3, h = 0.7$ .

<sup>c</sup>The log of the bolometric luminosity (in erg s $^{-1}$ ). See details in §5.3.2 of the text.

<sup>d</sup>The log of the ionizing photon rate (in s $^{-1}$ ). See details in §5.3.2 of the text.

<sup>e</sup>Absorption system redshift determined by finding the wavelength that divides the optical depth of the N V  $\lambda 1239$  line in half.

<sup>f</sup>Ejection velocity relative to emission lines. A positive sign denotes a blueshift.

<sup>g</sup>Wavelength range covered by the Keck spectrum.

Table 5.2. Results of Decomposition of Nv Doublet Profiles <sup>a</sup>

Quasar	Kinematic Component	$v_{\text{ej}}^{\text{b}}$ (km s <sup>-1</sup> )	$v_{\text{rel}}^{\text{c}}$ (km s <sup>-1</sup> )	$\log N_{\text{Nv}}/\text{cm}^{-2}$ <sup>d</sup>	$b^{\text{e}}$ (km s <sup>-1</sup> )	$C_f^{\text{f}}$
HE0130-4021	1	4324	-120	$14.18 \pm 0.04$	$16.6 \pm 0.9$	$0.43 \pm 0.03$
	2	4275	-71	$14.2 \pm 0.1$	$20 \pm 8$	$0.16 \pm 0.04$
	3	4236	-32	$13.9 \pm 0.1$	$13 \pm 2$	$0.49 \pm 0.08$
	4	4209	-5	$14.8 \pm 0.3$	$32 \pm 5$	$0.13 \pm 0.03$
	5	4084	+121	$15.9 \pm 0.5$	$28 \pm 5$	$0.10 \pm 0.02$
	6	4020	+184	$15.0 \pm 0.3$	$17 \pm 2$	$0.13 \pm 0.03$
	7	3891	+313	$15.2 \pm 0.6$	$6 \pm 1$	$0.37 \pm 0.02$
	8	3952	+252	$14.62 \pm 0.06$	$34 \pm 3$	$0.21 \pm 0.02$
	9	4210	-6	$13.51 \pm 0.01$	$12.6 \pm 0.3$	1.00 <sup>g</sup>
Q1009+2956	1	-453	-5	$13.70 \pm 0.29$	$14.3 \pm 1.7$	$0.62 \pm 0.11$
	2	-455	+2	$14.79 \pm 1.48$	$3.40 \pm 1.90$	$0.70 \pm 0.06$
	3	-461	+7	$14.63 \pm 0.28$	$13.0 \pm 2.1$	$0.22 \pm 0.05$
HS1700+6416	1	771	-9	$14.18 \pm 0.05$	$11.0 \pm 0.5$	$0.43 \pm 0.02$
	2	766	+2	$14.62 \pm 0.04$	$25.6 \pm 0.5$	$0.33 \pm 0.02$
	3	450	+314	$13.8 \pm 0.1$	$18.4 \pm 0.8$	$0.36^{+0.07}_{-0.06}$

<sup>a</sup>The errors are a combination of the continuum fitting errors and Voigt profile fitting errors.<sup>b</sup>Ejection velocity relative to the emission lines. A positive sign denotes a blueshift.<sup>c</sup>Velocity offset relative to the redshift of the entire absorption line complex. The sign of the velocity is set to be positive, if the lines are redshifted from the system center.<sup>d</sup>The Nv column density inferred from the Voigt profile fit.<sup>e</sup>Doppler parameter.<sup>f</sup>Effective coverage fraction.<sup>g</sup>There is no error bar for this entry because our fitting yields  $C_f > 1$ , which is unphysical. Thus, we set  $C_f = 1.00$  when we fit this component.

Table 5.3. Spectral Indices of Different SED Models Compared to Observed Values

SED <sup>a</sup>	$\alpha_o$	$\alpha_x$	$\alpha_{ox}$
Observed, HS1700+6416	-0.47	-1.00	-1.87
Elvis 1994	-0.91	-0.93	-1.37
Mathews 1987	-0.50	-0.70	-1.41
Small BB I	-1.24	-0.99	-1.40
Small BB II	-1.27	-0.99	-1.60
Small BB III	-1.27	-0.99	-1.90
Medium BB I	-0.54	-1.00	-1.40
Medium BB II	-0.56	-1.00	-1.60
Medium BB III	-0.56	-1.00	-1.90
Big BB I	-0.49	-1.82	-1.40
Big BB II	-0.50	-2.30	-1.60
Big BB III	-0.51	-3.06	-1.90

<sup>a</sup>The Roman numerals I,II,III represent different values of  $\alpha_{ox}$ .

Table 5.4. Summary of Preliminary Fitting Results

Quasar	Kinematic Component	$\log(Z/Z_{\odot})$	$\log U$	$\log(n_{\text{H}}/\text{cm}^{-3})$	Abundance Pattern
HE0130-4021 <sup>b</sup>	1-4,9	1.6	-0.6 to 0	2	solar
		1.6-1.7	-1.4 to 0	8	solar
		1.5-1.6	-1.4 to -0.3	10	solar
		1.5	-1.3	12	solar
		1.3	-1.0	14	solar
	5-8	$\geq 1.7$	$\geq -0.6$	2	solar
		1.9	-0.9 to -0.2	8	solar
		$\geq 1.8$	-0.8 to 0.4	10	solar
		1.6	-0.8	12	solar
		1.8	-0.6	14	solar
Q1009+2956	1-3	2.5	0.2	2-8	solar, [S/H] $\leq 1.5$
		2.5	0.3	9	solar, [S/H] $\leq 1.5$
		2.4	0.2	10	solar, [S/H] $\leq 1.4$
		2.3	0.3	11-12	solar, [S/H] $\leq 1.5$
		2.3	0.7	12	solar, [C/H] $\sim 2.8$
		2.5	0.4	13	solar, [S/H] $\leq 1.5$
HS1700+6416 <sup>a</sup>	1-2	1.4	0.5	2-8	solar
		1.3	0.4	9	solar
		1.3	0.5	10	solar
	3	1.7	1.0	2-8	solar

Note. — Here we assume that blended kinematic components, as indicated in the second column of the table, are described by the same combination of model parameters.

<sup>a</sup>The results presented in this table are also a summary of the final results for this quasar. See the discussion in §5.4.1 of the text.

Table 5.5. Best model parameters for the Q1009+2956 absorption systems

Kinematic Component	$\log (Z/Z_{\odot})$	$\log U$	$\log(n_{\text{H}}/\text{cm}^{-3})$	Abundance Pattern	$r^{\text{a}}$ (pc)
1	2.4	0.0	2-8	solar, [S/H]=1.3	
2	2.6	0.3		solar, [S/H]=1.5	10–10,000
3	2.3	0.0		solar, [S/H]=1.4	
1	2.3	0.1	9	solar, [S/H]=1.3	
2	2.5	0.3		solar, [S/H]=1.5	~ 0.2
3	2.7	0.1		solar, [S/H]=1.7	
1	2.4	0.0	10	solar, [S/H]=1.4	
2	2.5	0.3		solar, [S/H]=1.5	~ 0.8
3	2.3	0.0		solar, [S/H]=1.3	
1	2.2	0.2	11-12	solar, [S/H]=1.2	
2	2.4	0.3		solar, [S/H]=1.4	0.04–0.2
3	2.4	0.5		solar, [S/H]=1.4	
1	2.2	0.5	12	solar, [C/H]=2.7	
2	2.4	0.8		solar, [C/H]=2.9	~ 0.04
3	2.5	0.5		solar, [C/H]=3.0	
1	2.7	0.3	13	solar, [S/H]=1.7	
2	2.6	0.5		solar, [S/H]=1.6	~ 0.004
3	2.6	0.4		solar, [S/H]=1.6	

<sup>a</sup>The distance of the absorbing gas from the ionizing source for this combination of  $U$  and  $n_{\text{H}}$ .

Table 5.6. Best model parameters for the HS1700+6416 absorption systems

Kinematic Component	$\log (Z/Z_{\odot})$	$\log U$	$\log(n_{\text{H}}/\text{cm}^{-3})$	Abundance Pattern	$r^{\text{a}}$ (pc)
1	1.3	0.5	2–8	solar	9–9000
	1.6	0.6		solar	
1	1.2	0.5	10	solar	$\sim 1$
	1.5	0.6		solar	
3	1.7	1.0	2–8	solar	5–5000

<sup>a</sup>The distance of the absorbing gas from the ionizing source for this combination of  $U$  and  $n_{\text{H}}$ .

Table 5.7. Constraints on Continuum and BELR Coverage Fractions

Quasar	Ion	$W^a$	$C_c$	$C_e$
HE0130–4021	N V	$1.1 \pm 0.4$	unconstrained	...
			$< 0.5$	...
			unconstrained	...
			$< 0.4$	...
			$< 0.3$	...
			$< 0.4$	...
			unconstrained	...
			$< 0.6$	...
			$> 0.3$	...
			unconstrained	unconstrained
Q1009+2956	N V	$0.66 \pm 0.12$	unconstrained	...
			$> 0.2$	...
			$< 0.5$	...
			unconstrained	unconstrained
HS1700+6416	N V	$0.96 \pm 0.06$	$0.2\text{--}0.8$	...
			$< 0.7$	...
			$< 0.8$	...
	Ly $\alpha$ ( $C_f = 0.4$ ) <sup>b</sup>	$0.6 \pm 0.2$	$< 0.7$	$> 0.2$
	C IV	$0.41 \pm 0.02$	$0.2\text{--}0.8$	$> 0.4$

<sup>a</sup>The values of  $W$  were obtained from the following spectra: for HE0130–4021 we used the spectrum published by Osmer & Smith (1976), for Q1009+2956 we used the spectrum published by Burles & Tytler (1998), and for HS1700+6416 we used the SDSS DR5 spectrum.

<sup>b</sup>In the case of Ly $\alpha$  we cannot determine  $C_f$  from the data, therefore we assign a value equal to the average of all the N V kinematic components.

## Chapter 6

### Summary of Thesis Results

In this thesis, we constructed a set of 149 composite quasar spectra from SDSS DR3 quasar catalog using the spectral binning, averaging and weighting techniques in Vanden Berk et al. (2001). The composite spectrum set covers a wide range of redshift with  $0 \lesssim z \lesssim 5$  and wide range of luminosity with  $40.5 \lesssim l_\lambda(2200 \text{ \AA}) \lesssim 43.5$ . Using this spectrum, we studies the general trend of UV spectral shape. For quasars at a similar redshift, the ones with lower luminosity tend to be redder than the ones with higher luminosity. For quasars with a similar luminosity, the ones at higher redshift tend to be redder than the ones at lower redshift. The Baldwin Effect is an anti-correlation between the emission line EW and UV luminosity. Using the composite spectrum set, we found that while the slopes of the Baldwin Effect do not change over redshift, the amplitudes exhibit a significant trend over redshift. At the certain luminosity, the EW of C IV  $\lambda 1549$  may change 0.2 dex from  $z \sim 1.4$  to  $z \sim 4.5$ ; the EW of Mg II  $\lambda 2798$  may change 0.3 dex from  $z \sim 0.4$  to  $z \sim 2.2$ . The redshift dependence of this correlation is an important contributor to the scatter of the Baldwin Effect of a quasar sample covering a wide range of redshifts. There is a monotonic trend of the mean emission line shifts with luminosity for several emission lines such as H I  $\lambda 5877$ , C IV  $\lambda 1549$ , Al III  $\lambda 1857$ , C III]  $\lambda 1908$ , and Mg II  $\lambda 2798$ . The emission line width, parameterized by FWHM is positively correlated to luminosity. Similar to the Baldwin Effect, the amplitude of this correlation also depends on redshift. At a given luminosity, the FWHM of C IV  $\lambda 1549$  trends to be lower at higher redshifts (see Chapter 2).

The C IV Baldwin Effect is driven by both  $\alpha_{\text{ox}}$  and  $l_\nu(2500 \text{ \AA})$  or equivalently, by  $l_\nu(2 \text{ keV})$  and  $l_\nu(2500 \text{ \AA})$ . This result implies that changes in the ionizing flux induce changes in the ionization state of the BELR, producing more C IV ions when the SED becomes harder, and vice versa. No correlation is found between EW(Fe K $\alpha$ ) and either  $\alpha_{\text{ox}}$  or EW(C IV), implying that Fe K $\alpha$  is not likely to have the same origin as C IV. Most of the scatter of the Baldwin Effect as well as the EW(C IV)- $\alpha_{\text{ox}}$  relation is contributed by variability (Chapter 3).

We have constructed a quasar catalog with simultaneous UV, optical and X-ray observations from *Swift*. This catalog contains 843 quasars, among which 637 have both UVOT and XRT observations and 354 are detected by both UVOT and XRT. Using a sample selected from this catalog, we found that the dispersion of the  $\alpha_{\text{ox}}-l_\nu(2500 \text{ \AA})$  relation can be reduced by  $\sim 10\%$  (Chapter 4).

By modeling the intrinsic NAL systems in the spectra of three quasars, we found that they processed super-solar metallicity of  $\log(Z/Z_\odot) \gtrsim 2$ , regardless of constraints on ionization parameter  $U$  and hydrogen volume density  $n_{\text{H}}$ , implying intensive star formation before quasars become observable. We find a high ionization parameter,

$\log U \sim 0$ , consistent over three absorption systems. By constraining the coverage fraction of the continuum source and the BELR separately, we found that the continuum source is likely to be partially covered by absorbers. We conclude that the accretion rate  $\dot{M} \lesssim \text{few } M_{\odot} \text{ yr}^{-1}$ , and the corresponding power is a negligible fraction of the electromagnetic luminosity. The O VI  $\lambda\lambda 1032, 1038$  doublet is underproduced by the models that reproduce the other absorption lines, suggesting that the O VI lines are either contaminated by Lyman forest or they arise in a separate gas phase with a larger coverage fraction (Chapter 5).

We verified the following properties and correlations, which were previously generally accepted.

- In general, the emission line shift is correlated with the ionization potential; emission lines produced by high ionization potential ions are more blue shifted than lines from low ionization ions. (Chapter 2)
- The Baldwin Effect does not only exist for C IV  $\lambda 1549$  but also for other broad emission lines such as Ly $\alpha$ , Si IV  $\lambda 1396$ , He II  $\lambda 1640$ , and Mg II  $\lambda 2798$ , but this trend is not present for N V  $\lambda 1239$ , C III]  $\lambda 1908$ , and H $\beta$  (Chapter 2).
- The  $\alpha_{\text{ox}} - l_{\nu}(2500 \text{ \AA})$  correlation is seen in the samples we constructed to study the C IV Baldwin Effect and the Cleaned sample we constructed from the *Swift* quasar catalog (Chapter 4).

Again, this work is intended to provide observational constraints to quasar models. Although a general super massive black hole – accretion disk model has been established for many years, detailed structure and physical conditions of each component are still not well understood, such as the volume density of the BELR and the outflow, the exact locations of the broad emission lines and the states of the accretion disk which controls the energy distribution shape as well as the variability. In addition, models should explain the redshift-dependent properties of quasars. A possible solution to these problems could be building a large code to numerically modeling the geography, dynamics, radiative transfer, ionization and chemical compositions of the central engine. The physical conditions must be tuned so that the output emission matches the observational property constraints.

Observationally, to obtain further constraints on AGN physics, it is not only important to continue performing many wavelengths and programs of long term monitoring to study the variability, which are directly informative on the central black mass (e.g., Peterson et al. 2004; Denney et al. 2009, 2010) and outflow dynamics (e.g., Misawa et al. 2005). In addition, larger ground based and space telescopes are critical to spatially resolve the BELR and even the accretion disk. The ongoing Thirty-Meter-Telescope project (Trimble & Barton 2007) and Large Synoptic Survey Telescope (LSST), could be very helpful on this. Although *Chandra* and *XMM-Newton* have produced tremendous successes since launch (see a review by Brandt & Hasinger (2005)), there may not be large X-ray missions in the next decade except the Astro-H (Takahashi et al. 2010).

## Appendix A

# Fitting the Continuum of a Composite SDSS Quasar Spectrum Using CMA-ES

### A.1 Introduction

This appendix describes an algorithm to fit a composite SDSS quasar spectrum using CMA-ES. Although this algorithm is different from the one used in our thesis, it has some advantages and produces better results.

Described in Chapter 2, we have constructed a set of 161 composite quasar spectra binned in redshift and luminosity space from  $\sim 80,000$  typical SDSS quasars. Spectra in the same bin are normalized and averaged, so each composite spectrum represents the *average* properties of the spectra in each bin. This is the first time this technique has been used to study the evolution over a wide range of luminosity ( $38.25 \lesssim \log l_\lambda(2200 \text{ \AA}) \lesssim 44.00$ ) and redshift ( $0 \lesssim z \lesssim 5$ ).

We have tried two types of algorithms to derive a set of measurements from the composite spectra.

- The Levenberg-Marquardt algorithm (LMA) is designed to solve multivariate least squares curve fitting problem (More & Wright 1993). It interpolates between the Gauss-Newton algorithm (GNA) (Björck 1996) and the method of gradient descent (Avriel 1993), but it is more robust than the GNA, which means that in many cases, LMA finds a solution even if it starts far from the final minimum. The fitting results using this algorithm are displayed in Figure A.1. This method has some limitations
  - We found that to solve a problem with more than 10 free parameters, this algorithm is sensitive to initial values of parameters. We must provided an initial guess of the power-law (PL) component, and this estimate must be as close to the optimal value as possible.
  - This algorithm does not work well for the iron templates, which do not have analytical expressions. The features where the iron emission and the small blue bump (SBB) cannot be fit well simultaneously (see the over-production of model flux around 3700 Å). In addition, we could not use this method to find the proper velocity dispersions of the iron template because the available dispersion values are discrete.
- Genetic algorithm (GA). GA is a search heuristic that mimics the process of natural evolution (Banzhaf et al. 1998). This heuristic is routinely used to generate useful solutions to optimization and search problems. A typical GA algorithm consists

of initialization, selection, reproduction and termination. The solutions asymptotically converge to an optimal value under a given criteria. Vanden Berk wrote a program (not published) using this algorithm to fit the SDSS quasar spectra using GA, but this program has two disadvantages

- Instead of using the iron emission templates, each low ionization iron emission line is simulated using one or multiple Gaussian profiles. This requires over a hundred free parameters, which significantly slows down the fitting process. GA requires  $\sim 5 - 10$  minutes (a typical desktop computer with a 2.0 GHz processor) to fit a composite quasar spectrum.
- Although the program can produce a nearly perfect fit, the result is unstable. There are a number of parameter combinations that can produce equally good fits and it is difficult to determine which is the physical solution.

The limitations and disadvantages of these two methods motivate us develop an evolutionary algorithm to fit the quasar spectrum. The CMA-ES is an evolutionary algorithm for difficult non-linear non-convex optimization problems in a continuous domain (Hansen & Kern 2004). It is a second order approach to estimate a positive definite matrix within an iterative procedure (the covariance matrix). This approach makes the method feasible on non-separable and/or ill conditioned problems. Because this method does not require gradients, it is feasible on non-smooth and even non-continuous problems, as well as “noisy” problems. Previous results using LMA imply that the space of the objective function is not smooth as the solution varies depending on the initial guess (Hansen & Ostermeier 2001). The CMA-ES approach is can overcome this problem.

