
**A comprehensive analysis
of optical and near-infrared spectroscopy
of $z \sim 6$ quasars**

Gisella De Rosa
Max-Planck-Institut für Astronomie

Heidelberg 2011

Dissertation in Astronomy
submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola-University of Heidelberg, Germany,
for the degree of
Doctor of Natural Sciences

Put forward by
M.Sc. *Gisella De Rosa*
born in Vasto (CH), Italy

Oral examination: 1.12.11

**A comprehensive analysis
of optical and near-infrared spectroscopy
of $z \sim 6$ quasars**

Gisella De Rosa
Max-Planck-Institut für Astronomie

**Referees: Prof. Dr. Hans-W. Rix
Prof. Dr. Jochen Heidt**

Abstract

High redshift ($z \sim 6$) quasars (QSOs) are unique probes of the early growth of accreting supermassive black holes (BHs) and are fundamental tools to study the ionization state of the intergalactic medium and its chemical enrichment. Until now, only the most luminous QSOs have been studied, often one object at the time. In this thesis we present the most extended consistent analysis to date of optical and near-infrared spectra of $z \sim 6$ QSOs. Our NIR sample is composed of 22 spectra of QSOs at $4.5 < z < 6.45$, covering the Mg II and Fe II emission features, and includes new spectroscopic VLT-ISAAC observations of three faint $z \sim 6$ SDSS QSOs. The optical sample consists of 26 high quality spectra of QSOs at $5.70 < z < 6.45$, covering the Ly α complex, and the O I+Si II, C II and Si IV+O IV] broad lines.

Based on the NIR spectra we estimate the BH mass (M_{BH}), the Eddington ratio and the Fe II/Mg II line ratio, a proxy for the chemical enrichment of the BLR. The QSOs in our sample host BHs with masses of $\sim 10^9 M_{\odot}$ that are accreting significantly faster than a luminosity-matched comparison sample at lower redshifts. The measured Fe II/Mg II line ratios show no evolution with cosmic time in the redshift range $4.5 < z < 6.45$, implying that the QSOs in our sample have undergone a major episode of Fe enrichment in the few 100s Myr preceding the cosmic age at which they are observed.

An analysis of the wavelength shifts of the optical lines shows that the redshifts inferred from the O I+Si II complex, visible in most of the optical spectra, are consistent with those obtained from the Mg II line. This implies that z_{OI} is a reliable proxy for the QSO systemic redshift. We create a composite template of the $z \sim 6$ QSOs that, red-ward the Ly α line, remarkably resembles the one obtained for low-redshift QSOs. Investigating the existence of possible correlations amongst optical emission properties, we do not detect the expected anti-correlation between the EW and the QSO continuum luminosity (*Baldwin effect*) for the high ionization lines (N V & Si IV+O IV]). We show that a strong correlation holds between the line FWHM and its luminosity. Given the narrow distribution of Eddington ratios that characterizes the $z \sim 6$ QSOs population, this implies that QSOs with brighter lines are the ones that host more massive BHs. We further confirm this statement by testing the relation between the line luminosity and M_{BH} .

Finally we use our optical spectra to study the ionization state of the IGM at high redshift, by measuring the evolution of the *near zone* (NZ) sizes and that of the Gunn Peterson (GP) effective optical depth. Our comprehensive analysis shows that the correlation between the NZ radii and the QSO systemic redshifts is significantly reduced with respect to previous results. This implies that changes of the IGM ionization state are difficult to constrain through such measurements. We also outline a new method to detect the evolution of the GP effective optical depth through the analysis of the Ly α and Ly β flux ratios. For each QSO we build a model based on the SDSS QSO composite spectrum and on the optical depth parametrization of lower-redshift absorbers ($z < 5.7$). Comparing the QSOs flux ratios with the ones obtained from the models we detect a steep increase in the evolution of the effective optical depth.

Zusammenfassung

Hoch rotverschobene ($z \sim 6$) Quasare (QSOs) stellen einzigartige Proben des frühen Anwachsens von akkretierenden supermassereichen Schwarzen Löchern (BHs) dar und sind essentiell, um den Ionisationsgrad des Intergalaktischen Mediums (IGM) und dessen chemische Anreicherung zu studieren. Bis heute konnten nur die allerhellsten QSOs studiert werden, zumeist nur in individuellen Studien. In dieser Arbeit präsentieren wir die zum heutigen Zeitpunkt umfassendste einheitliche Analyse von optischen und nah-infraroten (NIR) Spektren von $z \sim 6$ QSOs. Unser NIR Datensatz umfasst 22 Spektren von QSOs bei $4.5 < z < 6.45$, welcher die Mg II und Fe II Emissionslinien beinhaltet, inklusive neuer spektroskopischer VLT-ISAAC Beobachtungen von drei leuchtschwachen $z \sim 6$ SDSS QSOs. Der optische Datensatz besteht aus 26 hoch qualitativen Spektren von QSOs bei $5.70 < z < 6.45$, welcher den Ly α Komplex und die O I+Si II, C II und Si IV+O IV] Emissionsbanden abdeckt. Basierend auf den NIR Spektren schätzen wir die BH Masse (M_{BH}), den Eddington Faktor und das Fe II/Mg II Linienverhältnis ab, welches eine Abschätzung der chemischen Anreicherung der 'Broad Line Region' ermöglicht. Die QSOs in unserem Datensatz beinhalten BHs mit Massen von $\sim 10^9 M_{\odot}$ welche signifikant schneller Materie akkretieren als ein Vergleichsdatsatz ähnlicher Helligkeit bei niedrigerer Rotverschiebung. Das gemessene Fe II/Mg II Linienverhältnis zeigt keine Veränderung mit kosmischer Zeit in dem Rotverschiebungsbereich $4.5 < z < 6.45$. Dies legt nahe, dass die QSOs in unserem Datensatz eine Episode von rascher Fe Anreicherung durchlaufen haben. Eine Untersuchung der Wellenlängenverschiebung des in den meisten optischen Spektren sichtbaren O I+Si II Komplex liefert übereinstimmende Werte mit denen der Mg II Linie, die typischerweise eine verlässliche Bestimmung der Rotverschiebung von QSOs erlaubt. Wir erstellen ein 'Template' für ein $z \sim 6$ QSO Spektrum, welches für Wellenlängen länger als die von Ly α bemerkenswerterweise den Spektren von QSOs geringerer Rotverschiebung ähnelt. Eine Untersuchung zum Vorhandensein möglicher Korrelationen zwischen optischen Emissionsbanden ergibt, dass wir die erwartete Antikorrelation (*Baldwin Effekt*) zwischen Äquivalentbreite (EW) und QSO Kontinuumshelligkeit von hoch-ionisierten Linien (N V & Si IV+O IV]) nicht bestätigen können. Dafür zeigen wir, dass eine starke Korrelation zwischen Linienbreite (FWHM) und Linienhelligkeit besteht. Die geringe Streuung der Eddington Faktoren, welche für die $z \sim 6$ QSO Population typisch ist, impliziert dass QSOs mit helleren Linien die Objekte sind, welche BHs grösserer Masse beheimaten. Darüberhinaus können wir diese Aussage durch einen Vergleich von Linienhelligkeit und M_{BH} bestätigen. Schließlich nutzen wir unsere optischen Spektren, um den Ionisationsgrad des IGM bei hoher Rotverschiebung zu studieren, indem wir die Evolution der Größe der 'Near-Zone' (NZ) und die zeitliche Änderung der Gunn Peterson (GP) optischen Tiefe messen. Unsere erkenntnisreiche Analyse ergibt, dass die Korrelation zwischen NZ Radien und systemischen QSO Rotverschiebungen im Vergleich zu früheren Resultaten signifikant schwächer ist. Dies impliziert, dass Änderungen des IGM Ionisationsgrads nur schwer durch solche Messungen einzugrenzen sind. Wir zeigen zudem eine neue Methode auf, um die Evolution der GP effektiven optischen Tiefe über die Analyse des Ly α und Ly β Linienflussverhältnisses zu ermitteln. Mit Hilfe des SDSS QSO Kompositspektrums und der optischen Tiefe-Parametrisierung weniger rotverschobener ($z < 5.7$) Absorber modellieren wir jeden QSO einzeln. Der Vergleich der beobachteten QSO Linienflussverhältnisse mit denjenigen, die wir den Modellen entnehmen, ergibt ein starkes Anwachsen in der Entwicklung der effektiven optischen Tiefe.