The paper is outlined in the following way. In Section A.2, we describe the problem in detail; in Section A.3, we present the solution strategy in terms of each step applied in CMA-ES; in Section A.4, we briefly describe some important issues that arose when developing the program; in Section A.5, we present our optimization results; and finally in Section A.6, we discuss the limitations of our approach and potential extension of this work.

## A.2 Problem Description

We are aiming at fitting the underlying continuum of a composite quasar spectrum at  $z = 0.738$ , and  $\log l_\nu(2500 \text{ \AA})$  (see Figure A.1). This continuum is fit by four components

1. A power-law (PL) continuum, which takes the form  $f_\lambda = 10^\beta \lambda^\alpha$ . This component has two free parameters: the spectral index  $\alpha$  and the scaling factor  $\beta$ . In principle, the variables do not have any constraints but previous studies have found that  $\alpha \sim -1.6$  and  $\beta \sim 6-7$ .
2. A small blue bump (SBB), which takes the form

$$f_\lambda = \begin{cases} A_B \cdot \frac{1-e^{\tau_B}}{e^{hc/(k\lambda_B T_B)}-1} \cdot e^{-(\lambda-\lambda_B)/\Delta\lambda}, & \lambda \geq \lambda_B \\ A_B \cdot \frac{1-e^\tau}{e^{hc/(k\lambda T_B)}-1} \cdot \left(\frac{\lambda}{\lambda_B}\right)^{-5}, & \lambda \leq \lambda_B, \tau = \tau_B \left(\frac{\lambda}{\lambda_B}\right)^3 \end{cases} \quad (\text{A.1})$$

This component has three free parameters: Balmer temperature  $T_B$ , Balmer optical depth  $\tau_B$ , and scale factor  $A_B$ . The scale factor should be positive; the other two parameters do not have constraints. Empirically,  $\tau_B \sim 1$ , and  $T_B \sim 10^4$ . Both  $\lambda$  and  $\lambda_{rmB}$  are in unit of Å and  $\Delta\lambda = 100$  Å.

3. UV iron emission forest. This is a blend of many low ionization iron emission lines ranging from  $\sim 1075$  Å to  $3090$  Å. This component has three free parameters, the scale factor  $A_U$ , velocity dispersion  $v_U$  and relative shift  $\Delta z_U$ <sup>1</sup>. The scale factor should be positive. The velocity dispersion can only be one of  $1500, 2000, \dots, 9500$  km s<sup>-1</sup> (set by template grid). In principle,  $\Delta z_U$  can be any value but typically  $|\Delta z_U| \leq 0.005$ .
4. Optical iron emission forest. This is a blend of many low ionization iron emission lines ranging from  $\sim 3535$  Å to  $\sim 7534$  Å. Similar to the UV template, this component also has three free parameters, the scale factor  $A_O$ , velocity dispersion  $v_O$  and relative shift  $\Delta z_O$ . Again,  $v_O$  can only be one of  $1500, 2000, \dots, 9500$  km s<sup>-1</sup>, and  $|\Delta z_O| \leq 0.005$ .

A summary of the free parameters and their permitted values are listed in Table A.1. We use the *reduced*  $\chi^2$  value as the objective function:

$$\chi^2 = \frac{1}{N - n - 1} \sum_{i=1}^N \frac{(f_i - f_{e,i})^2}{\sigma_i^2}$$

in which  $N$  is the number of benchmark wavelength points, which is used to compare the model and the observed fluxes;  $n$  is the number of variables;  $f_i$  is the interpolated observed flux value at wavelength point  $i$ ;  $f_{e,i}$  is the expected (calculated) flux value at wavelength point  $i$ , and  $\sigma_i$  is the interpolated observational uncertainty at wavelength point  $i$ . In this problem, we have  $N = 50$  benchmark points and  $n = 11$  variables.

### A.3 Solution Strategy

We apply the CMA-ES as the method to solve this problem. This algorithm involves sampling, selection, cumulation and updating operations (Hansen & Ostermeirer 1996). The problem is initialized as Table A.2 and each parameter shown below such as  $\mathbf{x}$ ,  $\mathbf{m}$  and  $\mathbf{z}$  are also described in Table A.2. Each generation loop consists the following steps.

1. Generate and evaluate  $\lambda$  offspring

$$\begin{aligned} \mathbf{x}_i &= \mathbf{m} + \sigma \mathbf{z}_i, \mathbf{z}_i \sim \mathcal{N}_i(\mathbf{0}, \mathbf{C}), \quad i = 1, 2, \dots, \lambda \\ f_i &= \chi^2(\mathbf{x}_i) \end{aligned}$$

---

<sup>1</sup>This is the shift of the template *with respect to* the quasar spectrum itself, not the redshift of the quasar relative to the observer.

Vector  $\mathbf{m}$  has  $N$  elements, and  $\mathbf{C}$  is an  $N \times N$  matrix (the covariant matrix). In this step, we apply a *death penalty* to the two velocity dispersion values  $v_U$  and  $v_O$ . If they fail to fall into the range of [1500, 9500], we simply discard this value and re-draw a new set of random numbers. However, if the number of failures is greater than 10, we stop drawing and adopt the boundary value. When evaluating the fitness function, we apply a *penalty term*  $f_p(\mathbf{x})$  to the fitness function based on the value of redshift  $\Delta z_U$  and  $\Delta z_O$ , so that

$$\begin{aligned}\tilde{f}(\mathbf{x}_i) &= f(\mathbf{x}_i) + f_p(\mathbf{x}_i) \\ f_p(\mathbf{x}_i) &= (C \cdot t)^2 \cdot G(\mathbf{x}_i) \\ G(\mathbf{x}_i) &= \sum_{k=1}^4 \max[0, 5 \times g_k(\mathbf{x}_i)]^2\end{aligned}$$

The  $g_k(\mathbf{x}_i)$  are the constrained items [ $g_k(\mathbf{x}_i) \leq 0$ ]:

$$\begin{aligned}g_1 &= \Delta z_U - 0.005 \\ g_2 &= -\Delta z_U - 0.005 \\ g_3 &= \Delta z_O - 0.005 \\ g_4 &= -\Delta z_O - 0.005\end{aligned}$$

In the equations above,  $C = 0.5$  and,  $t$  is the generation number.

2. Sort offspring by the penalized function  $\tilde{f}(\mathbf{x}_i)$  and compute the weighted mean. Although we sort the offspring using the penalized function, we still output the unpenalized function  $f_p(\mathbf{x}_i)$ , because it reflects the *real* goodness-of-fit regardless of penalty term. In addition,  $\tilde{f}(\mathbf{x}_i)$  involves the generation number  $t$ , so even if the best  $\tilde{f}(\mathbf{x}_i)$  decreases in the first few generations, will increase after a certain point. After sorting the offspring ascendingly, we select the first  $\mu$  offspring as the parents to calculate the weighted mean. The weighted mean after selecting  $\mu$  offspring is

$$m_j^s = \sum_{i=1}^{\mu} w_i x_{ij}^s, \quad j = 1, 2, \dots, N$$

Note that  $\mathbf{x}^s$  is a matrix of  $\mu \times N$  and  $\mathbf{x}^s$  is selected from  $\mathbf{x}$  based on the fitness function so  $x_{0j}, j = 1, 2, \dots, N$  contains the best offspring and thus has the heaviest weight  $w_0$  (see appendix).

3. Cumulation: update the evolution paths. Conceptually, the evolution path is the path the strategy takes over a number of generation steps, which can be expressed as a sum of consecutive steps of the (weighted) mean  $\mathbf{m}$ . To accomplish this task, we iterate two path vectors,  $\mathbf{p}_\sigma$  and  $\mathbf{p}_c$ , which represent the path of  $\sigma$  and covariance matrix  $\mathbf{C}$ . Both  $\mathbf{p}_\sigma$  or  $\mathbf{p}_c$  has two terms. The first term is the vector itself in the last iteration multiplied by a decay factor. This term causes the vector norm to decrease. The second term is based on the weighted mean of the selected offspring with a normalization factor and directs the evolution path to the optimal

value. This step is a preparation to the covariance matrix adaption and step size adaption in the succeeding steps. The equations used to calculate these paths and the updates of evolution paths  $\mathbf{p}_\sigma$  and  $\mathbf{p}_c$ , are shown below.

$$\begin{aligned} p_j^\sigma &= (1 - c_s)p_j^\sigma + \sqrt{c_s(2 - c_s)\mu_{\text{eff}}}\sum_{k=1}^N C_{jk}^{-1/2}z_k \\ z_k &= \frac{m_k^s - m_k}{\sigma}, \quad j = 1, 2, \dots, N \\ p_j^c &= (1 - c_c)p_j^c + h^\sigma \sqrt{c_c(2 - c_c)\mu_{\text{eff}}}z_j \\ z_j &= \frac{m_j^s - m_j}{\sigma}, \quad j = 1, 2, \dots, N \\ h^\sigma &= \begin{cases} 1, & \frac{\|\mathbf{p}_c\|}{N\sqrt{1-(1-c_s)^2t/\lambda}} < 1.4 + \frac{2}{1+N} \\ 0, & \frac{\|\mathbf{p}_c\|}{N\sqrt{1-(1-c_s)^2t/\lambda}} \geq 1.4 + \frac{2}{1+N} \end{cases} \end{aligned}$$

in which  $C_{jk}^{-1/2}$  represents the array element in the inverse square root of matrix  $\mathbf{C}$ .

4. Adapt covariance matrix  $\mathbf{C}$ . This is the essential part of the algorithm and the reason why it is named as covariance-matrix adaption evolutionary algorithm. The adaption equation includes three terms. The first term is the covariance matrix itself multiplied by a decay factor. Again, this term decreases the norm of the covariance matrix. The second term is *rank one update* plus a minor correction based on the evolutionary path  $\mathbf{p}_c$ ; This term can learn straight ridges in  $O(n)$  rather than  $O(n^2)$  function evaluations. The third term is *rank  $\mu$  update* which is based on the weighted cross product of the selected offspring. The rank  $\mu$  update increases the possible learning rate in large populations and also reduces the number of necessary generations roughly from  $O(n^2)$  to  $O(n)$  Hansen & Koumoutsakos (2003). Therefore, the rank  $\mu$  update is the primary mechanism whenever a large population size is used (say  $\lambda > 3n + 10$ ). The quations for adaption of covariance matrix  $\mathbf{C}$  is

$$C_{jk} = (1 - c_1 - c_\mu)C_{jk} + c_1 \left[ p_j^c p_k^c + (1 - h^\sigma)c_c(2 - c_c)C_{ij} \right] + \sum_{k=0}^N c_\mu T_{jp} W_{pp} T_{pk}$$

in which  $\mathbf{W} = \text{diag}(w_1, w_2, \dots, w_\mu)$ ,  $T_{jq} = x_{jq}^s - M_{jq}$ ,  $q = 1, 2, \dots, \mu$ , and  $\mathbf{M}$  is an array of which each column is  $\mathbf{m}$ . Note that  $\mathbf{x}^s$  represents the *selected* and *sorted* offspring.

5. Finally, we update the step size  $\sigma$ . The reason for a step control is that the covariance matrix update can hardly increase the variance in *all* directions simultaneously. In other words, the *overall* scale of search, the *global step-size* cannot be increased effectively (Andreas Ostermeier 1995). In our investigation, we apply

the *path length control* (cumulative step-size adaption) as recommended by Beyer & Arnold (2006). The equation for adaption of step size  $\sigma$  is.

$$\sigma = \sigma \exp \left[ \frac{c_s}{d_\sigma} \left( \frac{\|\mathbf{p}_s\|}{\|\mathcal{N}(\mathbf{0}, \mathbf{I})\|} \right) - 1 \right].$$

6. Correct  $\mathbf{C}$ . To ensure that  $\mathbf{C}$  is a symmetric matrix, we enforce symmetry by replacing the lower triangle elements with the upper triangle elements.
7. We apply three criteria to terminate the generation loop
  - (a) The best fitness is below the pre-defined value.
  - (b) The number of generations is greater than the maximum number of generations allowed ( $10^4$ ).
  - (c) The change of the best fitness value is smaller than 0.05 for 20 consecutive generations.
  - (d) The change of the best fitness value is positive for 20 consecutive generations. Theoretically, this could never happen, but it exists when the objective function approaches the optimal value.

#### A.4 Program Design

The coding design basically follows the flow of the loop in Section A.3; here we briefly mention some special aspects that arose to when developing the code.

- To avoid reading the iron emission templates from hard drive each time we evaluate the fitness function, we upload template spectra of *all* the velocity dispersions before the generation loop. This requires a few seconds of some overhead computational time, but it significantly reduces the computational time required to evaluate the objective function. In each generation, we must evaluate the objective function for at least 220 times (in the default situation).
- The available values of  $v_U$  and  $v_O$  are not continuous. As a result, we must perform a discretization to make sure that they are integer multiples of  $500 \text{ km s}^{-1}$  before evaluating the fitness function.
- To compute  $\mathbf{C}^{-1/2}$ , we first must find the eigenvectors of  $\mathbf{C}$ , and construct an array  $\mathbf{P}$  whose column vectors are these eigenvectors. Given that  $\mathbf{C}$  is *symmetric and positive deterministic*, it can be diagonalized as  $\mathbf{D} = \mathbf{P}^{-1}\mathbf{C}\mathbf{P}$  whose diagonal elements are eigenvalues of  $\mathbf{C}$ . Therefore,  $\mathbf{C}^{-1/2} = \mathbf{P} \left( \mathbf{D}^{1/2} \right)^{-1} \mathbf{P}^{-1}$ .
- It is recommended to use the eigenvalues to construct the diagonal matrix  $\mathbf{D}$ , instead of obtaining  $\mathbf{D}$  using matrix multiplication, *i.e.*,  $\mathbf{D} = \mathbf{P}^{-1}\mathbf{C}\mathbf{P}$ . This is simply because of the limited precision of computers so that after a series of computations, the number zero is expressed in terms of a small number such as  $1.1\text{E-}11$ , but after a number of iterations, this number may accumulate, increase and eventually destroy the computation.

- Because the termination criteria mentioned above, the last generation is not necessarily the best. As a result, after terminating the loop, the program reads the entire output file and selects the generation with the minimum fitness value as the final best result.

## A.5 Optimization Results

### A.5.1 Testing Program

Instead of using an existing code, we wrote our own code because the objective function evaluation cannot be performed separately from the optimization code. The code is written in interactive data language (IDL) 8.0 Mac version and run on a Mac Pro with two 2.66 GHz Dual-Core Intel Xeon processors and 2 GB 667 MHz memory. Before using the code to fit the quasar spectrum, we perform a series of tests by optimizing the generalized Rosenbrock function (Rosenbrock 1960):

$$f(\mathbf{x}) = \sum_{i=1}^{n-1} \left[ 100 \left( x_i^2 - x_{i+1} \right)^2 + (x_i - 1)^2 \right]$$

The global minimum of this function occurs when each  $x_i$  equals 1 yielding  $f(\mathbf{x}) = 0$ . The global minimum is inside a long, narrow parabolic-shaped flat valley. This function is widely used as a test function because to find the valley is trivial but to converge to the global minimum is difficult, e.g., Storn & Price (1997). The goal of this test is to prove that this code is working and achieve the desired performance. We terminate the generation loop when the function values decrease below  $10^{-10}$ . We perform three tests. In the first test, we fix the number of variables  $n = 2$  while change the population size  $\lambda = 10, 20, 50, 100$ . The evolution curves (Figure A.2) indicate that the code is working and as the population size becomes larger, the results converge faster.

In the second test, we fix the population size at 50 and change the number of variables  $N = 2, 5, 10, 15$ . The evolution curves are presented in Figure A.3. Again, these curves indicate that the code is working well up to  $N = 15$ , which is above the size of the quasar spectrum fitting problem ( $N = 11$ ). The number of generations requested to converge increases with problem size.

In the final test, we fix both population size  $\lambda = 50$  and DOF  $N = 10$ , and run the code 100 times. We want to determine (1) whether the function can converge to the same value every run; (2) if it can, the distribution of the number of generations it needs to converge (Figure A.3). We find that the objective function is below  $10^{-10}$  and  $|x_i - 1| < 10^{-3}$  for all the runs. The distribution resembles a Gaussian with a median of 247 (Figure A.4), and a dispersion of  $\sim 20$ . Only 3 runs need more than 300 generations.

The three tests verifies that the code is working well for the Rosenbrock function optimization problem up to  $N = 15$  with acceptable reliability. In the following sections, we use this program to fit the quasar spectrum.

### A.5.2 Results in Default Setup

In the default setup, we fix the population size  $\lambda = 220$  and the number of selected offspring  $\mu = 110$ . The initial values of the free parameters are, in general, of the same order of magnitude as the typical values (but *not* exactly the fitting results using other methods). The optimization process terminates at the 72nd generation. The values of basic parameter are presented in Table A.2. The fitting results in the default setups are presented in Figure A.5. The optimal parameters we obtained are presented and compared with previous results using LMA in Table A.3. Comparing with the fitting result in Figure A.1, we find a significant improvement around the emission lines between 2500 Å and 3000 Å. Another evident improvement is the region around 3600 Å and 3700 Å. The LMA overproduces the flux around this wavelength range but the CMA-ES produces a much better fit. However, the CMA-ES over-produces the PL continuum between 4500 Å and 5000 Å. In order to fit the broad bump (contributed by iron emission) around 4500 Å, the overall scale factor of the optical iron template is large so that the total flux beyond 5000 Å is over-produced. However, in general, the CMA-ES algorithm produces a good balance between all the emission components and the overall quality is indeed improved comparing with LMA method.

### A.5.3 Varying Parameters

In this section, we compare the performance of CMA-ES on this particular problem by changing population size  $\lambda$  and number of offspring  $\mu$ . First, we fix the proportion of selected offspring with respect to the entire population, which is 50%, but change the population size. The evolutionary curves are shown in Figure A.6 and the number of generations for each population size is listed in Table A.4.

The population size changes from 55 to 440 as the color changes from red to blue. We can see that neither the red nor the blue curve has the best performance. The red curve, which represents the case  $\lambda = 55$ , converges to the optimal values at  $t \sim 110$  but it is not the fastest option. The blue curve, which represents  $\lambda = 440$ , actually does not converge to the optimal value; it bounces back at generation 100, returns at generation 170, and then remains around  $f = 110$ . It even exhibits a gradual increase after  $t \sim 170$ . The curve that shows the fastest convergence represents the case  $\lambda = 330$ , which is 30 times the total number of free parameters.

Next, we fix the population size  $\lambda = 22$ , but change the proportion of selected offspring. The evolutionary curves are shown in Figure A.7 and the number of generation for each proportion is listed in Table A.5. The color changes from red to blue as the proportion increases from 1/10 to 1/1.5. From Figure A.7, we do not see a significant difference among these cases, although the case in which  $\mu/\lambda = 1/3$  converges the fastest with the minimum number of generations (64).

Finally, we test the reliability of CMA-ES on the quasar spectrum fitting problem by running the program under the default setups ( $\lambda = 220$ ,  $\mu = 110$ ) 100 times. The distribution of the number of generations the evolution terminates is represented in Figure A.8. This figure illustrates that the in general the CMA-ES algorithm is reliable; about 60% of runs converge to the optimal value.

## A.6 Conclusion and Discussion

The CMA-ES method has achieved a preliminary success in the continuum fitting problem of a typical composite quasar spectrum. In the best case, it can finish fitting a spectrum in about 1 minute under the default setups. Although this is about 5 times longer than the LMA algorithm, it is still within the acceptable time scale. The most important aspect is that *the overall fitting quality is significantly improved and the optimal results do not depend on the initial parameters*. It is then feasible to extend its application to more quasar spectra and more complicated cases.

The real case can be more complicated. For instance, an iron emission template (either UV or optical) can be subdivided into a number of sections and each section may have a different scale factor. For some objects, a single PL is not enough and we usually need to include another PL at  $\lambda > 5600 \text{ \AA}$  (if covered). Because of effective exposure time, spectra may have different signal-to-noise ratio. All of these factors may complicate the fitting process. Although the task in this work is simplified, the spectrum we are using for the test is representative of the entire sample.

The fact that the theoretical reduced  $\chi^2$  value ( $\chi^2 \sim 1$ ) is not reached is proved not because of the algorithm but the model used the fit the spectrum. A single power-law may not be accurate description to the overall spectral profile. The iron emission shape may vary with different quasars, so a single iron emission template may not be exact. The emission lines can also contribute some flux at the benchmark wavelength points so it may not correct to only count the contribution from the continuum at these wavelength points. Another issue is the non-uniformity of the benchmark wavelength points. When selecting these wavelength points, we were trying to use *as many* constraints as possible. However, because the benchmark points must avoid *emission lines* and there is a region in which the iron emission template is *not available* (between 3000  $\text{\AA}$  and 3500  $\text{\AA}$ ), these benchmark points are not uniformly distributed; this could affect the evaluation.

Despite these limitations, the CMA-ES produces us a limit we can reach under the best fitting model we can provide. It provides an acceptable fitting of a quasar spectrum in a timely manner. Based on these results, we can further fit the *emission lines* seen in the spectra.

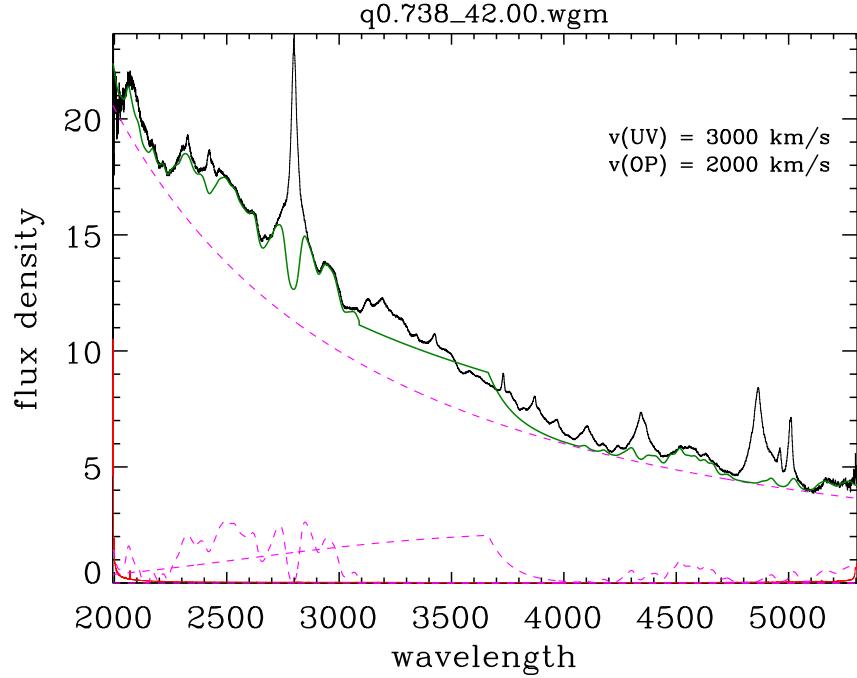


Fig. A.1 The LMA fit to the SDSS composite quasar spectrum at  $z = 0.738$  and  $\log l_\lambda(2200 \text{ \AA}) = 42.00$ . The original spectrum is shown in black. The  $x$ -axis is rest-frame wavelengths in angstroms ( $\text{\AA}$ ) and the  $y$ -axis is flux density in  $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ . The green curve is the underlying continuum fit using the LMA; the dashed magenta lines are the individual continuum components.

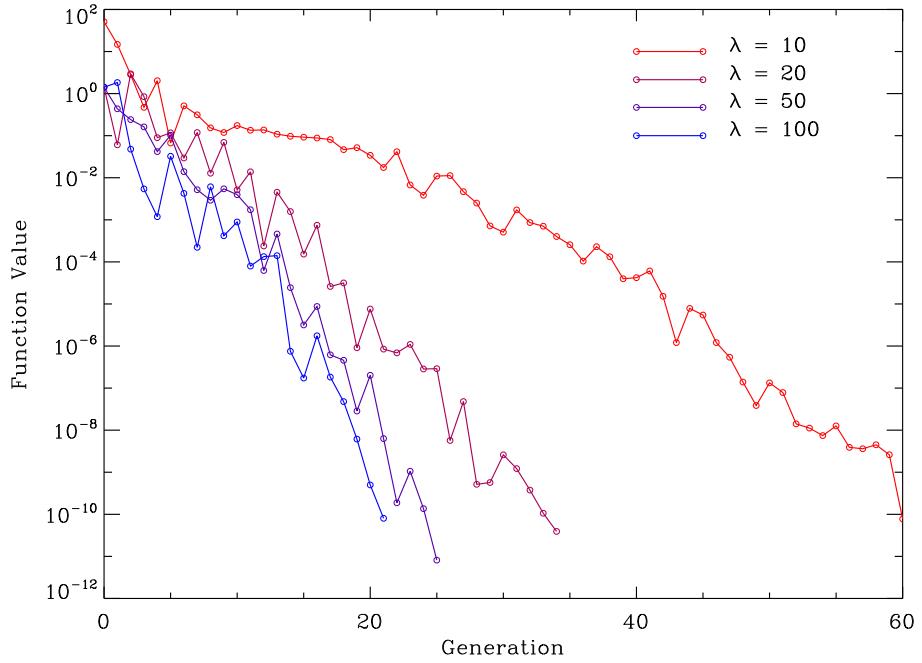


Fig. A.2 Testing results of Rosenbrock function by varying population size using CMA-ES. The colors of curves change from red to blue as the population size changes from 10 to 100. We fix degree of freedom (DOF) as 2. It is clear that a smaller population takes much more generations to converge than a large population.

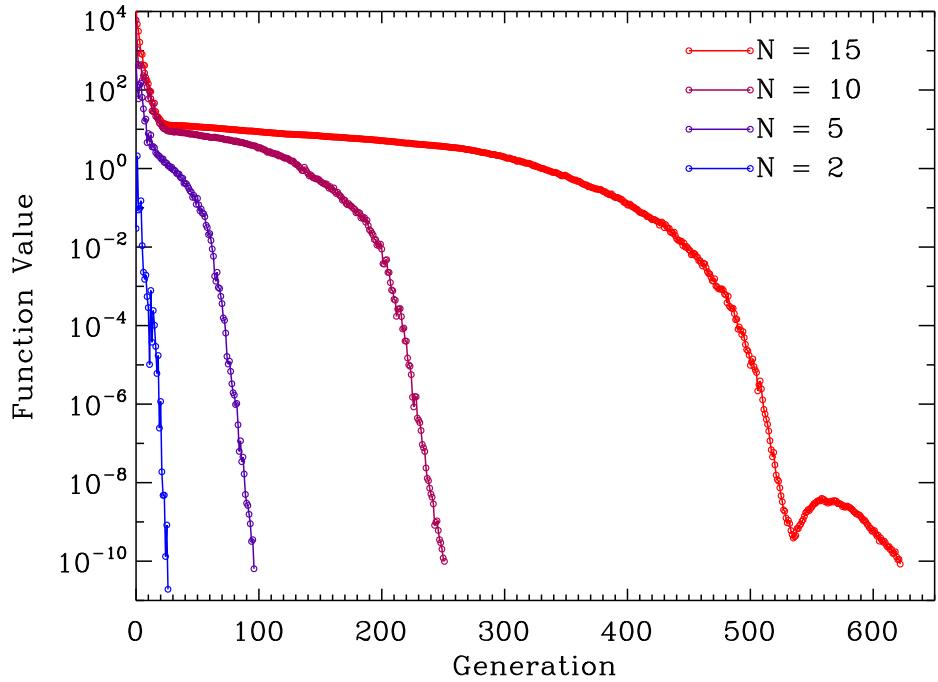


Fig. A.3 Testing results of Rosenbrock function by varying DOF. The colors of curves change from red to blue as the DOF decreases from 15 to 2. We fix the population size as  $\lambda = 50$ . It is clear that an evolutionary process with a higher DOF converges much more slowly than an evolutionary process with a lower DOF.

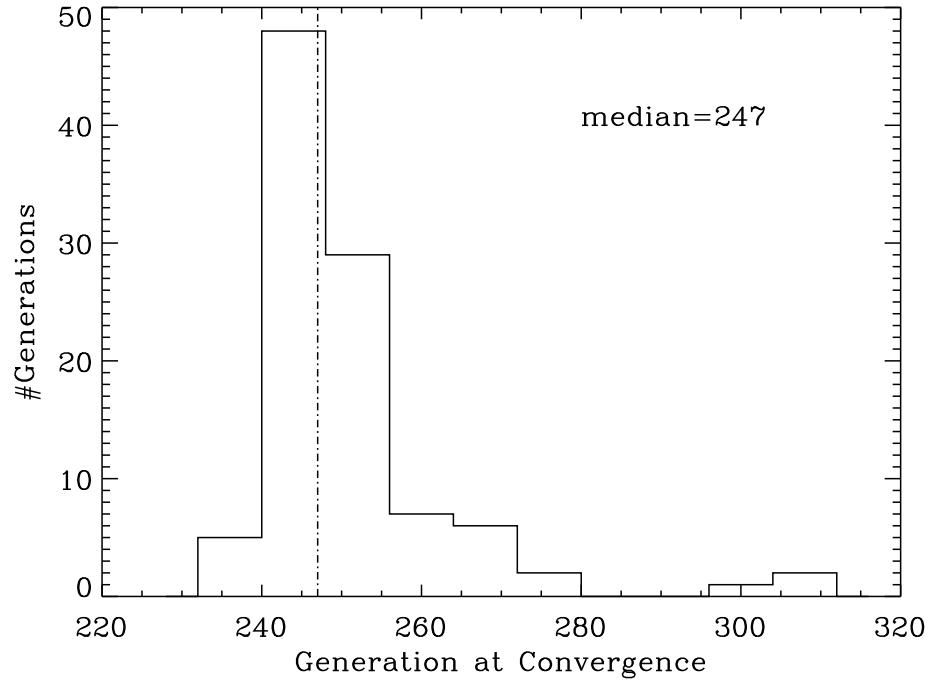


Fig. A.4 Distribution of the number of generations the Rosenbrock function takes to converge to  $10^{-10}$  using CMA-ES. We set the population size  $\lambda = 50$  and DOF  $N = 10$  and run the code 100 times. The median value (247) is presented as a dot-dashed line.

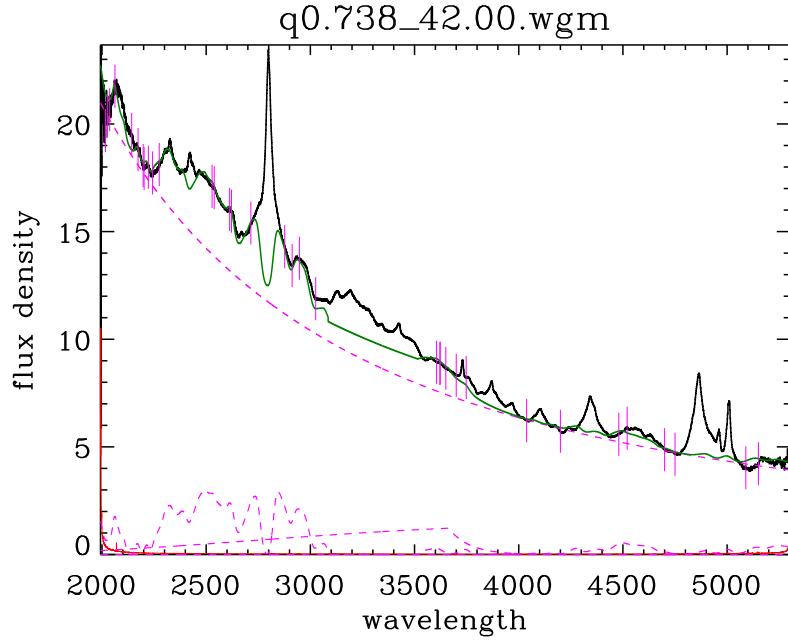


Fig. A.5 Fitting results using CMA-ES under the default setup. The legends are the same as Figure A.1, except that we over plot vertical lines in magenta at the benchmark wavelength points.

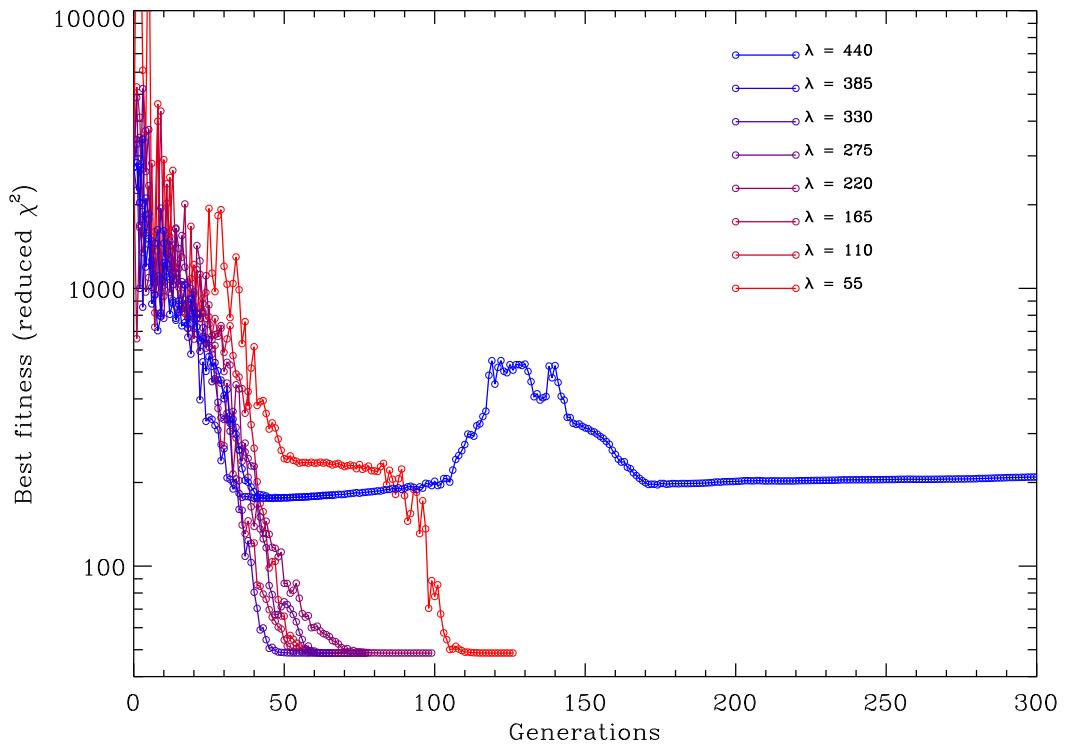


Fig. A.6 Evolutionary curves as the population size changes. The colors of curves change from red to blue as the population increases from 55 to 440. The evolutionary curve at  $\lambda = 440$  does not coverage. The curve with  $\lambda = 55$  converges the most slowly; the curve with  $\lambda = 330$  converges the fastest.

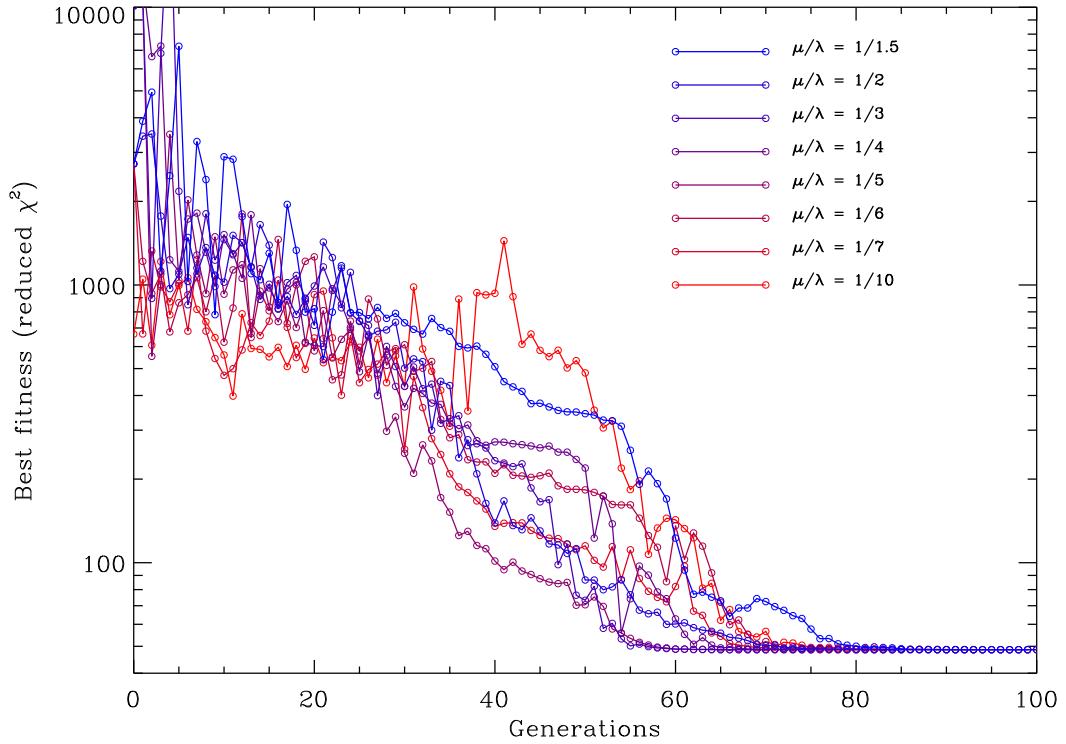


Fig. A.7 Evolutionary curves for different proportions of selected offspring with respect to population size. The colors of curves change from red to blue as  $\mu/\lambda$  increases from 1/10 to 1/1.5. These curves do not show a significant differenct convergence speed.