CONTENTS

1	INTRODUCTION	1
1.1	QUASAR PHENOMENOLOGY	1
1.1.1	THE DISCOVERY	1
1.1.2	WHAT IS POWERING QSQS?	2
1.1.3	AGN ZOOLOGY	2
1.1.4	UNIFIED MODEL FOR AGN	4
1.2	COSMOLOGY	5
1.2.1	FRIEDMAN EQUATIONS	5
1.2.2	COSMOLOGICAL DISTANCES	6
1.2.3	THE HISTORY OF THE UNIVERSE IN A NUTSHELL	7
1.3	THIS WORK	8
2	EVIDENCE FOR NON-EVOLVING FeII/MgII RATIOS IN RAPIDLY ACCRETING HIGH REDSHIFT QSOs	11
2.1	INTRODUCTION	12
2.2	SAMPLE, OBSERVATIONS AND DATA REDUCTION	14

2.2.1	NEW DATA	14
2.2.1.1	Data reduction	15
2.2.2	LITERATURE DATA	16
2.3	DATA ANALYSIS	17
2.3.1	POWER-LAW	17
2.3.2	BALMER PSEUDO-CONTINUUM	19
2.3.3	Fe II TEMPLATE	19
2.3.4	FITTING PROCEDURE	20
2.4	RESULTS	21
2.4.1	BLACK HOLE MASS ESTIMATE	21
2.4.2	DEPENDENCE ON THE S/N AND ON THE SYSTEMATICS IN THE OBSERVATIONS	25
2.4.3	COMPARISON WITH SDSS QSO SAMPLE AT $0.35 < z < 2.25$	25
2.4.4	COMPARISON WITH LUMINOSITY MATCHED SDSS QSO SAMPLE	28
2.4.5	IMPACT ON BH FORMATION HISTORY	32
2.4.6	Fe II/Mg II	35
2.5	SUMMARY AND CONCLUSIONS	36
2.A	FIT RESULTS: LITERATURE DATA	39
3	OPTICAL PROPERTIES OF HIGH REDSHIFT QUASARS	51
3.1	A SPECTROSCOPIC SAMPLE OF $z \sim 6$ QSOs	52
3.2	DATA ANALYSIS	55
3.2.1	CORRECTING FOR GALACTIC EXTINCTION	55
3.2.2	SPECTRAL MODEL	56
3.2.3	PROCEDURE FOR LINE FITTING	57
3.2.4	SYSTEMIC REDSHIFT	59

3.2.5	SUMMARY OF THE FITTING PROCEDURE	60
3.2.6	FIT RESULTS AND NOTES ON INDIVIDUAL SOURCES	60
3.3	COMPOSITE SPECTRUM	62
3.4	REDSHIFT ESTIMATE	69
3.5	EMISSION PROPERTIES	77
3.5.1	CONTINUUM LUMINOSITY	77
3.5.2	BLACK HOLE MASS	81
3.6	SUMMARY AND CONCLUSIONS	84
4	IONIZATION NEAR ZONES AND GUNN PETERSON ABSORPTION	93
4.1	IONIZATION NEAR ZONES	94
4.1.1	TESTING THE DEFINITION OF RNZ	97
4.1.2	RNZ AS A FUNCTION OF THE QSO SYSTEMIC REDSHIFT	99
4.2	ABSORPTION FROM NEUTRAL HYDROGEN	104
4.2.1	BUILDING A MODEL	106
4.3	FLUX RATIOS	107
4.4	SUMMARY AND CONCLUSIONS	110
5	SUMMARY AND FUTURE PERSPECTIVES	113
	ACKNOWLEDGMENTS	123
	BIBLIOGRAPHY	127

1.1 QUASAR PHENOMENOLOGY

Quasars are amongst the most luminous, powerful objects in the universe: they radiate from 10 to 100,000 times the energy radiated by our Galaxy. For this reason we are able to observe them up to very high redshifts.