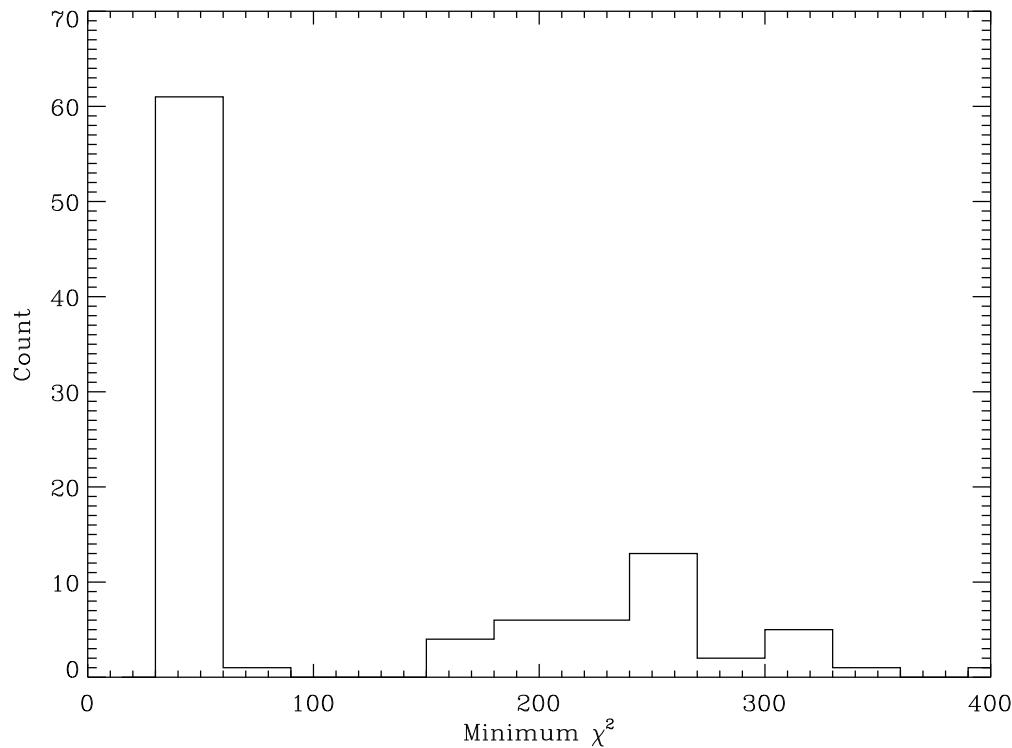


Fig. A.8 Distribution of  $\chi^2$  values at termination in default setups  $\lambda = 220, \mu = 110$ . About 60% of cases converge (the high bar on the left).

Table A.1. Spectral components and free parameters.

N	Name	Expression	Free Para.	Constraint
1	Power-law (PL)	analytical	$\alpha$ $\beta$	$\alpha \sim -1.6$ $\beta \sim 6-7$
2	Small Blue Bump (SBB)	Analytical <sup>1</sup>	$A_B$ $T_B$ $\tau_B$	$A_B > 0$ $T_B \sim 10^4$ $\tau_B \sim 1$
3	UV iron emission	templates	$A_U$ $v_U$ $\Delta z_U$	$A_U > 0$ $v_U = 1500, 2000, \dots, 9500$ $ \Delta z_U  \leq 0.005$
4	Optical iron emission	templates	$A_O$ $v_O$ $\Delta z_O$	$A_O > 0$ $v_O = 1500, 2000, \dots, 9500$ $ \Delta z_O  \leq 0.005$

<sup>1</sup>The exact form is seen in Equation A.1.

Table A.2. Initialization. Values of  $\mathbf{m}$  are typical values from previous results.

Para.	Initialization	Remarks
$\mathbf{m}$	$[-1.6, 7, 150, 10000, 1, 0.02, 3000, 0, 0.005, 3000, 0]$	(weighted) means of parameters, including $[\alpha, \beta, A_B, T_B, \tau_B, A_U, v_U, \Delta z_U, A_O, v_O, \Delta z_O]$
$\lambda$	$20 \times N = 220$	population size
$\mu$	$\lambda/2 = 110$	number of parents for recombination
$\mathbf{w}$	$w_j = \log(\mu + 1/2) - \log j, \quad j = 1, 2 \dots, \mu$	normalized weights
$\mu_{\text{eff}}$	$1/\sum w_i^2 = 57.6986$	variance-effectiveness
$\mathbf{C}$	$C_{jk} = 1(j=k), = 0(j \neq k) \quad j, k = 1, 2, \dots, N$	covariance matrix
$c_c$	$(4 + \mu_{\text{eff}}/N)/(N + 4 + 2\mu_{\text{eff}}/N) = 0.3627$	time constraint for cumulation for $\mathbf{C}$
$c_\sigma$	$(\mu_{\text{eff}} + 2)/(N + \mu_{\text{eff}} + 5) = 0.81$	const for cumulation for $\sigma$ control
$c_1$	$2/\left \frac{(N+1.3)^2 + \mu_{\text{eff}}}{(N+1)^2 + \mu_{\text{eff}}}\right  = 0.01$	learning rate for rank-one update of $\mathbf{C}$
$c_\mu$	$2(\mu_{\text{eff}} - 2 + 1/\mu_{\text{eff}})/\left \frac{(N+1)^2 + \mu_{\text{eff}}}{(N+1.3)^2 + \mu_{\text{eff}}}\right  = 0.49$	learning rate for rank- $\mu$ update
$d_\sigma$	$1 + 2 \max\left(0, \sqrt{[(\mu_{\text{eff}} - 1)/(N - 1)] - 1}\right) = 4.157$	damping for $\sigma$
$\mathbf{p}^c$	$p_j^c = 0, \quad j = 1, 2, \dots, N$	evolution paths for $\mathbf{C}$
$\mathbf{p}^\sigma$	$p_j^\sigma = 0, \quad j = 1, 2, \dots, N$	evolution paths for $\sigma$
$\ \mathcal{N}(\mathbf{0}, \mathbf{I})\ $	$\sqrt{N} \left 1 - 1/(4N) + 1/(21N^2)\right  = 3.24$	expectation of $\ \mathcal{N}(\mathbf{0}, \mathbf{I})\ $

Table A.3 Optimal parameter values using LMA and CMA-ES (under the default setup).

Para.	LMA	CMA-ES
$\alpha$	-1.768	-1.713
$\beta$	7.147	6.974
$A_B$	134.827	150.681
$\tau_B$	1	0.5153
$T_B$	10000	10001
$A_U$	0.028	0.044
$v_U$	3000	3000
$\Delta z_U$	0.0000	-0.0008
$A_O$	0.015	0.015
$v_O$	3000	3000
$\Delta z_O$	0.000	-0.0046

Table A.4 Changes of termination generation as a function of population size  $\lambda$ . Proportion of selected offspring is fixed to be 50%.

Population Size	Number of Generation
55	127
110	100
165	82
220	72
275	78
330	66
385	76
440	NA

Table A.5 Numbers of generations as a function of the proportion of selected offspring with respect to the fixed total population size  $\lambda = 220$ .

Proportion	Number of Generations
1/10	101
1/7	81
1/6	94
1/5	82
1/4	69
1/3	64
1/2	72
1/1.5	102

## Appendix B

# Software Design & User Manual for QSFIT

### B.1 Introduction

The SDSS Quasar Spectral Fitting Software Package (QSFIT) is designed to fit a typical quasar spectrum in the rest-frame UV and/or optical band. The wavelength coverage of an *observed* SDSS quasar spectrum is 3700 Å-9200 Å; the spectral resolution is  $\sim 2000$ . The continuum of the spectrum can be fit with two power-laws (PLs), intersecting at  $\sim 5600$  Å. The spectrum between Ly $\alpha$  ( $\sim 1216$  Å) and  $\sim 5600$  Å is called UV part; the wavelength longer than  $\sim 5600$  Å is denoted as the optical part. The power-law can be represented as

$$f_\lambda = 10^\beta \lambda^\alpha \quad (\text{B.1})$$

in which  $\alpha$  is called *spectral index* while  $\beta$  is called spectral *coefficient* or *scale factor*.

#### B.1.1 Software Availability

**The user is always recommended to use the most recent version of this software package.** Due to historical reasons, the software parts are not assembled as a single compressed file. Users must copy/download necessary programs from the designated directories/urls.

- Install **IDL 5.5** or newer version.
- Copy IDL astro library and add it to your IDL path: <http://idlastro.gsfc.nasa.gov/>.
- Copy the IDL application collection:  
`nprmac.astro.psu.edu:jxw394/idl/app/` and add it to your IDL path.
- Copy the UV iron template collection:  
`nprmac.astro.psu.edu:jxw394/research/sdss/project/composite/data/FeTempl_VC/`.  
 You are recommended to copy this folder to a separate directory. You will need to specify this directory in `qconfit.par` or `qconfita.par`.
- Copy the optical iron template collection:  
`nprmac.astro.psu.edu: jxw394/research/sdss/project/composite/data/Veron2004_VC/`.  
 Again, you are recommended to copy this folder to a separate directory. You will need to specify this directory in `qconfit.par` or `qconfita.par`.
- Copy the whole continuum and emission line fitting package:  
`nprmac.astro.psu.edu: jxw394/research/sdss/project/qsfit/`. This folder contains

all the programs, parameter, input files and list files needed for this software package. After copying, rename the folder `wgm/` as a name the user wants, which is related to the project of the user, and empty all the sub-folders, *e.g.*, `bicheck.out/`. Some folders are used for the user to save input/parameter files, others are used to save programming running results.

### B.1.2 Convention

We strongly recommend the user keep the conventional extensions defined in the original software package, although the input spectral files are allowed to have any extension. For example, the output of `qconfita.pro` is always named as `.tmp`. These extensions are used to distinguish files input/output for different procedures/functions. Changing these extensions may cause the program to malfunction.

### B.1.3 Copyright and Maintenance

This software is designed, written, and maintained by Jian Wu (Department of Astronomy and Astrophysics, Pennsylvania State University). All rights reserved. The user should contact Jian Wu for most updated code and issue credit for the use of this software package. Please send inquiries or problems to `jwu@astro.psu.edu`.

## B.2 Software Structure

### B.2.1 Data Structure and Algorithms

The intersection point of two power-law is computed as follows. Assuming  $f_1 = 10^{\beta_1} \lambda^{\alpha_1}$  and  $f_2 = 10^{\beta_2} \lambda^{\alpha_2}$  intersect at  $(\lambda_0, f_0)$ , then

$$\lambda_0 = 10^{(\beta_2 - \beta_1)/(\alpha_1 - \alpha_2)}$$

The uncertainty of  $\lambda_0$  given the uncertainties of  $\alpha_1$ ,  $\alpha_2$ ,  $\beta_1$ , and  $\beta_2$  are

$$\frac{\delta \lambda}{\lambda \ln 10} = \frac{1}{\alpha_1 - \alpha_2} \sqrt{\delta \beta_2^2 + \delta \beta_1^2 + \left( \frac{\beta_2 - \beta_1}{\alpha_1 - \alpha_2} \right)^2 (\delta \alpha_2^2 + \delta \alpha_1^2)}$$

We assume the uncertainties of all parameters follow Gaussian distributions.

### B.2.2 Software Parts

This software package consists of three main programs, a number of data files, and parameter files. In this section, we will introduce these programs and their functions. We suggest the user copy all the codes and files before start. The programs and their associated data/parameter files are selectively described below.

- Main programs.
  1. **bicheck.pro**: Performs an initial estimation of the power-law (PL) indices and coefficients of a quasar continuum.

- 2. **qconfit.pro/qconfita.pro.** Manually/automatically fits a quasar spectrum from SDSS.
- 3. **mlfit.pro/vmlfit.pro:** Fits multiple emission lines in a quasar spectrum. A variation of this code is **vmlfit.pro**, which has the same function.
- Check list of functions/procedures and their input/parameter/list files. Some of them are in my personal IDL application library (marked with \*), because they are more general and are more frequently invoked.
  - 1. *bicheck.pro.* Estimate the PL indices and coefficients of quasar spectra in UV/optical band. Associated files include
    - (a) *bicheck.par.*
    - (b) *bicheck.list.*
    - (c) *nod.list.*
  - 2. *bicheckplot.pro.* Plot the results of *bicheck.pro* into a PS file. Associated file include
    - (a) *bicheck.par.* This is the same file used by *bicheck.pro*. The program needs to know output directory from this file.
    - (b) *line.in.* Emission line list.
  - 3. *bip.pro.* Calculate the total flux contributed by a power-law continuum, a small blue bump and the iron continuum.
  - 4. *bipower.pro.*
  - 5. *bipower2.pro.*
  - 6. *chisq.pro.*
  - 7. *continuum.pro.*
  - 8. *ewpow.pro.* Calculate the equivalent width of an emission line standing on top of a power-law continuum. Invoked by *mlfit.pro*.
  - 9. *ewbipower.pro.* Same as *ewpow.pro* but generalized for a broken power-law. Invoked by *vmlfit.pro*.
  - 10. *filesearchr.pro\**.
  - 11. *filesearchw.pro\**.
  - 12. *filterfun.pro.*
  - 13. *gauss1\_sh.pro.*
  - 14. *gauss2\_sh.pro.*
  - 15. *gauss3\_sh.pro.*
  - 16. *gauss4\_sh.pro.*
  - 17. *gauss5\_sh.pro.*
  - 18. *gsfit\_m.pro.*
  - 19. *gsfit\_s.pro.*

20. *indgens.pro*.
21. *ironscls.pro*.
22. *linesm.pro\**.
23. *loadcolors.pro\**.
24. *loadpscolor.pro\**.
25. *mpfitfun.pro*. Kernel fitting program. Contained in IDL astro library.
26. *mysymb.pro\**.
27. *nearmax.pro*.
28. *nearmin.pro*.
29. *print\_progress.pro\**.
30. *qconfita.pro*. Associated files are
  - (a) *qconfita.par*.
  - (b) *bicheck.par*.
  - (c) *qconfita.in*.
  - (d) *benchmark.list*.
31. *qconfitaloop.pro*. Fit a list of quasar spectra in a loop, using *qconfita.pro*.
32. *qconfitfun.pro*. Compute the flux of the model spectrum at the desired wavelength points when the z-parameter is not fixed.
33. *qconfitfun\_fz.pro*. Compute the flux of the model spectrum at the desired wavelength points when the z-parameter is fixed.
34. *qconfitfun\_fzop.pro*. Compute the flux of the model spectrum at the desired wavelength points when the z-parameter is only fixed for the *optical* iron template.
35. *qconfitfun\_fzuv.pro*. Compute the flux of the model spectrum at the desired wavelength points when the z-parameter is only fixed for the *UV* iron template.
36. *qconsyn.pro*. Synthesize and plot the continuum spectrum output by *qconfita.pro* (.tmp file). Associated files are
  - (a) *qconsyn.par*.
  - (b) *qconsyn.in*.
  - (c) *benchmark.list*.
37. *qconsynplot.pro*. Plot a list of quasar spectra by invoking *qconsyn.pro*. Associated files are
  - (a) *qconsyn.list*.
38. *qconsynfun.pro*.
39. *qsview.pro\**. View a quasar spectrum in UV/optical band. Optimized to view a spectrum within 900–8000 Å.
40. *readin.pro\**.

41. *readpar.pro*\*.
  42. *readqslist.pro*.
  43. *readqslist2.pro*. Read the QSO list file and return the desired spectral file names.
  44. *rmprefix.pro*\*.
  45. *rms.pro*\*.
  46. *rmsuffix.pro*\*.
  47. *rmwidcard.pro*.
  48. *sbb.pro*. Associated files are
    - (a) *sbb.par*.
  49. *speccut.pro*\*.
  50. *status.pro*\*.
  51. *strarrcontain.pro*\*.
  52. *strcontain.pro*\*.
  53. *timeelapsed.pro*\*. Calculate the time elapsed when running a code.
  54. *touchbound.pro*.
  55. *vmlfit.pro*. Associated files include
    - (a) *vmlfit.par*.
    - (b) *vmlfit.in*.
    - (c) *vmlfit.list*.
    - (d) *broadline.list*.
    - (e) *sdsswave.sav*.
- Input>List/Parameter files (\*.in, \*.par, \*.list). The package adopts a special format of parameter files (\*.par) and input files (\*.in). A parameter file contains adjustable parameters used in a corresponding program in which each parameter has a *single* possible value. An input file contains adjustable parameters used in a corresponding program in which each parameter has *multiple* possible values. All the \*.par files use the same format and so do the \*.in files. The program used to read the \*.par files and \*.in files are *readpar.pro* and *readin.pro*. Both formats are within the IDL application collection.
    - A parameter file (.par file) contains four columns separated by semicolons. **Empty lines are not permitted in .par files.**
      1. Parameter value. Scientific notion supported.
      2. Parameter type. We support four types: **int**, **flt**, **dbl** and **str**. For example, a directory name is of type **str**. **All directories must end with a slash “/”.**
      3. Parameter name used in program.
      4. Notes/Explanation of a parameter. Space permitted.

- An input file (\*.in) contains
  1. Quasar spectra. These could be original SDSS quasar spectra or composite quasar spectra. They should contain three columns separated by space. Column 1 is the wavelength; Column 2 is the flux density  $f_\lambda$ ; Column 3 is the uncertainty in  $f_\lambda$ . These spectra can be placed in a folder but you need to specify the folder in the parameter files of fitting programs (see below).
  2. *bicheck.par*. Parameter file for *bicheck.pro*.
  3. *nod.list*. A list of wavelength windows for PL estimation; used by *bicheck.pro*. This file has three columns, separated by hyphens (“-”). In each row, the first number is the starting wavelength of the window; the second number is the ending wavelength of the window; the third number is the decrease factor. This is a *default* wavelength window list. If you cannot achieve a good fit using the default window and decrease factors, you can write a nod file for a specific quasar (see the next item).
  4. *[qsoname].nod*. The same as *nod.list* except that it is designed for a particular quasar. *bicheck.pro* first searches for this file. If it does not exist, *bicheck.pro* uses the general nod file *nod.list*. It is recommended that you use the “qsoname” which is part of the original spectrum. For example, if the original spectrum is named as “Mrk727.dat”, this file should be “Mrk727.nod”, and put it into a special directory designated in *bicheck.par* (see § 1).
  5. *line.in*. A list of UV and optical emission lines appearing in a quasar spectrum, used by several programs. This file has two columns separated by space. The first column is the name of species producing an emission line, *e.g.*, , CIII, NIII], Ly1. Lyman and Balmer series are expressed in numerical numbers rather than Greek letters, so Ly1 stands for Ly $\alpha$ . The second column is the central vacuum rest-frame wavelength of the emission line in angstroms.
  6. *benchmark.list*. A list of wavelengths at which the original and the model spectra are compared. It contains only one column.