1.1.1 THE DISCOVERY

Quasars were already observed during the first half of the 20th century as radio-sources, but their nature remained unclear for decades. In the late 50s, radio observations revealed that these sources were characterized by very small angular sizes: they were star-like objects, or *quasi stellar radio sources*, later contracted into *quasars*. The optical counter-parts of some of these radio sources were observed for the first time in the 60s. Their star-like nature was quickly contradicted by their atypical spectral properties. The optical spectra were not resembling any known type of star and presented broad emission lines at wavelengths not corresponding to any known atomic transition. The extragalactic nature of these radio sources was discovered only in the 1963, by Maarten Schmidt, who identified lines of the Balmer series in the spectrum of 3C 273. The spectral lines presented the same relative distances expected for the hydrogen lines, but were red-shifted ($z = 0.16$) with respect to the reference transition wavelengths. Interpreting this red-shift as entirely due to the Doppler effect, Schmidt realized that this source was moving away from our Galaxy at a velocity of $\sim 48000 \text{ km s}^{-1}$. Since such a velocity was too high for any Galactic object, he concluded that this quasar was an extra-galactic source. 3C 273, with a redshift of $z = 0.16$ was the most distant source known at that time. After 3C 273, the extragalactic nature was confirmed for many other quasars. Later on, many more of these sources

were discovered and observed throughout the full electromagnetic spectrum, and astronomers realized that only $\sim 10\%$ of them showed emission at radio wavelengths. As a consequence the name was changed from quasars to *quasi stellar objects* or QSOs (today, both names are usually used equivalently). Given their bright apparent magnitudes and their large distances, we can conclude that QSOs are characterized by an enormous emitting power, up to 100,000 times higher than the power radiated by the entire Milky Way.

1.1.2 WHAT IS POWERING QSOs?

The most important clue regarding the nature of the central engine that powers QSOs is their rapid time variability. Quasars can show significant variations in their luminosity on very short timescales, going from months to hours. Variability on such short time scales implies that the radiation is emitted from very compact regions. To obtain coherent variations in the luminosity of a source of a given size, the information must propagate throughout its extension. Since the information cannot travel faster than the speed of light, the maximum size of an object varying with some characteristic timescale is equal to the distance covered by the light during that time. As a consequence, the QSO radiation must be emitted from regions of the order of light-hours to light-months. The high luminosities originating from such small regions can only be explained by violent and exotic phenomena. The dominant explanation is that QSOs are the most luminous manifestation of the class of objects known as Active Galactic Nuclei (AGN): accreting super-massive black holes (BHs) hosted in the center of galaxies (Rees 1984). As matter spirals-in towards a BH through an accretion disk, a substantial fraction of its gravitational energy can be released as viscosity converts gravitational and kinetic energy into heat and radiation. The accretion of matter through a disk around a black hole is an extremely efficient way to release large amounts of radiation.

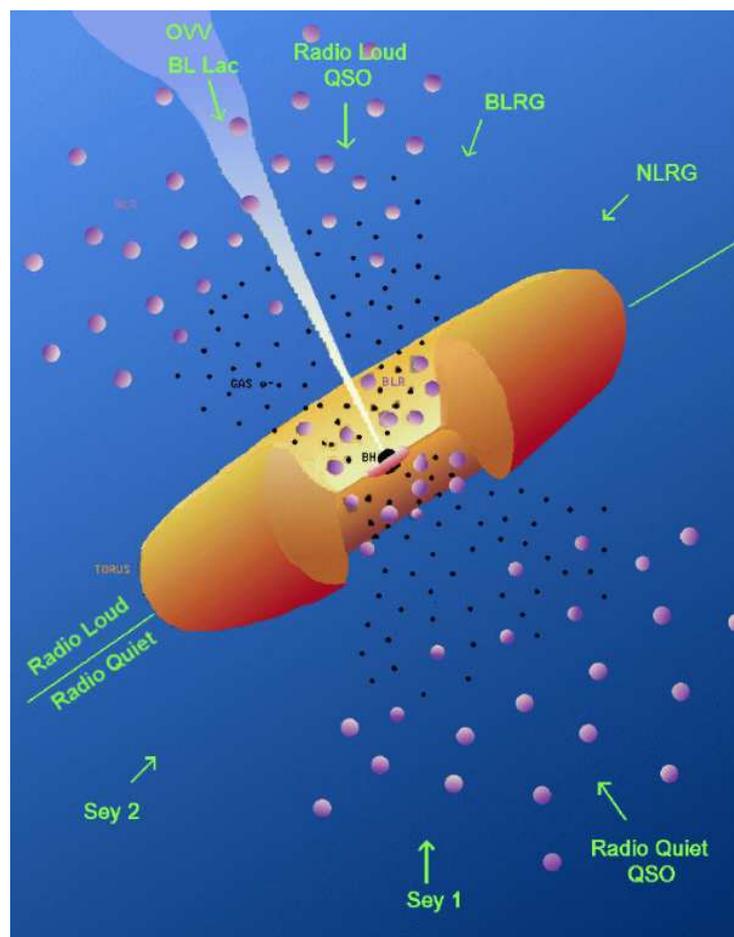
1.1.3 AGN ZOOLOGY

AGN emit radiation throughout the whole electromagnetic spectrum. Their rest-frame UV-optical spectra are characterized by a non-stellar continuum that can be parametrized as a double, broken power-law (Vanden Berk et al. 2001), very broad emission lines ($\sigma \gtrsim 1000 \text{ km s}^{-1}$) and narrow emission lines ($\sigma \sim 500 \text{ km s}^{-1}$). Many sub-classes of AGN are identified, according to their observed photometric and spectral characteristics. **QSOs** are defined as AGN having an absolute magnitude in the visual band $M_V < -23 \text{ mag}$. Objects weaker than this reference magnitude are instead known as **Seyfert galaxies**. Concerning the spectral features instead, AGN are classified as:

- **Type I:** broad and narrow emission lines and bright (rest-frame) UV-optical continuum;
- **Type II:** only narrow emission lines and weaker (rest-frame) UV-optical continuum;
- **LINERS** (Low Ionization Nuclear Emission Line Regions): only low-ionization narrow emission lines and weak (rest-frame) UV-optical continuum.

The properties of rapid variability and high degree of polarization at visible wavelengths define instead the class of AGN known as **blazars**.

Figure 1.1: Figure re-adapted from Urry & Padovani (1995): illustration of the AGN structure. The continuum emission is originated in the accretion disk surrounding the central BH, while the broad and narrow emission lines are originated by clouds of cold gas orbiting around the central engine. Different projections result in different observed photometric and spectral characteristics.



1.1.4 UNIFIED MODEL FOR AGN

Even though AGN are characterized by a large variety of phenomena, their main characteristics can be interpreted through a unified model (Antonucci 1993; Urry & Padovani 1995): a single structure coupled with a line-of-sight effect (see fig. 1.1). The different structural regions in the standard AGN model are:

- **Accretion disk:** hot gas orbiting around a rapidly rotating supermassive BH. It is responsible for the AGN continuum emission in the (rest-frame) optical, UV, and soft X-rays. The size of the accretion disk can be estimated through variability arguments and it is of the order of milli-parsecs.
- **Broad Line Region:** cold gas near the accretion disk orbiting around the BH. The broad (permitted and semi-forbidden) emission lines originate from this region. Their width is mainly due to Doppler broadening produced by the high rotational velocities. From reverberation mapping studies (e.g. Peterson et al. 2004) we know that this region is a hundred times larger than the accretion disk. The geometry and dynamical properties of the BLR are still under debate (see e.g., Decarli et al. 2011).
- **Dusty torus:** donut-like structure (1–10 pc) made of warm and hot dust. The dust absorbs the inner accretion disk radiation and re-emits it in the infrared as a black body. Since this torus obscures the central parts of the AGN over a wide solid angle, it holds a key role in the unification theory. For instance, depending on the viewing angle it could shield the light coming from the inner accretion disk and from the broad lines, which are emitted *inside* the torus, thus originating the Type-I/II dichotomy.
- **Narrow Line Region:** cold gas at large distances from the central BH (1–1000 pc). The larger distances result in a much lower rotational velocity. The narrow (permitted and forbidden) lines are originated in this region.
- **Jets:** highly collimated streams of energetic particles originating from a region close to the accretion disc. The high level of collimation is most probably reached thanks to the presence of strong magnetic fields. They can propagate for distances larger than the QSO host galaxy size (~ 1 Mpc). Jets emit in the whole electromagnetic spectrum, but they are brighter at radio wavelength (synchrotron radiation by accelerated particles).

If such a structure is observed with a line of sight close to the polar axis, avoiding the obscuring torus, the object will appear as a Type I AGN. On the other hand if the line of sight is intercepting the obscuring torus, the object will present only narrow line features, appearing as a Type II AGN. Finally, objects observed with a line of sight intercepting the polar jets will appear as BLAZARs.

In this latter case, the continuum emission is amplified by relativistic effects and over-shines the emission lines. The actual size of the different regions is a function of the BH mass. This geometric unification is still a matter of debate.

1.2 COSMOLOGY

From the present collection of empirical evidence the favored model describing the universe and its evolution is the Λ CDM model. This model explains the observed accelerated expansion of the universe, as well as the formation of structures at high redshift and the power spectrum of the observed fluctuations in the CMB. In this section we introduce some fundamental cosmological concepts that will be used throughout this work (Peebles 1993; Hogg 1999).

1.2.1 FRIEDMAN EQUATIONS

The cosmological principle states that, on large-scales, the universe is homogeneous and isotropic. The metric describing a universe with these characteristics is the Robertson-Walker metric (1933):

$$ds^2 = c^2 dt^2 - R^2(t) \left(\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2(\theta) d\phi^2) \right), \quad (1.1)$$

where ds is the line element and c is the speed of light. $[t, r, \theta, \phi]$ are the co-moving coordinates, i.e. coordinates fixed with respect to the matter flow, so that a particle does not change its comoving coordinates as the universe expands or contracts. $R(t)$ is the scale parameter or scale factor and describes the evolution (expansion or contraction) of the spatial part of the universe $([r, \theta, \phi])$. k is the curvature parameter: positive values of the curvature parameter correspond to an open universe, while negative values correspond to a closed universe. The flat minkowskian universe corresponds to $k = 0$.

In general relativity the gravitational field is described by the Einstein's equations as the curvature of the space-time due to the matter-energy density distribution. Combining Einstein's equations and the Robertson-Walker metric, one can obtain the Friedmann-Lemaître equations, that regulate the dynamics of the universe:

$$\left(\frac{\dot{R}}{R} \right)^2 = \frac{\Lambda c^2}{3} - \frac{k c^2}{3} + \frac{8\pi G}{3} \rho \quad (1.2)$$

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2} \right) + \frac{\Lambda c^2}{3}, \quad (1.3)$$

where G is the gravitational constant and Λ is the cosmological constant. These equations of state of the universe relate pressure (p) and density (ρ) through the scale parameter $R(t)$. The ratio between the cosmological time derivative of the scale parameter and itself, is the Hubble parameter:

$$H(t) = \frac{\dot{R}(t)}{R(t)}. \quad (1.4)$$

$H(t)$ has the units of the inverse of a time and gives an estimate of the expansion rate. At the present epoch ($t = t_0$), H_0 is known as the Hubble constant and it is conventionally parametrized as $H_0 = 100 h^{-1} \text{ km s}^{-1} \text{ Mpc}^{-1}$, with $0.6 < h < 0.9$ (from observational constraints). The inverse of the hubble constant ($1/H_0 = 9.78 \times 10^9 h^{-1} \text{ yr}$) is the Hubble time and it represents an estimate of the age of the universe (assuming a constant rate of expansion equal to the one measured today). Multiplying the first Friedmann's equation by $(1/H^2)$ and considering the present time ($t = 0$) one can obtain that:

$$\Omega_\Lambda + \Omega_k + \Omega_M = 1 \quad (1.5)$$

where $\Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2}$ is the cosmological constant density parameter, $\Omega_k = -\frac{kc^2}{3R_0^2 H_0^2}$ is the curvature density parameter and $\Omega_M = \frac{8\pi G}{3H_0^2} \rho_0$ is the matter density parameter. Throughout this thesis we will follow the concordance model for a Λ CDM universe and fix the density parameters to the following values: $\Omega_\Lambda = 0.7$, $\Omega_M = 0.3$, $\Omega_k = 0$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ (e.g. Spergel et al. 2007)

1.2.2 COSMOLOGICAL DISTANCES

GRAVITATIONAL REDSHIFT

The gravitational redshift is caused by the expansion of the space. The wavelength increases as it traverses the expanding universe between its point of emission and its point of detection by the same amount that space has expanded during the crossing time:

$$z + 1 = \frac{\lambda_{obs}}{\lambda_{em}} = \frac{R(t_{obs})}{R(t_{em})}. \quad (1.6)$$

CO-MOVING DISTANCE

The *proper distance* is the radial distance (ideally measurable with a ruler) between the observer and the emitting source, measured along a geodesic at the time the light was emitted. The *co-*

moving distance is the proper distance between observer and source at the present time, i.e. is the proper distance at the time the light was emitted times the ratio of scale factors now to then:

$$D_C = c \int_0^z \frac{dz'}{H(z')}. \quad (1.7)$$

ANGULAR DIAMETER DISTANCE

The *angular diameter distance* D_A is defined as the ratio between the physical transverse size of a source, D , and its angular size, θ , in radians:

$$D_A = \frac{D}{\theta}. \quad (1.8)$$

It does not increase indefinitely as $z \rightarrow \infty$: it starts decreasing since $z \sim 1$ and thereafter more distant objects actually appear larger in angular size.