## B.3 User Manual

Fitting a quasar spectrum can be divided into three major steps: estimating the power-law, continuum joint fit, and emission line fit. Currently, these three steps can only be run separately.

### B.3.1 Power-law Estimation

The major programs to run this step are *bicheck.pro* and *bicheckplot.pro*. Before running these codes, a few initial setups are required on the *.par* and *.in* files. Most parameters do not need any change. Some information needs be changed only one time, *e.g.*, file path. In the context below, we define the *working directory* as the same directory where you run the code.

1. Editing *bicheck.par*. This file must be copied to the working directory.

- *bpoint* - no need to change.
- *nodfile* - no need to change, but you must copy this file to working directory.
- *dirin* - **change**. This is the directory where you put all the rest-frame quasar spectra you need to fit.
- *dirnod* - **change**. This is the directory where you put the wavelength window files for specific quasars. It is recommended to be a sub-directory of *dirin* named *nod*.
- *outfile* - no need to change. This is the output file of *bicheck.pro*. This file has a header and contains 14 columns, with one row per spectrum. The meaning of each column is given in Table B.1.

Table B.1 Columns in *bicheck.pro* output file.

Header	Explanation
SPECFILE	Original quasar spectrum file name
PINDEX1	UV PL spectral index
PINDEX2	Optical PL spectral index
PCOEFF1	UV PL spectral coefficient
PCOEFF2	Optical PL spectral coefficient
BPOINT	Wavelength point where the two PL intersects
NOD1_0	Left wavelength of UV window 1
NOD1_1	Right wavelength of UV window 1
NOD1_2	Left wavelength of UV window 2
NOD1_3	Right wavelength of UV window 2
NOD2_0	Left wavelength of Optical window 1
NOD2_1	Right wavelength of Optical window 1
NOD2_2	Left wavelength of Optical window 2
NOD2_3	Right wavelength of Optical window 2

- *dfact* - no need to change. Decremental factor by default.
  - *smf* - **may change**. Smooth factor by default.
2. Editing *bicheck.list*. This is the input quasar spectrum list by default. It has two columns, the first column is a flag, there are typically 3 flag values (Table B.2) You need to flag the spectra you want to fit with “r”.
  3. Run *bicheck.pro*. There is a parameter and a keyword to run the program (Table B.3). This program is run as below:

**IDL> bicheck[,file][,/auto]**

4. You should immediately check the output file (*bicheck.tmp.out* by default) after running and rename it as necessary. This program will *not* remind you if this file

Table B.2 Flags in *bicheck.list*.

Flag	Explanation
x	Skip this spectrum
r	Run/Fit this spectrum
f	Spectral fitting completed

Table B.3 Parameters and keywords of *bicheck.pro*.

Name	Property	Option	Default	Explanation
file	parameter	optional	<i>bicheck.list</i>	list of quasar spectrum names
auto	keyword	optional	0	If set, program pauses at each spectrum; If not, program continues without stop.

exists. It will simply *overwrite* the file. If you want to write a postscript (PS) plot of the original spectrum with the PL fit into a file, we recommend renaming it as *bicheck.out* because it is the default input file of *bicheckplot.pro*.

5. Edit *bicheckplot.par*. There are only two parameters in this file.
  - *dirin* - **Change**. This is the directory where you put all the rest-frame quasar spectra you need to fit.
  - *dirout* - no need to change. The figures will be plotted into a single file and written into this directory. By default *dirout* is the current directory but you can certainly change it.
6. Run *bicheckplot.pro*. There is only one keyword - FILE. It is optional and has a default value of *bicheck.out*. This should be the output file by *bicheck.pro*. However, because the default file of *bicheck.pro* is *bicheck.tmp.out*, you may need to rename that file to *bicheck.out*. The user can specify an alternative name as long as it was output by *bicheck.pro*. The program is run as

**IDL> bicheckplot[ ,file=file]**

There are two output files produced by this program

- *bicheckplot.ps* - the original quasar spectra overplot with the estimated PL, one spectrum per page.
- *[qsoname].noc* - PL continuum subtracted spectra written in *dirin* as specified in *bicheckplot.par*. This file will be used to estimate the iron template scaling factor.

### B.3.2 Continuum Join-Fit

The major code to run for this step is *qconfit.pro* and *qconfita.pro*. The former is for manual fitting and the latter is for automatic fitting. Because of the complexity of the manual operations, this section only covers the automatic fitting at this moment. As before, some initial setups need to be done before we run *qconfita.pro*.

1. Edit *qconfita.par*. This parameter file consists many parameters.
  - *indir* - **Change**. This is the directory where you put all the rest-frame quasar spectra you need to fit.
  - *outdir* - no need to change. This is the output directory and is the current directory by default.
  - *pardir* - no need to change. Directory for *[qsoname].par*.
  - *infilendir* - no need to change. Directory for *[qsoname].in*.
  - *uvdir* - **Change**. The UV iron template directory.
  - *opdir* - **Change**. The optical iron template directory.
  - *bicheckout* - no need to change. The output file of *bicheck.pro*. Note that you must rename this output file as *bicheck.out* (the original file name is *bicheck.tmp.out*).
  - *f\_unit* - no need to change. Scale factor for the flux density of the SDSS spectra. It is  $10^{-17}$  by default.
  - *bmfile* - no need to change. A file containing benchmark wavelengths where the original and the model spectra are compared.
  - *smf* - **May change**. Smooth factor by default. This default smooth factor is suitable for composite spectra. For real SDSS spectra, you may need a higher value.
  - *snlim* - **May change**. Limit of signal-to-noise (S/N) ratio. Flux pixels are discarded if the S/N ratio is below this value. You may need another value depending on your spectral quality.
  - *woffset* - **May change**. Wavelength offset with respect to a benchmark wavelength. Pixel fluxes between (benchmark–*woffset*,benchmark+*woffset*) are evaluated (see *bmfile*).
  - *tnm\_uv* - no need to change. Prefix of UV iron template files.
  - *norm\_uv* - no need to change. This is a universal normalization factor of UV iron template. Another factor will be multiplied to obtain the real flux level of the iron emission.
  - *norm\_op* - no need to change. This is a universal normalization factor of optical iron template. Another factor will be multiplied to obtain the real flux level of the iron emission.
  - *pidx\_uv* - no need to change. PL index  $\alpha_{UV}$  at UV band ( $f_\lambda = 10^\beta \lambda_{UV}^\alpha$ ). This is a default spectral index if *qconfita.pro* cannot find the output of *bicheck.pro*.

- *pscl\_uv* - no need to change. PL coefficient  $\beta$  at UV band (see *pidx\_uv*).
- *pidx\_op* - no need to change. This is a default spectral index  $\alpha_{op}$  in optical band if *qconfita.pro* cannot find the output of *bicheck.pro*.
- *pscl\_op* - no need to change. This is a default spectral coefficient  $\beta$  in optical band if *qconfita.pro* cannot find the output of *bicheck.pro*.
- *fix\_pidx\_uv* - **May change.** Set to 1 if you want to fix the PL index at UV band (either from *bicheckout* or the default value), otherwise, set it to 0.
- *fix\_pscl\_uv* - **May change.** Set to 1 if you want to fix the PL coefficient at UV band (either from *bicheckout* or the default value), otherwise, set it to 0.
- *fix\_pidx\_op* - **May change.** Set to 1 if you want to fix the PL index at optical band (either from *bicheckout* or the default value), otherwise, set it to 0.
- *fix\_pscl\_op* - **May change.** Set to 1 if you want to fix the PL coefficient at optical band (either from *bicheckout* or the default value), otherwise, set it to 0.
- *z\_uv* - no need to change. Initial redshift offset of UV iron template.
- *fix\_z\_uv* - **May change.** Set to 1 to fix *z\_uv* and 0 to fit it.
- *z\_op* - no need to change. Initial redshift offset of optical template.
- *fix\_z\_op* - **May change.** Set to 1 to fix *z\_op* or 0 to fit it.
- *v\_uv* - no need to change. Initial velocity dispersion of UV iron template.
- *fix\_v\_uv* - **May change.** Set to 1 to fix *v\_uv* or 0 to fit it.
- *v\_op* - no need to change. Initial velocity dispersion of optical iron template.
- *fix\_v\_op* - **May change.** Set to 1 to fix *v\_op* or 0 to fit it.
- *tscl\_uv* - no need to change. Initial scale factor of UV iron template.
- *fix\_tscl\_uv* - **May change.** Set to 1 to fix *tscl\_uv* or 0 to fit it.
- *btemp* - no need to change. Balmer temperature  $T_{BBB}$ .
- *btau* - no need to change. Balmer optical depth.
- *bscl* - no need to change. Initial scaling factor of small blue bump.
- *fix\_bscl* - **May change.** Set to 1 to fix *bscl* or 0 to fit it.
- *c4200* - no need to change. An imperial adjustment factor at the 4200 Å benchmark wavelength.

If you want a *.par* file for a particular quasar spectrum, you can create a copy of *qconfita.par* and modify, and rename it. It is strongly recommended that you save the renamed *.par* file in the directory designated by *pardir*. The program will preferentially use this parameter file. If it does not exist, it will use the default one.

## 2. Edit *qconfita.in*.

- *trange\_uv* - no need to change. UV iron template wavelength range.

- *trange\_op* - no need to change. Optical iron template wavelength range.
- *wdiv\_uv* - no need to change. Empirical dividing wavelengths of UV iron template.
- *wdiv\_op* - no need to change. Empirical dividing wavelengths of optical iron template.
- *vval\_uv* - no need to change. Possible velocity dispersion values for UV iron template.
- *vval\_op* - no need to change. Possible velocity dispersion values for optical iron template.

Similarly to the .par file, you can make a copy of *qconfita.in*, modify it and save it in the directory designated by *infilendir*. The program will preferentially use this parameter file. If it does not exist, it will use the default one.

3. Edit *qconfita.list*. The meanings of flags are the same as *bicheck.list*.
4. Run *qconfita.pro* or *qconfitaloop.pro*. To run a single spectrum, use the former; to run a number of spectra in a loop, use the latter. The running command *qconfita.pro* contains one parameter and two keywords (Table B.4). The commands

Table B.4 Parameters and keywords in *qconfita.pro*.

Name	Property	Option	Default	Explanation
specfile	parameter	mandatory		original quasar spectrum file
plotg	keyword	optional	0	Set to 1 to launch window display or else figures are plot into a PS file
replace	keyword	optional	0	Set to 1 to overwrite the existed result file

to run these two programs are

```
IDL> qconfita, specfile[ ,/plotg][ ,/replace]
IDL> qconfitaloop
```

For each spectrum, the continuum fitting results are written into a file named *[qsoname].tmp*. If *plotg* is not set, a figure containing the underlying continuum component over the original spectrum is plotted into a file named *[qsoname].ps*.

5. Checking the results. The result file of *qconfita.pro* may contain multiple rows, each of them corresponds to one combination of velocity dispersions for the iron templates in UV and optical bands. Currently, we have not found an effective way to find the proper value of velocity dispersion for the iron templates. As a result, we plot and print all the results out for a visual inspection.

Alternatively, the user can simply fix the velocity dispersion of UV or optical template to be  $3000 \text{ km s}^{-1}$  if the sample size is so big that a visual inspection is impossible to implement. This is an empirical “typical” or “average” value. It may not apply to every spectrum, but should be a reasonable value for most quasars.

6. Run *qconsyn.pro* and *qconsynplot.pro*. The former program is used to synthesize and plot the continuum spectrum output by *qconfita.pro*. Before running this program, the user must edit *qconsyn.par* and *qconsyn.in*. These two files are imported from *qconfita.par* and *qconfita.in*, respectively, so most of the parameters have the same values. The only parameter to possibly change is *tmpdir*, which is the directory of spectrum parameter files. The command to run *qconsyn.pro* has one parameter and three keywords (see Table B.5). You must also edit *qconsynplot.list*, which has the same format as *qconfita.list*. The program *qconsyn.pro* can only deal

Table B.5 Parameters and keywords in *qconsyn.pro*.

Name	Property	Option	Default	Explanation
specfile	parameter	mandatory		original quasar spectrum file
psplot	keyword	optional	0	Set to 1 to plot figures into a PS file or else plot figures in an IDL window
modelf	keyword (output)	optional	0	Model flux. <i>modelf[0,*]</i> - total continuum <i>modelf[1,*]</i> - PL continuum <i>modelf[2,*]</i> - Small blue bump <i>modelf[3,*]</i> - UV iron emission <i>modelf[4,*]</i> - optical iron emission
tmpdir	keyword	optional	0	Same as <i>tmpdir</i> in <i>qconsyn.par</i> Just in case of difference from default value

with a single spectrum; it is invoked by *qconsynplot.list* to run a number of spectra.

```
IDL> .compile qconsynplot.pro
IDL> .go
```

The output files are *[qsoname].noc*, which is the continuum subtracted spectra (ready for emission line fitting), and *[qsoname-syn].ps*, which is the underlying continuum over the original quasar spectrum. Note that *[qsoname].noc* is different from the files output by *bicheckplot.pro*. By default the outputs by *bicheckplot.pro* are placed into the same directory as the original spectra while the outputs by *qconsyn.pro* are written into the current directory.

### B.3.3 Emission Lines Fitting

Emission lines are fit using either of the two codes: *mlfit.pro* and *vmlfit.pro*. They fit emission lines in a similar way except that they create different outputs. The program *mlfit.pro* writes line parameters for one quasar spectrum into *one* file, while *vmlfit.pro* writes line parameters of an emission line for *all* quasar spectra into one file. **We recommend using *vmlfit.pro* other than *mlfit.pro*.** We introduce how to run *mlfit.pro* first.

1. Edit *mlfit.par*. This file has only two parameters.
  - *dir\_rst* - Directory where you put all the rest-frame quasar spectra you need to fit.
  - *dir\_noc* - Underlying continuum-subtracted spectrum directory.
2. Edit *mlfit.in*. This input file is *different* from traditional input files described above. It is especially designed for *mlfit.pro* or *vmlfit.pro*. Each column is described below:
  - (a) Flags - **Change**. Two possible values: + or -. A + means including this line in fitting queue while a - means not fitting this line.
  - (b) Line ID number - no need to change. Line sequence number adopted from *line.in*,
  - (c) Emission line rest-frame wavelength in angstroms - no need to change.
  - (d) Emission line fitting window left boundary - **may change**. If you want to *group* several *adjacent* lines and let them be fit simultaneously, you *must* use the same value for this parameter.
  - (e) Emission line fitting window right boundary - **may change**. If you want adjacent lines to be fit simultaneously, you must use the same value for this parameter.
  - (f) Number of Gaussian profiles needed to fit the emission line - **may change**.
  - (g) Left boundary of a single Gaussian profile range - **may change**. This specifies the wavelength range of a single line even if it is blended with other transitions. The code will fit the spectrum within this wavelength range so as to determine the Gaussian  $\sigma$  of the line profile. If multiple Gaussian profiles are needed, this sigma will be used as the one of the narrowest component.
  - (h) Right boundary of a single Gaussian profile range - **may change**.
  - (i) Continuum choice - **may change**. This specifies which type of continuum is used; “mc” means *global* continuum while “lc” means *local* continuum.
3. Setup the mask files and are named as *[qsoname].mask*. These files are used in *mlfit.pro* if the user wants to mask certain wavelength ranges when fitting the emission lines. The spectrum may show strong associated absorption features that may lead to underestimations of emission line fluxes. This file can contain as many lines as needed. It contains only two columns delimited by hyphens (“-”).
4. Run *mlfit.pro*. The command has one parameter and one keyword.

- *qsolist* - mandatory. Spectrum file name *without* the suffix. For example, if the spectrum file name is *Mrk766.dat*, this parameter should be *Mrk766*.
- *auto* - optional. This program has two modes. If *auto* is set, it runs all the lines in a spectrum; otherwise, it is in the manual mode. In this mode, you can select and fit a specific emission line. If an emission line is grouped with other emission lines, all the lines in this group will be fit.

The command to run is

**IDL> mlfit, qsoname[, /auto]**

The steps to run *vmlfit.pro* are similar to that to run *mlfit.pro*.

1. Edit *vmlfit.par*. This is the default setup file for all emission lines, but if you need a parameter file for a certain emission line different from the default one, you can create a similar file called *[linename].vml.par*. The program will search for this file first. If it does not exist, the default parameter file is used. Here, *[linename]* is a name of an emission line after removing the “irregular” characters. For example, the semi-forbidden line name CIII]1908 is renamed as CIII1908.
  - *filterpass* - **may change**. The smaller it is, the less the high frequency component (noise) there is.
  - *dir\_noc* - **change**. The directory of underlying continuum-subtracted spectra.
  - *file* - **may change**. A file containing a list of spectra names whose emission lines are to be fit. This file has three columns, separated by spaces. The first column is a flag. Differently from previous list files, **spectrum files marked with an “x” is going to be run**. The second column is the numbers of quasar spectra. The third column is the name of quasar spectra **without suffix**. For example, if the spectrum file name is *Mrk766.dat*, this name should be *Mrk766*.
  - *sufnoc* - **may change**. The suffix of the continuum-subtracted spectra. Of course, it is *.noc* by default.
  - *dir\_ori* - **change**. Directory of the original quasar spectra.
  - *sufori* - **may change**. Suffix of the original spectra, *e.g.*, *.dat*.
2. Edit *vmlfit.in*. This is the same as *mlfit.in*.
3. The outputs of this program include four types of files.
  - *vmlfit.[linename].log*. This is the log file recording the run status of this program.
  - *vmlfit.[linename].gss*. Parameters for Gaussian profiles of the best fit. The first line is the title for each column. Here, a Gaussian profile takes the form

$$g(\lambda) = g_0 e^{-(\lambda - \lambda_0)^2 / 2\sigma^2} \quad (\text{B.2})$$

There are eight columns in this file (Table B.6).

Table B.6. Columns in *vmlfit.[linename].gss*

Title	Explanation
SPECFILE	Stem <sup>1</sup> of quasar spectra
N_GSS	Number of Gaussian profiles used
PEAK	Gaussian profile peak flux $g_0$
MEAN	Gaussian profile central wavelength $\lambda_0$
SIGMA	Gaussian profile dispersion $\sigma$
E_PEAK	Uncertainty of $g_0$
E_MEAN	Uncertainty of $\lambda_0$
E_SIGMA	Uncertainty of $\sigma$

Table B.7. Columns in *vmlfit.[linename].par*

Title	Explanation
SPECFILE	Stem of quasar spectra
CHISQ	$\chi^2$ values of line fitting
FWHM(km/s)	Full width of half maximum in km s <sup>-1</sup>
FWHM(A)	Full width of half maximum in Å
LSFT(A)	Emission line shift in Å
E_LSFT(A)	Uncertainty of emission line shift in Å
LSFT(km/s)	Emission line shift in km s <sup>-1</sup>
E_LSFT(km/s)	Uncertainty of emission line shift in km s <sup>-1</sup>
EW(A)	Equivalent width in Å
E_EW(A)	Uncertainty of equivalent width in Å

- *vmlfit.[linename].par*. There are ten columns in this file (see Table B.7).
- *[linename].vml.ps* This file contains four panels.
  - (a) The first plot is the region of the original spectrum around the desired emission line.
  - (b) The panel to the right of the first one is the global continuum subtracted spectrum around the desired emission line.
  - (c) The panel below the first one shows the global continuum subtracted spectrum with the “low pass band filter” signal (in blue). If the local continuum fit is required, it will also display the shape of this local continuum.
  - (d) The last panel displays the emission line fit in blue. The dotted lines are Gaussian components and the dot-dashed lines are the boundary within which the emission line fitting is performed.