LUMINOSITY DISTANCE

The *luminosity distance* D_L is defined by the relationship between bolometric flux F and bolometric luminosity L :

$$D_L = \sqrt{\frac{L}{4\pi F}}. \quad (1.9)$$

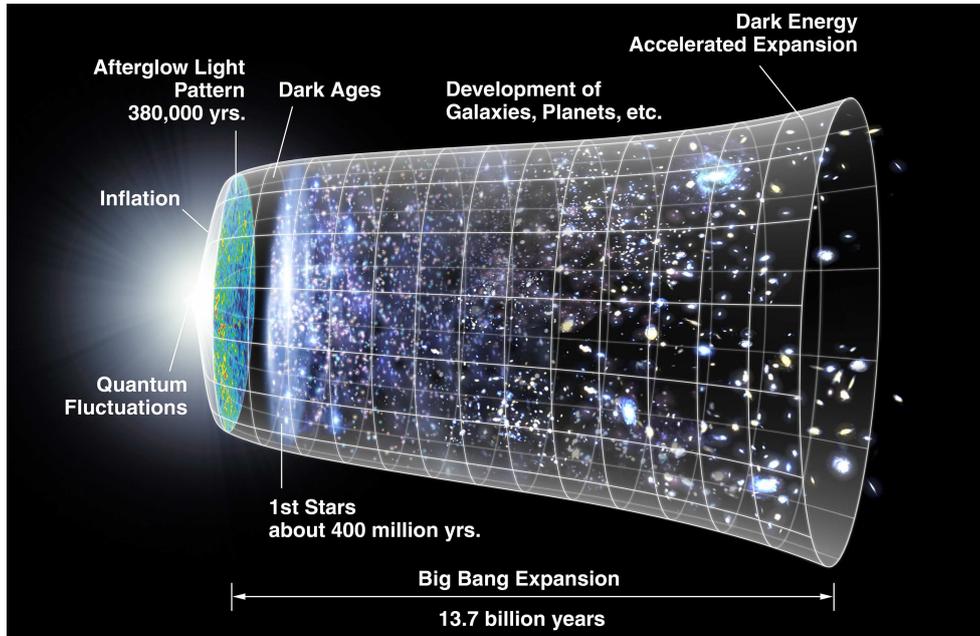
It is related to the angular diameter distance in the following way:

$$D_L = (1 + z)^2 D_A. \quad (1.10)$$

1.2.3 THE HISTORY OF THE UNIVERSE IN A NUTSHELL

The currently accepted cosmological model states that the universe has originated from a configuration in which all the matter was concentrated in a singularity. At $t=0$ (~ 13.7 Gyr ago), with the Big Bang, the universe began its expansion that is still ongoing (see fig. 1.2).. After $t \sim 300,000$ years, corresponding to a redshift $z \sim 1000$, the temperature of the universe dropped below the point at which hydrogen atoms were no longer ionized by the electromagnetic radiation. At that point, given the drastic drop in the number of free electrons, matter and radiation decoupled and started evolving independently. The radiation continued cooling down due to the universe expansion and we now observe it as the Cosmic Microwave Background (CMB). Matter instead started forming the first structures under its gravitational pull.

Figure 1.2: An artist’s view of the history of the universe. At $t=0$ (~ 13.7 Gyr ago), with the Big Bang, the universe began its expansion that is still ongoing. At redshift $z \sim 1000$ matter and radiation decoupled and the universe became opaque. With the formation of the first objects that were able to emit energetic radiation, the universe started to re-ionize.



At the moment of decoupling the universe became opaque to its own radiation. Therefore we are not able to probe this era, known as *the dark ages*. With the formation of the first objects that were able to emit energetic radiation, the universe started to re-ionize. How and when the reionization took place is still a matter of debate, as studying sources at the end of the dark ages means pushing the limits of today’s observational cosmology. First, metal-pure galaxies and quasars at $z \sim 10$ were probably main actors in this process. QSOs at $z \sim 6 - 7$ (1 Gyr after the Big Bang) are probing the era at which the universe became transparent again.

1.3 THIS WORK

In the past 10 years, more than 60 QSOs at redshift $5.7 < z < 6.5$ have been discovered thanks to large multivalength surveys such as the Sloan Digital Sky Survey (SDSS). Only recently, the $z = 6.5$ redshift barrier has been broken and the first $z \sim 7$ QSO has been discovered in the near-infrared (NIR) UKIDSS large area survey (ULAS). Studying the absorption systems along the line of sight of bright high redshift QSOs, we can put constraints on the cosmological evolution of the inter galactic medium (IGM) ionization state and of its chemical enrichment. At the same

time, high-redshift QSOs are key probes in understanding the formation of the AGN structure and in defining the nature of the link between the evolution of the central BH and that of the host galaxy. Detailed studies of QSOs require spectra covering the highest possible number of key-emission features. Moreover, to be able to put stringent constraints on such complex and unknown processes, the analysis needs to be performed on statistically significant samples, covering a wide range of QSO properties (luminosities, accretion rates, BH masses).

At $z \sim 6$ key diagnostic UV emission features are redshifted to the optical and NIR bands. Until now, only the most luminous QSOs have been studied, often one object at the time. In this thesis we present the most extended consistent analysis up to date of optical and NIR spectra of $z \sim 6$ QSOs. The thesis is organized as follows:

- in chapter 2 we present the analysis of 21 NIR spectra of QSOs at redshift $4 < z < 6.5$, covering the Mg II and Fe II emission features. The latter are powerful probes of the BH mass and of the chemical enrichment of the BLR. Three of the spectra are presented in this thesis for the first time and are extending the existing SDSS sample towards the faint end of the QSO luminosity function;
- in chapter 3 we study the optical emission properties of a sample of QSOs at redshifts $5.7 < z < 6.45$, analyzing 26 high quality spectra covering the Ly α complex (Ly α , N V, Si II), and the O I+Si II, C II and Si IV+O IV] broad lines;
- in chapter four we analyze the evolution of the ionization state of the IGM at high redshifts through measurements of the Gunn Peterson effective optical depth and of the *near zone* sizes;
- in chapter five we summarize our work and outline future perspectives of studying the highest redshift QSOs in the universe.