By now, both of the continuum and emission line fittings are completed.

## Bibliography

- Abazajian, K., et al. 2003, AJ, 126, 2081  $\checkmark$
- . 2005, AJ, 129, 1755  $\checkmark$
- Abazajian, K. N., et al. 2009, ApJS, 182, 543  $\checkmark$
- Abdo, A. A., et al. 2010, ApJ, 716, 30  $\checkmark$
- Adelman-McCarthy, J. K., et al. 2006, ApJS, 162, 38  $\checkmark$
- . 2007, ApJS, 172, 634  $\checkmark$
- . 2008, ApJS, 175, 297  $\checkmark$
- Akritas, M. G., & Siebert, J. 1996, MNRAS, 278, 919  $\checkmark$
- Anderson, S. F., & Margon, B. 1987, ApJ, 314, 111  $\checkmark$
- Anderson, S. F., et al. 2003, AJ, 126, 2209  $\checkmark$
- Andreas Ostermeier, A. G. . N. H. 1995, Evolutionary Computation, 2, 369  $\checkmark$
- Antonucci, R. 1993, ARA&A, 31, 473  $\checkmark$
- Arav, N., Korista, K. T., Barlow, T. A., & Begelman. 1995, Nature, 376, 576  $\checkmark$
- Arav, N., et al. 2007, ApJ, 658, 829  $\checkmark$
- Armitage, P. J. 2004, in Astrophysics and Space Science Library, Vol. 308, Supermassive Black Holes in the Distant Universe, ed. A. J. Barger, 89+  $\checkmark$
- Arnaud, K. A. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 101, Astronomical Data Analysis Software and Systems V, ed. G. H. Jacoby & J. Barnes, 17+  $\checkmark$
- Arnaud, K. A., et al. 1985, MNRAS, 217, 105  $\checkmark$
- Artymowicz, P., Lin, D. N. C., & Wampler, E. J. 1993, ApJ, 409, 592  $\checkmark$
- Avni, Y., & Tananbaum, H. 1982, ApJ, 262, L17  $\checkmark$
- . 1986, ApJ, 305, 83  $\checkmark$
- Avni, Y., Worrall, D. M., & Morgan, Jr., W. A. 1995, ApJ, 454, 673  $\checkmark$
- Avriel, M. 1993, ACM Trans. Program. Lang. Syst., 15, 745  $\checkmark$
- . 2003, Nonlinear Programming: Analysis and Methods (Dover Publications)  $\checkmark$

$\checkmark$

$-B^2'$   
 $in B^2$

$B$

- Bachev, R., Marziani, P., Sulentic, J. W., Zamanov, R., Calvani, M., & Dultzin-Hacyan, D. 2004, ApJ, 617, 171  $\Sigma$
- Bade, N., Fink, H. H., Engels, D., Voges, W., Hagen, H., Wisotzki, L., & Reimers, D. 1995, A&AS, 110, 469  $\Sigma$
- Baker, J. G., & Menzel, D. H. 1938, ApJ, 88, 52  $\Sigma$
- Baldwin, J., Ferland, G., Korista, K., & Verner, D. 1995, ApJ, 455, L119+  $\Sigma$
- Baldwin, J. A. 1977, ApJ, 214, 679  $\Sigma$
- Baldwin, J. A., Phillips, M. M., & Terlevich, R. 1981, PASP, 93, 5  $\Sigma$
- Baldwin, J. A., Wampler, E. J., & Gaskell, C. M. 1989, ApJ, 338, 630  $\Sigma$
- Ballantyne, D. R. 2010, ApJ, 716, L27  $\Sigma$
- Banzhaf, W., Francone, F. D., Keller, R. E., & Nordin, P. 1998, Genetic programming: an introduction: on the automatic evolution of computer programs and its applications (San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.)  $\Sigma$
- Barlow, T. A., Hamann, F., & Sargent, W. L. W. 1997, in Astronomical Society of the Pacific Conference Series, Vol. 128, Mass Ejection from Active Galactic Nuclei, ed. N. Arav, I. Shlosman, & R. J. Weymann, 13+  $\Sigma$
- Barlow, T. A., & Sargent, W. L. W. 1997, AJ, 113, 136  $\Sigma$
- Barth, A. J., & Shields, J. C. 2000, PASP, 112, 753  $\Sigma$
- Barthelmy, S. D., et al. 2005, Space Sci. Rev., 120, 143  $\Sigma$
- Baskin, A., & Laor, A. 2004, MNRAS, 350, L31  $\Sigma$
- Becker, R. H., White, R. L., & Helfand, D. J. 1995, ApJ, 450, 559  $\Sigma$
- Beckmann, V., Engels, D., Bade, N., & Wucknitz, O. 2003, A&A, 401, 927  $\Sigma$
- Begelman, M. C. 1985, in Astrophysics of Active Galaxies and Quasi-Stellar Objects, ed. J. S. Miller, 411–452  $\Sigma$
- Bennert, V. N., Auger, M. W., Treu, T., Woo, J., & Malkan, M. A. 2011, ApJ, 726, 59  $\Sigma$
- Bennett, A. S. 1962, MNRAS, 125, 75  $\Sigma$
- Beyer, H.-G., & Arnold, D. V. 2006, Evolutionary Computation, 11, 19  $\Sigma$
- Bianchi, S., Guainazzi, M., Matt, G., & Fonseca Bonilla, N. 2007, A&A, 467, L19  $\Sigma$
- Bicknell, G. V., Dopita, M. A., & O'Dea, C. P. O. 1997, ApJ, 485, 112  $\Sigma$
- Bjorck, A. 1996, Numerical Methods for Least Squares Problems (SIAM: Society for Industrial and Applied Mathematics)  $\Sigma$

- Björck, Å. 1996, Numerical Methods for Least Squares Problems (Philadelphia: SIAM)  $J$
- Blandford, R. D., & McKee, C. F. 1982, ApJ, 255, 419  $J$
- Blandford, R. D., & Rees, M. J. 1978, in BL Lac Objects, ed. A. M. Wolfe, 328–341  $\text{in } \beta$
- Boggess, A., et al. 1978, Nature, 275, 372  $J$
- Boroson, T. A., & Green, R. F. 1992, ApJS, 80, 109  $J$
- Brandt, W. N., & Hasinger, G. 2005, ARA&A, 43, 827  $J$
- Brandt, W. N., & Kaspi, S. 2002, in American Institute of Physics Conference Series, Vol. 645, Spectral Line Shapes, ed. C. A. Back, 119–130  $\text{in } \beta$
- Brandt, W. N., Laor, A., & Wills, B. J. 2000, ApJ, 528, 637  $J$
- Brinkmann, W., Laurent-Muehleisen, S. A., Voges, W., Siebert, J., Becker, R. H., Brotherton, M. S., White, R. L., & Gregg, M. D. 2000, A&A, 356, 445  $J$
- Brotherton, M. S., De Breuck, C., & Schaefer, J. J. 2006, MNRAS, 372, L58  $J$
- Brotherton, M. S., Wills, B. J., Steidel, C. C., & Sargent, W. L. W. 1994, ApJ, 423, 131  $J$
- Browne, I. W. A., et al. 2003, MNRAS, 341, 13  $J$
- Buckley, J., & James, I. 1979, Biometrika, 1, 66  $J$
- Burbidge, G., & Hewitt, A. 1992, in Variability of Blazars, 4+  $\beta$
- Burles, S., & Tytler, D. 1998, ApJ, 507, 732  $J$ .
- Burrows, A. 2000, Nature, 403, 727  $J$
- Burrows, D. N., et al. 2005, Space Sci. Rev., 120, 165  $J$
- Cao, X. 2010, ApJ, 725, 388  $J$
- Carswell, R. F., & Smith, M. G. 1978, MNRAS, 185, 381  $J$
- Casebeer, D. A., Leighly, K. M., & Baron, E. 2006, ApJ, 637, 157  $J$
- Cash, W. 1979, ApJ, 228, 939  $J$
- Charlton, J. C. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 291, Hubble's Science Legacy: Future Optical/Ultraviolet Astronomy from Space, ed. K. R. Sembach, J. C. Blades, G. D. Illingworth, & R. C. Kennicutt Jr., 172+  $\text{in } \beta$
- Chartas, G., et al. 2009a, New Astronomy Review, 53, 128  $J$
- . 2009b, New Astronomy Review, 53, 128  $J$
- Cheung, C. C. 2007, AJ, 133, 2097  $J$

Churchill, C. W., Mellon, R. R., Charlton, J. C., Jannuzzi, B. T., Kirhakos, S., Steidel, C. C., & Schneider, D. P. 2000, ApJS, 130, 91 *J*

Clavel, J., et al. 1991, ApJ, 366, 64 *J*

Clemens, M. S., Bressan, A., Nikolic, B., & Rampazzo, R. 2009, MNRAS, 392, L35 *J*

Cohen, M., Wheaton, W. A., & Megeath, S. T. 2003, AJ, 126, 1090 *J*

Collin, S., & Zahn, J. P. 1999, in IAU Symposium, Vol. 194, Activity in Galaxies and Related Phenomena, ed. Y. Terzian, E. Khachikian, & D. Weedman, 246+ *in B*

Collinge, M. J., et al. 2001, ApJ, 557, 2 *J*

—. 2005, AJ, 129, 2542 *J*

Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J. 1998, AJ, 115, 1693 *J*

Congdon, A. B., Keeton, C. R., & Nordgren, C. E. 2010, ApJ, 709, 552 *J*

Corbin, M. R. 1990, ApJ, 357, 346 *J*

Crenshaw, D. M., & Kraemer, S. B. 2001, ApJ, 562, L29 *J*

Crenshaw, D. M., Kraemer, S. B., Boggess, A., Maran, S. P., Mushotzky, R. F., & Wu, C. 1999, ApJ, 516, 750 *J*

Crenshaw, D. M., Kraemer, S. B., & George, I. M. 2003, ARA&A, 41, 117 *J*

Crenshaw, D. M., Maran, S. P., & Mushotzky, R. F. 1998, ApJ, 496, 797 *J*

Cristiani, S., & Vio, R. 1990, A&A, 227, 385 *J*

Croom, S. M., et al. 2002, MNRAS, 337, 275 *J*

—. 2009, MNRAS, 399, 1755 *J*

Crummy, J., Fabian, A. C., Gallo, L., & Ross, R. R. 2006, MNRAS, 365, 1067 *J*

Dave, R., Hernquist, L., Weinberg, D. H., & Katz, N. 1997, ApJ, 477, 21 *J*

De Lucia, G., Springel, V., White, S. D. M., Croton, D., & Kauffmann, G. 2006, MNRAS, 366, 499 *J*

della Ceca, R., Lamorani, G., Maccacaro, T., Wolter, A., Griffiths, R., Stocke, J. T., & Setti, G. 1994, ApJ, 430, 533 *J*

Dempster, A. P., Laird, N. M., & Rubin, D. B. 1977, Royal Stat. Soc. B, 39, 1 *J*

Denney, K. D., et al. 2009, ApJ, 704, L80 *J*

—. 2010, ApJ, 721, 715 *J*

Dietrich, M., Hamann, F., Shields, J. C., Constantin, A., Vestergaard, M., Chaffee, F., Foltz, C. B., & Junkkarinen, V. T. 2002, ApJ, 581, 912  $\checkmark$

Ding, J., Charlton, J. C., Bond, N. A., Zonak, S. G., & Churchill, C. W. 2003, ApJ, 587, 551  $\checkmark$

Ding, J., Charlton, J. C., & Churchill, C. W. 2005, ApJ, 621, 615  $\checkmark$

D'Odorico, V., Cristiani, S., Romano, D., Granato, G. L., & Danese, L. 2004, MNRAS, 351, 976  $\checkmark$

Dong, X., Zhou, H., Wang, T., Wang, J., Li, C., & Zhou, Y. 2005, ApJ, 620, 629  $\checkmark$

Dopita, M. A., & Sutherland, R. S. 1995, ApJ, 455, 468  $\checkmark$

Dunn, J. P., et al. 2010, ApJ, 709, 611  $\checkmark$

Edge, D. O., Shakeshaft, J. R., McAdam, W. B., Baldwin, J. E., & Archer, S. 1959, MmRAS, 68, 37  $\checkmark$

Ekers, J. A. 1969, Australian Journal of Physics Astrophysical Supplement, 7, 3  $\checkmark$

Elvis, M., i. p. 2011  $\checkmark$

Elvis, M. 2000, ApJ, 545, 63  $\checkmark$

Elvis, M., et al. 1994, ApJS, 95, 1  $\checkmark$

Eracleous, M. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 311, AGN Physics with the Sloan Digital Sky Survey, ed. G. T. Richards & P. B. Hall, 183–+  $\checkmark$

Eracleous, M., Hwang, J. A., & Flohic, H. M. L. G. 2010, ApJ, 711, 796  $\checkmark$

Espey, B., & Andreadis, S. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 162, Quasars and Cosmology, ed. G. Ferland & J. Baldwin, 351–+  $\checkmark$

Espey, B. R., Carswell, R. F., Bailey, J. A., Smith, M. G., & Ward, M. J. 1989, ApJ, 342, 666  $\checkmark$

Fabian, A. C. 1999a, Proceedings of the National Academy of Science, 96, 4749  $\checkmark$   
—. 1999b, MNRAS, 308, L39  $\checkmark$

Fabian, A. C., Iwasawa, K., Reynolds, C. S., & Young, A. J. 2000, PASP, 112, 1145  $\checkmark$

Falcke, H., Gopal-Krishna, & Biermann, P. L. 1995, A&A, 298, 395  $\checkmark$

Falcone, A. D., et al. 2004, ApJ, 613, 710  $\checkmark$

Fan, J. H., Guo, Q., Hua, T. X., Liu, Y., Su, J. B., Zhang, Y. W., Zhang, J. Y., & Zhang, J. S. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 373, The Central Engine of Active Galactic Nuclei, ed. L. C. Ho & J.-W. Wang, 193–+  $\checkmark$

in  $\beta$

- Fan, X., et al. 2001, AJ, 122, 2833 *J*  
 —. 2003, AJ, 125, 1649 *J*  
 —. 2004, AJ, 128, 515 *J*  
 —. 2006, AJ, 131, 1203 *J*
- Fanaroff, B. L., & Riley, J. M. 1974, MNRAS, 167, 31P *J*
- Ferland, G. J. 2006, Hazy, A Brief Introduction to Cloudy 06.02, ed. Ferland, G. J. *J*
- Ferland, G. J., Baldwin, J. A., Korista, K. T., Hamann, F., Carswell, R. F., Phillips, M., Wilkes, B., & Williams, R. E. 1996, ApJ, 461, 683
- Ferland, G. J., Korista, K. T., Verner, D. A., Ferguson, J. W., Kingdon, J. B., & Verner, E. M. 1998, PASP, 110, 761
- Ferland, G. J., Peterson, B. M., Horne, K., Welsh, W. F., & Nahar, S. N. 1992, ApJ, 387, 95
- Fields, D. L., Mathur, S., Pogge, R. W., Nicastro, F., Komossa, S., & Krongold, Y. 2005, ApJ, 634, 928
- Fitzpatrick, E. L. 1999, PASP, 111, 63 *J*
- Flohic, H. M. L. G., Eracleous, M., Chartas, G., Shields, J. C., & Moran, E. C. 2006, ApJ, 647, 140
- Foltz, C. B., Weymann, R. J., Peterson, B. M., Sun, L., Malkan, M. A., & Chaffee, Jr., F. H. 1986, ApJ, 307, 504
- Francis, P. J., & Koratkar, A. 1995, MNRAS, 274, 504 *J*
- Frank, J., King, A., & Raine, D. 1992, Accretion power in astrophysics., ed. Frank, J., King, A., & Raine, D. *J*
- Fukugita, M., Ichikawa, T., Gunn, J. E., Doi, M., Shimasaku, K., & Schneider, D. P. 1996, AJ, 111, 1748
- Gabel, J. R., Arav, N., & Kim, T. 2006, ApJ, 646, 742 *J*
- Gabel, J. R., et al. 2005, ApJ, 623, 85 *J*
- Gallagher, S. C., Brandt, W. N., Chartas, G., & Garmire, G. P. 2002a, ApJ, 567, 37 *J*
- Gallagher, S. C., Brandt, W. N., Chartas, G., Garmire, G. P., & Sambruna, R. M. 2002b, ArXiv Astrophysics e-prints *preprint*
- Gallagher, S. C., Brandt, W. N., Chartas, G., Priddey, R., Garmire, G. P., & Sambruna, R. M. 2006, ApJ, 644, 709 *J*

Gallo, L. C., Lehmann, I., Pietsch, W., Boller, T., Brinkmann, W., Friedrich, P., & Grupe, D. 2006, MNRAS, 365, 688

Ganguly, R. et al., i. p. 2011

Ganguly, R., Bond, N. A., Charlton, J. C., Eracleous, M., Brandt, W. N., & Churchill, C. W. 2001, ApJ, 549, 133

Ganguly, R., & Brotherton, M. S. 2008, ApJ, 672, 102

Ganguly, R., Eracleous, M., Charlton, J. C., & Churchill, C. W. 1999, AJ, 117, 2594

Ganguly, R., Masiero, J., Charlton, J. C., & Sembach, K. R. 2003, ApJ, 598, 922

Gaskell, C. M. 1982, ApJ, 263, 79

—. 2009, New Astronomy Review, 53, 140

Gaskell, C. M., & Klimek, E. S. 2003, Astronomical and Astrophysical Transactions, 22, 661

Gehrels, N., et al. 2004, ApJ, 611, 1005

Gelbord, J. M., et al. 2008, in Astronomical Society of the Pacific Conference Series, Vol. 386, Extragalactic Jets: Theory and Observation from Radio to Gamma Ray, ed. T. A. Rector & D. S. De Young, 203–+

George, I. M., Turner, T. J., Yaqoob, T., Netzer, H., Laor, A., Mushotzky, R. F., Nandra, K., & Takahashi, T. 2000, ApJ, 531, 52

Gezari, S., Halpern, J. P., & Eracleous, M. 2007, ApJS, 169, 167

Ghosh, K. K., & Punshy, B. 2007, ApJ, 661, L139

Gibson, R. R., Brandt, W. N., & Schneider, D. P. 2008, ApJ, 685, 773

Gibson, R. R., et al. 2009, ApJ, 692, 758

Gierliński, M., & Done, C. 2004, MNRAS, 349, L7

Gill, P. E., & Murray, W. 1978, SIAM Journal on Numerical Analysis, 15, 977

Giustini, M., Cappi, M., Chartas, G., Eracleous, M., Palumbo, G. G. C., & Vignali, C. 2010, in Astronomical Society of the Pacific Conference Series, Vol. 427, Astronomical Society of the Pacific Conference Series, ed. L. Maraschi, G. Ghisellini, R. Della Ceca, & F. Tavecchio, 108–+