EVIDENCE FOR NON-EVOLVING Fe II/Mg II RATIOS IN RAPIDLY ACCRETING $z \sim 6$ QSOs*

CONTEXT

Quasars (QSOs) at the highest known redshifts ($z \geq 5.7$) are unique probes of the universe less than 1 Gyr after the Big Bang. They are fundamental for studying the early growth of supermassive black holes (BHs), the galaxy formation and the interstellar medium chemical evolution (e.g. [Kauffmann & Haenelt 2000](#); [Wyithe & Loeb 2003](#); [Hopkins et al. 2005](#)). Strong emission features excited by the central engine and its immediate surroundings can be used to infer properties of the powering BH and of the circumnuclear gas. Under the assumption that the broad-line region (BLR) dynamics is dominated by the central gravitational field, it is possible to estimate the BH mass (M_{BH}) using the velocity of the line-emitting gas and its distance from the central BH. Information about the bulk velocity of the BLR gas can be obtained by measurements of the emission line width. The distance of the emitting gas from the central BH (R) can instead be inferred from the continuum luminosity (L): from reverberation mapping studies of local active galactic nuclei (AGNs) we know that $R \propto L^{0.5}$ ($R - L$ relation, e.g. [Kaspi et al. 2000](#)). Measurements of the luminosity and of the gas velocity, via Doppler-broadened line-width, can thus be used to determine the M_{BH} from single epoch spectra. From the line flux ratios we can also derive chemical abundances of the BLR gas that are fundamental to set constraints on the star formation history of the QSO host galaxy. The abundance of Fe and α elements (e.g. O, Mg, Ne that are produced in the stars via α processes) is of particular interest for understanding the chemical evolution of galaxies at high redshift. Most of Fe in the solar neighborhoods has been produced via the explosions of type Ia supernovae (SNe Ia) while α elements are mainly produced by core collapse supernovae (SNe of types II, Ib and Ic). SNeIa are thought to originate

*This chapter is adapted from the paper [De Rosa et al. \(2011\)](#)

from intermediate-mass stars in close binary systems, characterized by long life-times, while core collapse SNe come from more massive stars which explode very soon after the initial starburst. In this picture the Fe/α ratio is expected to be a strong function of age in young systems.

2.1 INTRODUCTION

Under the assumption that the dynamics of the BLR is dominated by the central gravitational field, it is possible to estimate M_{BH} by using the velocity and the distance of the line-emitting gas from the central BH : $M_{\text{BH}} \propto R v^2$. From reverberation mapping studies of local active galactic nuclei (AGNs) it has been found that R is related to the continuum luminosity L ($R - L$ relation, e.g. Kaspi et al. 2000): $R \propto L^{0.5}$. Measurements of the luminosity and of the gas velocity, via Doppler-broadened line-width, can thus be used to determine the M_{BH} from single epoch spectra. Various strong emission lines can be used as M_{BH} estimators: $\text{H}\alpha$, $\text{H}\beta$, Mg II , C IV (for a review on the single-epoch spectrum method see Peterson 2010). For objects in the local universe, the $R - L$ relation and its intrinsic scatter have been investigated in detail using the $\text{H}\beta$ line and the relative continuum. These studies have shown that one can obtain accurate M_{BH} estimates via the $\text{H}\beta$ line. For sources at high- z a complication is that the $\text{H}\beta$ emission line is redshifted out of the visible window already at modest redshifts. Whether or not a particular UV line can be used to estimate M_{BH} depends on how well the respective UV line widths in AGN spectra are correlated. In the case of Mg II Shen et al. (2008), using a large collection of SDSS spectra, found that $\log([\text{FWHM}(\text{H}\beta)] / [\text{FWHM}(\text{Mg II})]) = 0.0062$, with a scatter of only 0.11 dex, suggesting that the Mg II can be used as a proxy for the $\text{H}\beta$ line, and thus for the BH mass. On the other hand, the accuracy achievable in the determination of the BH mass from the C IV line remains somewhat controversial, since C IV is a resonance line and absorptions are often detected in its blue wing due to outflows, which in turn affects the line width. With the single-epoch spectrum method it has been possible to estimate the M_{BH} for $z \sim 6$ QSOs (Barth et al. 2003; Willott et al. 2003; Jiang et al. 2007; Kurk et al. 2007, 2009; Willott et al. 2010). These studies have shown that these high- z sources from SDSS host BHs with $M_{\text{BH}} \sim 10^9 M_{\odot}$ and are accreting close to the Eddington limit ($L_{\text{bol}}/L_{\text{Edd}} \sim 1$). However, the $z \sim 6$ QSOs from SDSS only account for the brightest end of the QSO luminosity function. Only few faint QSOs ($z'_{\text{AB}} > 20$) have measured M_{BH} to date: two from the SDSS deep Stripe 82 (Jiang et al. 2008, 2009; Kurk et al. 2007, 2009), a deep imaging survey obtained by repeatedly scanning a stripe (260 deg^2) along the celestial equator, and nine from CFHQS (Willott et al. 2010). These sources are powered by less massive BHs (down to $M_{\text{BH}} \sim 10^8 M_{\odot}$) that also appear to accrete close to the Eddington limit. For QSOs at lower redshifts Shen et al. (2008) found that the most luminous QSOs ($L_{\text{bol}} > 10^{47} \text{ erg s}^{-1}$) at redshift $2 < z < 3$ are characterized by an Eddington ratio of only $L_{\text{bol}}/L_{\text{Edd}} \sim 0.25$ with a dispersion of 0.23 dex. It seems therefore that high-redshift QSOs have fundamentally different

properties than the lower-redshift ones (see also Trakhtenbrot et al. 2011).