González-Martín, O., Masegosa, J., Márquez, I., Guerrero, M. A., & Dultzin-Hacyan, D. 2006, A&A, 460, 45

Goodman, J., & Tan, J. C. 2004, ApJ, 608, 108

Goodrich, R. W. 1989, ApJ, 342, 224

*J228  
in B22*

- Gopal-Krishna, Kulkarni, V. K., & Mangalam, A. V. 1994, MNRAS, 268, 459 *J*
- Gower, J. F. R., Scott, P. F., & Wills, D. 1967, MmRAS, 71, 49 *J*
- Grandi, S. A. 1982, ApJ, 255, 25 *J*
- Green, P. J. 1996, ApJ, 467, 61 *J*
- . 1998, ApJ, 498, 170 *J*
- Green, P. J., Forster, K., & Kuraszkiewicz, J. 2001, ApJ, 556, 727 *J*
- Greene, J. E., & Ho, L. C. 2005, ApJ, 630, 122 *J*
- Greenstein, J. L., & Schmidt, M. 1964, ApJ, 140, 1 *J*
- Griffith, M. R., Wright, A. E., Burke, B. F., & Ekers, R. D. 1995, ApJS, 97, 347 *J*
- Grupe, D., Komossa, S., Leighly, K. M., & Page, K. L. 2010, ApJS, 187, 64 *J*
- Gunn, J. E., & Peterson, B. A. 1965, ApJ, 142, 1633 *J*
- Hall, P. B., Gallagher, S. C., Richards, G. T., Alexander, D. M., Anderson, S. F., Bauer, F., Brandt, W. N., & Schneider, D. P. 2006, AJ, 132, 1977 *J*
- Hall, P. B., et al. 2002, ApJS, 141, 267 *J*
- Hamann, F. 1997, ApJS, 109, 279 *J*
- . 1998, ApJ, 500, 798 *J*
- Hamann, F., Barlow, T. A., Cohen, R. D., Junkkarinen, V., & Burbidge, E. M. 1997, in Astronomical Society of the Pacific Conference Series, Vol. 128, Mass Ejection from Active Galactic Nuclei, ed. N. Arav, I. Shlosman, & R. J. Weymann, 187–+ *J*
- Hamann, F., Dietrich, M., Sabra, B. M., & Warner, C. 2004, Origin and Evolution of the Elements, 440 *J*
- Hamann, F., & Ferland, G. 1992, ApJ, 391, L53 *J*
- . 1993, ApJ, 418, 11 *J*
- . 1999, ARA&A, 37, 487 *J*
- Hamann, F., Korista, K. T., Ferland, G. J., Warner, C., & Baldwin, J. 2002, ApJ, 564, 592 *J*
- Hamann, F., & Sabra, B. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 311, AGN Physics with the Sloan Digital Sky Survey, ed. G. T. Richards & P. B. Hall, 203–+ *in B*
- Hamann, F. W., Barlow, T. A., Chaffee, F. C., Foltz, C. B., & Weymann, R. J. 2001, ApJ, 550, 142 *J*

Hansen, & Ostermeirer. 1996, in Proceedings of the 1996 IEEE International Conference on Evolutionary Computation, 312–317 C J.

Hansen, M., & Koumoutsakos. 2003, Evolutionary Computation, 11, 1 J.

Hansen, N., & Kern, S. 2004, in Eighth International Conference on Parallel Problem Solving from Nature PPSN VIII, Proceedings, Berlin: Springer, 282–291 C J.

Hansen, N., & Ostermeier, A. 2001, Evolutionary Computation, 9, 159 J.

Hao, L., et al. 2005, AJ, 129, 1795 J.

Hartmann, D., & Burton, W. B. 1997, Atlas of Galactic Neutral Hydrogen, ed. Hartmann, D. & Burton, W. B. J.

Healey, S. E., Romani, R. W., Taylor, G. B., Sadler, E. M., Ricci, R., Murphy, T., Ulvestad, J. S., & Winn, J. N. 2007, ApJS, 171, 61 J.

Heckman, T. M. 1980, A&A, 87, 152 J.

Hewett, P. C., & Wild, V. 2010, MNRAS, 405, 2302 J.

Ho, L. 1999, in Astrophysics and Space Science Library, Vol. 234, Observational Evidence for the Black Holes in the Universe, ed. S. K. Chakrabarti, 157–+ in B J.

Ho, L. C. 1996, in Astronomical Society of the Pacific Conference Series, Vol. 103, The Physics of Liners in View of Recent Observations, ed. M. Eracleous, A. Koratkar, C. Leitherer, & L. Ho, 103–+ in B

Ho, L. C. 2003, in Astronomical Society of the Pacific Conference Series, Vol. 290, Active Galactic Nuclei: From Central Engine to Host Galaxy, ed. S. Collin, F. Combes, & I. Shlosman, 379–+ in B

Hoehn, L., & Niven, I. 1985, Math. Mag., 58, 151 J.

Holweger, H. 2001, in American Institute of Physics Conference Series, Vol. 598, Joint SOHO/ACE workshop "Solar and Galactic Composition", ed. R. F. Wimmer-Schweingruber, 23–30 C J.

Hopkins, P. F., & Hernquist, L. 2009, ApJ, 694, 599 J.

Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Robertson, B., & Springel, V. 2006, ApJS, 163, 1 J.

Hopkins, P. F., et al. 2004, AJ, 128, 1112 J.

Hou, L. G., Wu, X., & Han, J. L. 2009, ApJ, 704, 789 J.

Inada, N., et al. 2010, AJ, 140, 403 J.

Ivanchik, A. V., Petitjean, P., Balashev, S. A., Srianand, R., Varshalovich, D. A., Ledoux, C., & Noterdaeme, P. 2010, MNRAS, 404, 1583 J.

- Iwasawa, K., & Taniguchi, Y. 1993, ApJ, 413, L15 *J*
- Jiang, L., Fan, X., Vestergaard, M., Kurk, J. D., Walter, F., Kelly, B. C., & Strauss, M. A. 2007, AJ, 134, 1150 *J*
- Jiang, P., Wang, J. X., & Wang, T. G. 2006, ApJ, 644, 725 *J*
- Jolley, E. J. D., Kuncic, Z., Bicknell, G. V., & Wagner, S. 2009, MNRAS, 400, 1521 *J*
- Jones, T. M., Misawa, T., Charlton, J. C., Mshar, A. C., & Ferland, G. J. 2010, ApJ, 715, 1497 *J*
- Juarez, Y., Maiolino, R., Mujica, R., Pedani, M., Marinoni, S., Nagao, T., Marconi, A., & Oliva, E. 2009, A&A, 494, L25 *J*
- Just, D. W., Brandt, W. N., Shemmer, O., Steffen, A. T., Schneider, D. P., Chartas, G., & Garmire, G. P. 2007, ApJ, 665, 1004 *J*
- Kalberla, P. M. W., Burton, W. B., Hartmann, D., Arnal, E. M., Bajaja, E., Morras, R., & Pöppel, W. G. L. 2005, A&A, 440, 775 *J*
- Kaspi, S., Brandt, W. N., Maoz, D., Netzer, H., Schneider, D. P., & Shemmer, O. 2007, ApJ, 659, 997 *J*
- Kaspi, S., Smith, P. S., Netzer, H., Maoz, D., Jannuzi, B. T., & Giveon, U. 2000, ApJ, 533, 631 *J*
- Kauffmann, G., & Haehnelt, M. 2000, MNRAS, 311, 576 *J*
- Kellermann, K. I., Sramek, R., Schmidt, M., Shaffer, D. B., & Green, R. 1989, AJ, 98, 1195 *J*
- Kelly, B. C., Bechtold, J., Siemiginowska, A., Aldcroft, T., & Sobolewska, M. 2007, ApJ, 657, 116 *J*
- Kelly, B. C., Vestergaard, M., Fan, X., Hopkins, P., Hernquist, L., & Siemiginowska, A. 2010, ApJ, 719, 1315 *J*
- Kembhavi, A. K., & Narlikar, J. V. 1999, Quasars and active galactic nuclei : an introduction, ed. Kembhavi, A. K. & Narlikar, J. V. *J*
- Kendall, M. G. 1938, Biometrika, 657, 116 *J*
- . 1970, Rank Correlation Methods, ed. Kendall, M. G. *B*
- King, L. J., Browne, I. W. A., Marlow, D. R., Patnaik, A. R., & Wilkinson, P. N. 1999, MNRAS, 307, 225 *J*
- Kinney, A. L., Rivolo, A. R., & Koratkar, A. P. 1990, ApJ, 357, 338 *J*
- Kollatschny, W. 2003, A&A, 407, 461 *J*

*J ~ \beta*  
*B ~ \beta*

- Kollgaard, R. I. 1994, *Vistas in Astronomy*, 38, 29 ↗
- Komossa, S., Voges, W., Xu, D., Mathur, S., Adorf, H., Lemson, G., Duschl, W. J., & Grupe, D. 2006, *AJ*, 132, 531 ↗
- Kondo, Y., Boggess, A., & Maran, S. P. 1989, *ARA&A*, 27, 397 ↗ I -
- Korista, K., Baldwin, J., & Ferland, G. 1998, *ApJ*, 507, 24 ↗
- Korista, K. T., et al. 1995, *ApJS*, 97, 285 ↗
- Kovalev, Y. Y., Kovalev, Y. A., Nizhelsky, N. A., & Bogdantsov, A. B. 2002, *PASA*, 19, 83 ↗
- Kraemer, S. B., Schmitt, H. R., Crenshaw, D. M., Meléndez, M., Turner, T. J., Guainazzi, M., & Mushotzky, R. F. 2011, *ApJ*, 727, 130 ↗
- Kraft, R. P., Burrows, D. N., & Nousek, J. A. 1991, *ApJ*, 374, 344 ↗
- Krolik, J. H., & Kallman, T. R. 1987, *ApJ*, 320, L5 ↗
- Krolik, J. H., Madau, P., & Zycki, P. T. 1994, *ApJ*, 420, L57 ↗
- Kundt, W. 1996, *Ap&SS*, 235, 319 ↗
- La Mura, G., Di Mille, F., Ciroi, S., Popović, L. Č., & Rafanelli, P. 2009, *ApJ*, 693, 1437 ↗
- Lacy, M. 2003, in *Astronomical Society of the Pacific Conference Series*, Vol. 290, Active Galactic Nuclei: From Central Engine to Host Galaxy, ed. S. Collin, F. Combes, & I. Shlosman, 343-+
- Laing, R. A. 1993, in *Astrophysics and Space Science Library*, Vol. 103, *Astrophysics and Space Science Library*, 95-119 ↗
- Laing, R. A., Riley, J. M., & Longair, M. S. 1983, *MNRAS*, 204, 151 ↗
- Laor, A., Bahcall, J. N., Jannuzzi, B. T., Schneider, D. P., & Green, R. F. 1995, *ApJS*, 99, 1 ↗
- Laor, A., Fiore, F., Elvis, M., Wilkes, B. J., & McDowell, J. C. 1997, *ApJ*, 477, 93 ↗
- Lavalley, M., Isobe, T., & Feigelson, E. 1992, in *Astronomical Society of the Pacific Conference Series*, Vol. 25, *Astronomical Data Analysis Software and Systems I*, ed. D. M. Worrall, C. Biemesderfer, & J. Barnes, 245-+ ↗ in B
- Lehmer, B. D., et al. 2007, *ApJ*, 657, 681
- Leighly, K. M., O'Brien, P. T., Edelson, R., George, I. M., Malkan, M. A., Matsuoka, M., Mushotzky, R. F., & Peterson, B. M. 1997, *ApJ*, 483, 767 ↗
- Levenberg, K. 1944, *The Quarterly of Applied Mathematics*, 164 ↗
- Lister, M. L., Hutchings, J. B., & Gower, A. C. 1994, *ApJ*, 427, 125 ↗

g 2 ↗  
in B 2 ↗

- Liu, F. K., & Zhang, Y. H. 2002, A&A, 381, 757 *J*
- Longair, M. S. 1992, High energy astrophysics. Vol.1: Particles, photons and their detection, ed. Longair, M. S. *I*
- . 1994, High energy astrophysics. Vol.2: Stars, the galaxy and the interstellar medium, ed. Longair, M. S. *I*
- Lonsdale, C. J., Farrah, D., & Smith, H. E. 2006, Ultraluminous Infrared Galaxies, ed. Mason, J. W. (Springer Verlag), 285+ *B*
- Lu, Y., Wang, T., Zhou, H., & Wu, J. 2007, AJ, 133, 1615 *J*
- Lupton, R. H., Gunn, J. E., & Szalay, A. S. 1999, AJ, 118, 1406 *J*
- Lusso, E., et al. 2010, A&A, 512, A34+ *J*
- Lynch, R. S., Charlton, J. C., & Kim, T. 2006, ApJ, 640, 81 *J*
- Maccacaro, T., Gioia, I. M., Wolter, A., Zamorani, G., & Stocke, J. T. 1988, ApJ, 326, 680 *J*
- Maiolino, R., & Risaliti, G. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 373, The Central Engine of Active Galactic Nuclei, ed. L. C. Ho & J.-W. Wang, 447+ *in B*
- Maraschi, L., & Rovetti, F. 1994, ApJ, 436, 79 *J*
- Marconi, A., Risaliti, G., Gilli, R., Hunt, L. K., Maiolino, R., & Salvati, M. 2004, MNRAS, 351, 169 *J*
- Marquardt, D. 1963, SIAM Journal on Applied Mathematics, 2, 431 *J*
- Marshall, H. L., Huchra, J. P., Tananbaum, H., Avni, Y., Braccesi, A., Zitelli, V., & Zamorani, G. 1984, ApJ, 283, 50 *J*
- Mason, K. O., et al. 1996, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 2808, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. O. H. Siegmund & M. A. Gummin, 438–447 *in B*
- Massaro, E., Giommi, P., Leto, C., Marchegiani, P., Maselli, A., Perri, M., Piranomonte, S., & Sclavi, S. 2009, A&A, 495, 691 *J*
- Mathews, W. G., & Ferland, G. J. 1987, ApJ, 323, 456 *J*
- Mathur, S., Elvis, M., & Wilkes, B. 1995, ApJ, 452, 230 *J*
- Matthews, T. A., & Sandage, A. R. 1963, ApJ, 138, 30 *J*
- McDowell, J. C., Canizares, C., Elvis, M., Lawrence, A., Markoff, S., Mathur, S., & Wilkes, B. J. 1995, ApJ, 450, 585 *J*

*in B*  
*in B*  
*in B*

- McIntosh, D. H., Rix, H., Rieke, M. J., & Foltz, C. B. 1999, ApJ, 517, L73 J
- McLure, R. J., & Dunlop, J. S. 2004, MNRAS, 352, 1390 J
- McLure, R. J., & Jarvis, M. J. 2002, MNRAS, 337, 109 J
- Meusinger, H., Hinze, A., & de Hoon, A. 2011, A&A, 525, A37+ J
- Miller, B. P., Brandt, W. N., Gibson, R. R., Garmire, G. P., & Shemmer, O. 2009, ApJ, 702, 911 J
- Misawa, T., Charlton, J. C., Eracleous, M., Ganguly, R., Tytler, D., Kirkman, D., Suzuki, N., & Lubin, D. 2007, ApJS, 171, 1 J
- Misawa, T., Eracleous, M., Charlton, J. C., & Tajitsu, A. 2005, ApJ, 629, 115 J
- Misawa, T., Eracleous, M., Chartas, G., & Charlton, J. C. 2008, ApJ, 677, 863 J
- More, J. J., & Wright, S. J. 1993, Optimization Software Guide (Philadelphia, PA, USA: Society for Industrial and Applied Mathematics) J
- Moretti, A., et al. 2005, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 5898, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, ed. O. H. W. Siegmund, 360–368 C
- Mukai, K. 1993, Legacy, vol. 3, p.21-31, 3, 21 J
- Murray, N., Chiang, J., Grossman, S. A., & Voit, G. M. 1995, ApJ, 451, 498 J
- Nandra, K., George, I. M., Mushotzky, R. F., Turner, T. J., & Yaqoob, T. 1997, ApJ, 488, L91+ J
- Narayanan, D., Hamann, F., Barlow, T., Burbidge, E. M., Cohen, R. D., Junkkarinen, V., & Lyons, R. 2004, ApJ, 601, 715 J
- Nemmen, R. S., Bower, R. G., Babul, A., & Storchi-Bergmann, T. 2007, MNRAS, 377, 1652 J
- Nenkova, M., Sirocky, M. M., Nikutta, R., Ivezić, Ž., & Elitzur, M. 2008, ApJ, 685, 160 J
- Netzer, H., Laor, A., & Gondhalekar, P. M. 1992, MNRAS, 254, 15 J
- O'Brien, P. T., et al. 1998, ApJ, 509, 163 J
- Oke, J. B., Shields, G. A., & Korycansky, D. G. 1984, ApJ, 277, 64 J
- Onken, C. A., & Peterson, B. M. 2002, ApJ, 572, 746 J
- Osmmer, P. S., Porter, A. C., & Green, R. F. 1994, ApJ, 436, 678 J
- Osmmer, P. S., & Shields, J. C. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 162, Quasars and Cosmology, ed. G. Ferland & J. Baldwin, 235–+ JNB

$\Sigma^2 \Sigma^0$   
 $\Sigma^2 \Sigma^1$   
 $\Sigma^2 \Sigma^2$   
 $\Sigma^2 \Sigma^3$   
 $\Sigma^2 \Sigma^4$