Regarding the chemical enrichment, photoionization models show that various emission line ratios are good estimators for the metallicity of the BLR gas (e.g. Hamann et al. 2002; Nagao et al. 2006). It is possible to estimate the BLR gas metallicity and to set constraints on the BLR enrichment history through measurements of the relative abundances of nitrogen (N) respect to carbon (C) and helium (He), since these three elements are formed by different astrophysical processes and on different timescales (Hamann & Ferland 1993): C is rapidly produced in the explosion of massive stars; N is a second generation element, i.e. slowly produced in stars from previously synthesized C and O; He is a primordial element and its abundance does not change significantly with cosmic time. Using these element ratios Jiang et al. (2007) estimated the BLR metallicity for a sample of six luminous QSOs with $5.8 < z < 6.3$, finding super solar metallicities with a typical value of $\sim 4 Z_{\odot}$ and no strong evolution up to $z \sim 6$.

The abundance of Fe and α elements (e.g. O, Mg, Ne that are produced in the stars via α processes) is of particular interest for understanding the chemical evolution of galaxies at high- z . Most of Fe in the solar neighborhoods has been produced via the explosion of type Ia supernovae (SNe Ia) while α elements such as Mg and O are mainly produced by core collapse supernovae (SNe of types II, Ib and Ic). SNeIa are thought to originate from intermediate-mass stars in close binary systems, characterized by long life-times, while core collapse SNe come from more massive stars which explode very soon after the initial starburst. A time delay between α elements and the Fe enrichment is thus expected. This delay depends on the Initial Mass Function (IMF) and on the galactic star formation history, and can vary from 0.3 Gyr, for massive elliptical galaxies, to 1-3 Gyr, for Milky-Way type galaxies (Matteucci & Recchi 2003). In this picture the Fe/ α ratio is expected to be a strong function of age in young systems. The observational proxy that is usually adopted to trace the Fe/Mg abundance ratio for the BLR gas is the Fe II/Mg II line ratio. Unfortunately this line ratio is only a second order proxy, since it depends not only on the actual abundance of Fe but also significantly on the excitation conditions that determine how strongly Fe II lines are emitted (Baldwin et al. 2004). However, even though to date no calibration is available to convert the observed relative line strengths into actual abundance ratios, the study of the Fe II/Mg II line ratio as a function of look-back time does in itself carry significant information about the BLR chemical enrichment history.

For $z > 5.7$ QSOs the Fe II and the Mg II lines are redshifted into the near-infrared (NIR). Numerous NIR-spectroscopy studies of QSO samples including $z \sim 6$ sources have been carried out in the past (e.g. Maiolino et al. 2001, 2003; Pentericci et al. 2002; Iwamuro et al. 2002, 2004; Barth et al. 2003; Dietrich et al. 2003; Freudling et al. 2003; Willott et al. 2003; Jiang et al. 2007; Kurk et al. 2007, 2009). Their results show an increase in the scatter of the measured Fe II/Mg II line ratios as a function of the redshift. One possible way to interpret the increase in the scatter is that some objects are observed such a short time after the initial starburst that the BLR is not yet fully

enriched with Fe (Iwamuro et al. 2004). Nonetheless, several possible systematics related to the adopted fitting procedure could potentially be the cause for the observed discrepancies between different authors: the iron template employed, the wavelength limits over which the template is integrated, the wavelength range over which the template is actually fitted and the wavelength coverage and S/N of the spectra used.

Our goal in this chapter is to study BH masses, Eddington ratios and to characterize the Fe II and Mg II emission in a large sample of high- z QSOs, by fitting their respective NIR spectra in a coherent and homogeneous way. With the aim of extending the existing SDSS sample of $z \sim 6$ QSOs towards the faint end of the QSO luminosity function, we observed three additional $z \sim 6$ sources (discovery papers: Jiang et al. 2008, De Rosa et al. in prep) with the Infrared Spectrometer And Array Camera (ISAAC) mounted on the Very Large Telescope (VLT). The three targets have $19.6 < z'_{AB} < 20.7$ and have been selected from the SDSS and the SDSS Stripe 82. The observed NIR spectra include the Mg II and Fe II emission lines. We have also collected literature spectra of $z > 4$ QSOs covering the restframe wavelengths between $2700 \text{ \AA} < \lambda < 3200 \text{ \AA}$, characterized by the presence of the Mg II line doublet ($\lambda = 2796, 2803 \text{ \AA}$) and of the Fe II line forest.

2.2 SAMPLE, OBSERVATIONS AND DATA REDUCTION

Our sample is composed of 22 targets: 3 new spectra observed with VLT-ISAAC (see tab. 2.1) and 19 sources from the literature (see tab. 2.2), kindly provided by the respective authors. Ten of the literature sources have redshift $4.50 < z < 5.70$ (Iwamuro et al. 2002), while the remaining 9 have $5.70 < z < 6.43$ and z -band magnitudes $z'_{AB} < 20.9$ (magnitudes are taken from the discovery papers).

2.2.1 NEW DATA

We have observed 3 SDSS QSOs with magnitudes $19.6 < z'_{AB} < 20.7$ and redshifts $6.05 < z < 6.08$ (see tab. 2.1). The two faintest ones have been selected from the SDSS Stripe 82 and extend the existing sample towards the faint end of the QSO luminosity function. The observations were carried out with ISAAC on Antu (VLT-UT1) in low resolution mode (LR), using the 1024x1024 Hawaii Rockwell array of the Short Wavelength arm. For each QSO the Mg II line and the Fe II complex were observed: given the redshift of the sources, these features fall in the K band. The selected slit had a width of 1" and combined with the order selection filter it gives a spectral resolution $\lambda/\Delta\lambda \sim 450$. Table 2.1 summarizes the exposure time for each object. For each ob-

Table 2.1: Summary of the new sources.

QSO name ¹	RA ²	DEC ³	z ⁴	z_{AB}^* ⁵	K ⁶	t_{exp} ⁷	Ref. ⁸
SDSS J0353 + 0104	03 53 49.72	+01 04 04.4	6.05	20.5	17.81	4.60 h	J08
SDSS J2054 – 0005	20 54 06.49	–00 05 41.8	6.06	20.7	17.87	5.92 h	J08
SDSS J0842 + 1218	08 42 29	+12 18 50.5	6.08	19.6	17.67	4.60 h	DRip

¹ QSO name;

^{2,3} QSO right ascension and declination in J2000.0 –from discovery papers;

⁴ optical redshift –from discovery papers;

⁵ z_{AB}^* magnitude –from discovery papers;

⁶ K band nominal magnitude, obtained from the redshifted SDSS QSO template scaled to match the observed J and H-band magnitudes – from discovery papers;

⁷ total on-source exposure time;

⁸ discovery papers: DRip, De Rosa et al. (in preparation); J08, Jiang et al. (2008)

servation block (OB), sixteen frames of 148 seconds were taken following an ABBA dithering pattern (with large offsets among the dithered positions: from 20'' to 30''). Further, small random offsets within a box of 4'' to 18'' were applied at each dithered position in order to avoid pixel-related artifacts (jittering). Given the faintness of the sources, the observation setup was chosen so that a bright star was always in the slit, in order to allow a correct centering of the target.