- Osmers, P. S., & Smith, M. G. 1976, ApJ, 210, 267 *I*
- Osterbrock, D. E. 1981, ApJ, 249, 462 *J*
- Osterbrock, D. E., & Ferland, G. J. 1989, Astrophysics of gaseous nebulae and active galactic nuclei: 2 edition, ed. Osterbrock, D. E. (University Science Books) *B*
- . 2006, Astrophysics of gaseous nebulae and active galactic nuclei, ed. Osterbrock, D. E. & Ferland, G. J. *B*
- Osterbrock, D. E., & Pogge, R. W. 1985, ApJ, 297, 166 *J*
- Padovani, P., Giommi, P., Landt, H., & Perlman, E. S. 2007, ApJ, 662, 182 *J*
- Page, K. L., O'Brien, P. T., Reeves, J. N., & Turner, M. J. L. 2004a, MNRAS, 347, 316 *J*
- Page, K. L., Reeves, J. N., O'Brien, P. T., & Turner, M. J. L. 2005, MNRAS, 364, 195 *J*
- Page, K. L., Reeves, J. N., O'Brien, P. T., Turner, M. J. L., & Worrall, D. M. 2004b, MNRAS, 353, 133 *J*
- Papovich, C., Norman, C. A., Bowen, D. V., Heckman, T., Savaglio, S., Koekemoer, A. M., & Blades, J. C. 2000, ApJ, 531, 654 *J*
- Park, S. Q., et al. 2010, ApJ, 717, 1181 *J*
- Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H. 2002, AJ, 124, 266 *J*
- . 2010, AJ, 139, 2097 *K*
- Peterson, B. M. 1993, PASP, 105, 247 *J*
- . 1997, An Introduction to Active Galactic Nuclei, ed. Peterson, B. M. (Cambridge University Press) *B*
- Peterson, B. M., & Wandel, A. 2000, ApJ, 540, L13 *J*
- Peterson, B. M., et al. 2004, ApJ, 613, 682 *J*
- Petitjean, P., Rauch, M., & Carswell, R. F. 1994, A&A, 291, 29 *J*
- Petitjean, P., Riediger, R., & Rauch, M. 1996, A&A, 307, 417 *J*
- Phillips, M. M. 1993, ApJ, 413, L105 *J*
- Pickering, T. E., Impey, C. D., & Foltz, C. B. 1994, AJ, 108, 1542 *J*
- Pilkington, J. D. H., & Scott, P. F. 1965, MmRAS, 69, 183 *J*
- Plotkin, R. M., Anderson, S. F., Hall, P. B., Margon, B., Voges, W., Schneider, D. P., Stinson, G., & York, D. G. 2008, AJ, 135, 2453 *J*
- Pogge, R. W. 1989, ApJ, 345, 730 *J*

*J ~ 20*  
*J ~ 3*

- Pogge, R. W., & Peterson, B. M. 1992, AJ, 103, 1084 *S*
- Pogson, N. 1857, MNRAS, 18, 47 *S*
- Poole, T. S., et al. 2008, MNRAS, 383, 627 *S*
- Poutanen, J. 1998, in Theory of Black Hole Accretion Disks, ed. M. A. Abramowicz, G. Bjornsson, & J. E. Pringle, 100--+ *in B*
- Proga, D., Stone, J. M., & Kallman, T. R. 2000, ApJ, 543, 686 *S*
- Rani, B., Gupta, A. C., Joshi, U. C., Ganesh, S., & Wiita, P. J. 2010, ApJ, 719, L153 *S*
- Rauch, M. 1998, ARA&A, 36, 267 *S*
- Reichard, T. A., et al. 2003, AJ, 126, 2594 *S*
- Reichert, G. A., et al. 1994, ApJ, 425, 582 *S*
- Reynolds, C. S. 1997, MNRAS, 286, 513 *S*
- Reynolds, C. S., & Nowak, M. A. 2003, Physics Report, 377, 389 *S*
- Richards, G. T. 2001, ApJS, 133, 53 *S*
- Richards, G. T., Vanden Berk, D. E., Reichard, T. A., Hall, P. B., Schneider, D. P., SubbaRao, M., Thakar, A. R., & York, D. G. 2002a, AJ, 124, 1 *S*
- Richards, G. T., et al. 2002b, AJ, 123, 2945
- . 2003, AJ, 126, 1131 *X*
- . 2006, ApJS, 166, 470 *F*
- . 2011, AJ, 141, 167 *T*
- Richstone, D. O., & Schmidt, M. 1980, ApJ, 235, 361 *S*
- Rieger, F. M., & Aharonian, F. A. 2008, A&A, 479, L5 *S*
- Rodriguez-Pascual, P. M., et al. 1997, ApJS, 110, 9 *S*
- Romano, D., Silva, L., Matteucci, F., & Danese, L. 2002, MNRAS, 334, 444 *S*
- Roming, P. W. A., et al. 2005, Space Sci. Rev., 120, 95 *S*
- . 2009, ApJ, 690, 163 *X*
- Rosenbrock, H. H. 1960, Computer J., 3, 175 *S*
- Saez, C., Chartas, G., & Brandt, W. N. 2009, ApJ, 697, 194 *S*
- Sakata, Y., et al. 2010, ApJ, 711, 461 *S*

*S<sup>2</sup> B<sup>2</sup> / in B<sup>2</sup>*

- Sambruna, R. M., Maraschi, L., & Urry, C. M. 1996, ApJ, 463, 444 *S*
- Sambruna, R. M., et al. 2006, ApJ, 646, 23 *S*
- Sanders, D. B., & Mirabel, I. F. 1996, ARA&A, 34, 749 *S*
- Sanders, D. B., Phinney, E. S., Neugebauer, G., Soifer, B. T., & Matthews, K. 1989, ApJ, 347, 29 *S*
- Scarpa, R., & Falomo, R. 1997, A&A, 325, 109 *S*
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525 *S*
- Schmidt, M. 1968, ApJ, 151, 393 *S*
- Schmidt, M. 1969, in Quasars and high-energy astronomy, ed. K. N. Douglas, I. Robinson, A. Schild, E. L. Schucking, J. A. Wheeler, & N. J. Woolf, 55+ *S*
- . 1970, ApJ, 162, 371 *S*
- Schmidt, M., & Green, R. F. 1983, ApJ, 269, 352 *S*
- Schmitt, H. R., & Kinney, A. L. 1996, ApJ, 463, 498 *S*
- Schneider, D. P., et al. 2002, AJ, 123, 567 *S*
- . 2003, AJ, 126, 2579 *X*
- . 2005, AJ, 130, 367 *X*
- . 2007, AJ, 134, 102 *X*
- . 2010, AJ, 139, 2360 *S*
- Seyfert, C. K. 1943, ApJ, 97, 28 *S*
- Shang, Z., Wills, B. J., Wills, D., & Brotherton, M. S. 2007, AJ, 134, 294 *S*
- Shang, Z., et al. 2005, ApJ, 619, 41 *S*
- Shao, J., & Tu, D. 1995 *S*
- Shemmer, O., Brandt, W. N., Netzer, H., Maiolino, R., & Kaspi, S. 2006a, ApJ, 646, L29 *S*
- . 2008, ApJ, 682, 81 *X*
- Shemmer, O., Brandt, W. N., Vignali, C., Schneider, D. P., Fan, X., Richards, G. T., & Strauss, M. A. 2005, ApJ, 630, 729 *S*
- Shemmer, O., Netzer, H., Maiolino, R., Oliva, E., Croom, S., Corbett, E., & di Fabrizio, L. 2004, ApJ, 614, 547 *S*
- Shemmer, O., et al. 2006b, ApJ, 644, 86 *S*
- S 2 / a*  
*in B 2 /*

- Shen, J., Vanden Berk, D. E., Schneider, D. P., & Hall, P. B. 2008a, AJ, 135, 928  
Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., & Schneider, D. P. 2008b, ApJ, 680, 169

Shen, Y., et al. 2010, ArXiv e-prints

Shields, G. A. 1978, Nature, 272, 706

—. 1996, ApJ, 461, L9+ ~~X~~

Shields, J. C. 2007, in Astronomical Society of the Pacific Conference Series, Vol. 373, The Central Engine of Active Galactic Nuclei, ed. L. C. Ho & J.-W. Wang, 355+ *in B*

Skrutskie, M. F., et al. 2006, AJ, 131, 1163 *J*

Small, T. A., & Blandford, R. D. 1992, MNRAS, 259, 725 *J*

Smith, H. J., & Hoffleit, D. 1963, Nature, 198, 650 *J*

Smith, P. S., Williams, G. G., Schmidt, G. D., Diamond-Stanic, A. M., & Means, D. L. 2007, ApJ, 663, 118 *J*

Srianand, R., Petitjean, P., Ledoux, C., & Hazard, C. 2002, MNRAS, 336, 753 *J*

Steffen, A. T., Strateva, I., Brandt, W. N., Alexander, D. M., Koekemoer, A. M., Lehmer, B. D., Schneider, D. P., & Vignali, C. 2006, AJ, 131, 2826

Stocke, J. T., Morris, S. L., Weymann, R. J., & Foltz, C. B. 1992, ApJ, 396, 487 *J*

Storn, R., & Price, K. 1997, Journal of Global Minimization, 11, 341 *J*

Stoughton, C., et al. 2002, AJ, 123, 485 *J*

Strateva, I. V., Brandt, W. N., Eracleous, M., Schneider, D. P., & Chartas, G. 2006, ApJ, 651, 749 *J*

Strateva, I. V., Brandt, W. N., Schneider, D. P., Vanden Berk, D. G., & Vignali, C. 2005, AJ, 130, 387 *J*

Sulentic, J. W., Marziani, P., & Dultzin-Hacyan, D. 2000, ARA&A, 38, 521 *J*

Takahashi, T., et al. 2010, in Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, Vol. 7732, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series *C*

- Tananbaum, H., et al. 1979, ApJ, 234, L9 J
- Telfer, R. C., Kriss, G. A., Zheng, W., Davidsen, A. F., & Green, R. F. 1998, ApJ, 509, 132 J
- Telfer, R. C., Zheng, W., Kriss, G. A., & Davidsen, A. F. 2002, ApJ, 565, 773 J
- Torres, D. F., & Anchordoqui, L. A. 2004, Reports on Progress in Physics, 67, 1663 J
- Treves, A., Maraschi, L., & Abramowicz, M. 1988, PASP, 100, 427 J
- Trimble, V., & Barton, E. 2007, PASP, 119, 1208 J
- Tripp, T. M., Lu, L., & Savage, B. D. 1996, ApJS, 102, 239 J
- Trump, J. R., et al. 2006, ApJS, 165, 1 J
- . 2009, ApJ, 700, 49 J
- Tueller, J., Mushotzky, R. F., Barthelmy, S., Cannizzo, J. K., Gehrels, N., Markwardt, C. B., Skinner, G. K., & Winter, L. M. 2008, ApJ, 681, 113 J
- Turner, T. J., & Pounds, K. A. 1989, MNRAS, 240, 833 J
- Turnshek, D. A. 1986, in IAU Symposium, Vol. 119, Quasars, ed. G. Swarup & V. K. Kapahi, 317–328 J
- Turnshek, D. A. 1988, in QSO Absorption Lines: Probing the Universe, ed. J. C. Blades, D. A. Turnshek, & C. A. Norman, 17+ in B
- Tytler, D., & Fan, X. 1992, ApJS, 79, 1 J
- Ueda, Y., et al. 2007, ApJ, 664, L79 J
- Urry, C. 2004, in Astronomical Society of the Pacific Conference Series, Vol. 311, AGN Physics with the Sloan Digital Sky Survey, ed. G. T. Richards & P. B. Hall, 49+ in B
- Urry, C. M., & Padovani, P. 1995, PASP, 107, 803 J
- Uttley, P., & McHardy, I. M. 2004, Progress of Theoretical Physics Supplement, 155, 170 J
- Vagnetti, F., & Spera, R. 1994, ApJ, 436, 611 J
- Vanden Berk, D. E., et al. 2001, AJ, 122, 549 J
- . 2004, ApJ, 601, 692 X
- . 2006, AJ, 131, 84 X
- Vanden Berk, D. E. et al., i. p. 2011 J
- Vasudevan, R. V., & Fabian, A. C. 2007, MNRAS, 381, 1235 J

*3220  
~V  
~B*

Veilleux, S. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 284, IAU Colloq. 184: AGN Surveys, ed. R. F. Green, E. Y. Khachikian, & D. B. Sanders, 111–+

in B

Veilleux, S., & Osterbrock, D. E. 1987, ApJS, 63, 295

✓

Vernaleo, J. C., & Reynolds, C. S. 2006, ApJ, 645, 83

✓

Véron-Cetty, M., Joly, M., & Véron, P. 2004, A&A, 417, 515

✓

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

✓

Vestergaard, M., & Wilkes, B. J. 2001, ApJS, 134, 1

✓

Vignali, C., Brandt, W. N., & Schneider, D. P. 2003, AJ, 125, 433

✓

Vignali, C., Brandt, W. N., Schneider, D. P., & Kaspi, S. 2005, AJ, 129, 2519

✓

Voges, W., et al. 1999, A&A, 349, 389

✓

—. 2000, IAUC, 7432, 3

✗

Wagner, S. J., & Witzel, A. 1995, ARA&A, 33, 163

✓

Waldram, E. M., Pooley, G. G., Grainge, K. J. B., Jones, M. E., Saunders, R. D. E., Scott, P. F., & Taylor, A. C. 2003, MNRAS, 342, 915

✓

Walter, R., & Fink, H. H. 1993, A&A, 274, 105

✓

✓  
in B

Wampler, E. J. 1980, in IAU Symposium, Vol. 92, Objects of High Redshift, ed. G. O. Abell & P. J. E. Peebles, 119–121

Wampler, E. J., Chugai, N. N., & Petitjean, P. 1995, ApJ, 443, 586

✓

Wandel, A., Peterson, B. M., & Malkan, M. A. 1999, ApJ, 526, 579

✓

Wanders, I., et al. 1997, ApJS, 113, 69

✓

Wang, J., Chen, Y., Ho, L. C., & McLure, R. J. 2006, ApJ, 642, L111

✓

Wang, T., Dong, X., Zhang, X., Zhou, H., Wang, J., & Lu, Y. 2005, ApJ, 625, L35

✓

Warner, C., Hamann, F., & Dietrich, M. 2004, ApJ, 608, 136

✓

Watanabe, S., et al. 2009, ApJ, 694, 294

✓

Weaver, K. A., et al. 1992, ApJ, 401, L11

✓

Weedman, D. W., et al. 2005, ApJ, 633, 706

✓

Weymann, R. J., Morris, S. L., Foltz, C. B., & Hewett, P. C. 1991, ApJ, 373, 23

✓

White, R. L., Becker, R. H., Helfand, D. J., & Gregg, M. D. 1997, ApJ, 475, 479

✓

✓  
in B  
✓

- Wilhite, B. C., Vanden Berk, D. E., Brunner, R. J., & Brinkmann, J. V. 2006, ApJ, 641, 78 J
- Wilhite, B. C., Vanden Berk, D. E., Kron, R. G., Schneider, D. P., Pereyra, N., Brunner, R. J., Richards, G. T., & Brinkmann, J. V. 2005, ApJ, 633, 638 J
- Wilkes, B. J. 1986, MNRAS, 218, 331 es
- Wilkes, B. J., Kuraszkiewicz, J., Green, P. J., Mathur, S., & McDowell, J. C. 1999, ApJ, 513, 76 J
- Wilkes, B. J., Tananbaum, H., Worrall, D. M., Avni, Y., Oey, M. S., & Flanagan, J. 1994, ApJS, 92, 53 J
- Williams, L. L. R., Foley, P., Farnsworth, D., & Belter, J. 2008, ApJ, 685, 725 J
- Wills, B. J., Netzer, H., & Wills, D. 1985, ApJ, 288, 94 J
- Winter, L. M., Mushotzky, R. F., Reynolds, C. S., & Tueller, J. 2009, ApJ, 690, 1322 J
- Wise, J. H., Eracleous, M., Charlton, J. C., & Ganguly, R. 2004, ApJ, 613, 129 J
- Wolf, C., Wisotzki, L., Borch, A., Dye, S., Kleinheinrich, M., & Meisenheimer, K. 2003, A&A, 408, 499 J
- Wolf, C., et al. 2004, A&A, 421, 913 J
- Worrall, D. M. 2009, A&AR, 17, 1 J
- Worrall, D. M., & Birkinshaw, M. 2006, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 693, Physics of Active Galactic Nuclei at all Scales, ed. D. Alloin, 39+ in progress
- Worrall, D. M., Tananbaum, H., Giommi, P., & Zamorani, G. 1987, ApJ, 313, 596 J
- Wu, J., Charlton, J. C., Misawa, T., Eracleous, M., & Ganguly, R. 2010, ApJ, 722, 997 J
- Wu, J., Vanden Berk, D. E., Brandt, W. N., Schneider, D. P., Gibson, R. R., & Wu, J. 2009, ApJ, 702, 767 J
- Wu, J., et al. 2011 J
- Wu, X., & Liu, F. K. 2004, ApJ, 614, 91 J
- Yip, C. W., et al. 2004, AJ, 128, 2603 J
- York, D. G., et al. 2000, AJ, 120, 1579 J
- Young, M., Elvis, M., & Risaliti, G. 2009, ApJS, 183, 17 J
- Yu, Q., Lu, Y., & Kauffmann, G. 2005, ApJ, 634, 901 J
- Yu, Q., & Tremaine, S. 2002, MNRAS, 335, 965 J

$\int \sim \sim \sim$   
 $\partial \sim \beta \sim \backslash$

Yuan, Q., Green, R. F., Brotherton, M., Tripp, T. M., Kaiser, M. E., & Kriss, G. A.  
2002, ApJ, 575, 687

J

Zakamska, N. L., et al. 2003, AJ, 126, 2125

J

Zamorani, G., Marano, B., Mignoli, M., Zitelli, V., & Boyle, B. J. 1992, MNRAS, 256,  
238

J

Zamorani, G., et al. 1981, ApJ, 245, 357

J

Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., & Davidsen, A. F. 1997, ApJ, 475,  
469

J

Zheng, W., & Malkan, M. A. 1993, ApJ, 415, 517

J

Zheng, W., & Sulentic, J. W. 1990, ApJ, 350, 512

J

Zhou, H., Wang, T., Wang, H., Wang, J., Yuan, W., & Lu, Y. 2006a, ApJ, 639, 716

J

Zhou, H., Wang, T., Yuan, W., Lu, H., Dong, X., Wang, J., & Lu, Y. 2006b, ApJS, 166,  
128

Zhou, X., & Wang, J. 2005, ApJ, 618, L83

J

J J J J

# Vita

## Jian Wu

### Education

**The Pennsylvania State University** University Park, PA, US 2004–2011  
Ph.D. in Astronomy & Astrophysics, expected in August, 2011  
Area of Specialization: Astronomy&Astrophysics  
Ph.D. Minor in Computational Science, expected in August, 2011

**University of Science & Technology of China** Hefei, Anhui, P.R.China 2000–2004  
B.S. in Physics & Astrophysics

### Awards and Honors

Zaccheus Daniel Fund 2009  
Zaccheus Daniel Fund 2007  
Stephen B. Brumbach Fellowship 2006  
USTC Excellent Graduate Student Award 2004

### Research Experience

**Doctoral Research** The Pennsylvania State University 2006–Present  
Thesis Advisor: Prof. Donald P. Schneider, Dr. Daniel E. Vanden Berk  
Continuum and emission line properties of quasars in UV, optical and X-ray bands  
using ground-based and space observations.

**Graduate Research** The Pennsylvania State University 2004–2006  
Research Advisor: Prof. Jane C. Charlton  
Modeling intrinsic narrow absorption line systems of quasars.

**Undergraduate Research** University of Science & Technology of China 2003–2004  
Research Advisor: Prof. Ting-gui Wang & Dr. Hong-yan Zhou  
Selection effect of radio quasars from SDSS quasar catalog

### Teaching Experience

**Teaching Assistant** The Pennsylvania State University 2004–2006  
Grading, proctoring and answering questions for students in ASTRO001, ASTRO291,  
ASTRO292. Lecturing for ASTRO11 (Astronomy Lab).