The QSO SDSS J2054-0005 has been discovered in the SDSS deep Stripe 82 by Jiang et al. (2008). Our new ISAAC spectrum confirms the weak-line nature of this source (see also the optical discovery spectrum by Jiang et al. 2008). Given the intrinsic weakness of the Mg II emission, we do not include this QSO in the following analysis.

2.2.1.1 DATA REDUCTION

The ESO ISAAC pipeline produces wavelength calibrated co-added 2-D spectra from the individual frames that were acquired during each OB. Subsequent reduction was carried out within IRAF. One-dimensional spectra were extracted using the `apall` task. The tracing of the 1-D spectra was performed first on the bright stars in the slit. The resulting tracing functions were then used for the extraction of the QSO spectra. Individual 1D spectra were corrected for the telluric absorptions using the `telluric` task. Usually the telluric correction is performed by dividing the observed spectrum by the one of a telluric standard star observed shortly after the science target. This ratio is subsequently multiplied by the model atmosphere corresponding to the spectral type of the telluric standard and scaled to its observed K magnitude, in order to re-

cover the correct slope of the QSO spectrum. This operation allows us also to flux-calibrate the QSO spectrum. Instead of using the observed telluric standard stars, we employed the ESO sky absorption spectrum measured on the Paranal site at a nominal airmass of 1. This choice was driven by two reasons: i) the spectral regions in the QSO spectrum where telluric absorptions are most severe are characterized by a very low signal-to-noise, i.e. insufficient to properly correct for the actual detailed shape of the night sky features; ii) the observed telluric standard stars were often characterized by a spectral and luminosity class for which accurate model atmosphere could not be computed; the derived QSO continuum slope would have hence been distorted by the stellar spectral shape. We thus decided to focus on the spectral regions with higher signal, where an accurate correction was possible: we assumed the template sky absorption spectrum, scaled to the airmass of our QSO spectra and ignored the stability of the sky transparency. This way, we could preserve the intrinsic shape of the QSO continua. The `telluric` task corrects for the difference in airmass between the science and calibration spectra via the Beer-Lambert-Bouguer law. The telluric absorptions are in any case well removed from the spectra of bright QSOs, while significant residuals are left in the spectra of faint QSOs at lower S/N. The relative flux calibration was obtained with the `sensfunction`, `standard` and `calib` tasks. The instrument sensitivity function was obtained from the observed telluric stars of luminosity class V (giants) and spectral class B. For these stars it is possible to compute reliable model atmospheres since effective temperature and surface gravity are well estimated. The model atmospheres were computed by interpolating the NIR spectral energy distributions available from Pickles (1998) in temperature and surface-gravity in order to match the observed spectral type and magnitudes. After the relative flux calibration, the individual 1D QSO spectra were averaged to form a single spectrum. The absolute flux calibration was performed scaling the observed spectra to match the QSOs K-band magnitudes. Since the K-band magnitude was not measured for the QSOs in tab. 2.1, we derived it from the SDSS QSO template scaled to the observed J and H-band magnitude. The whole reduction procedure was extensively tested on some of the observed telluric standard stars used as reference targets. The recovered spectra match the theoretical model atmospheres typically within 10%, even in the spectral regions more affected by telluric absorptions. The reduced spectra are shown in fig. 2.1.

2.2.2 LITERATURE DATA

The QSOs taken from the literature are summarized in tab. 2.2. We have collected a total of 22 spectra (19 different sources) which cover all the features of interest at sufficiently high S/N to perform the spectral decomposition (see sec. 2.4.2 and sec. 2.4.6). The literature sample is composed of 10 sources with $z < 5.70$ and 9 sources with $5.70 < z < 6.43$. The 10 QSOs with $z < 5.70$ are selected amongst the 13 sources published by Iwamuro et al. (2002). These QSOs have redshifts $4.4 < z < 5.3$ and were observed in the J and H bands with an OH-airglow

suppressor spectrograph (OH-S), mounted on the Subaru Telescope. The spectral resolution ($\lambda/\Delta\lambda$) is equal to 210 (J band) and 420 (H band). We also selected 2 of the 4 SDSS QSOs at $z \sim 6$ presented by Iwamuro et al. (2004): H and K observations were carried out with the Cooled Infrared Spectrograph and Camera (CISCO) mounted on the Subaru telescope, with a spectral resolution of 210 and 330 respectively. We used the K-band spectrum of the $z = 6.43$ QSO J1148+5251 published by Barth et al. (2003). Data were obtained with the NIRSPEC spectrograph on KeckII (spectral resolution of 1500). We selected 4 of the 5 sources presented by Kurk et al. (2007) at redshifts $z > 5.8$. The K-band observations were carried out with VLT-ISAAC with a spectral resolution of 450. We also included the K-band spectrum of J0303-0019 presented by Kurk et al. (2009). This faint $z \sim 6$ QSO was selected in SDSS Stripe 82 and its spectrum was taken with VLT-ISAAC. Finally, we added 4 SDSS QSOs at $z > 5.8$ published by Jiang et al. (2007), who observed them with the Gemini Near-Infrared Spectrograph (GNIRS) on Gemini South with a spectral resolution of 800.

2.3 DATA ANALYSIS

We focus on the spectral region with restframe wavelengths $2000 \text{ \AA} < \lambda_{rest} < 3500 \text{ \AA}$. This region is characterized by the presence of the Mg II emission line, the underlying non-stellar continuum, the Balmer pseudo continuum and the Fe II emission line forest. The last three emission features, that are overlapped in the spectral range of interest, are described in the following sections.

2.3.1 POWER-LAW

The dominant component of a QSO spectrum is the non-stellar continuum, modeled as a power-law:

$$F_\lambda = F_0 \left(\frac{\lambda}{2500 \text{ \AA}} \right)^\alpha \quad (2.1)$$

Typically, in our data, the determination of the slope coefficient α depends on the adopted fitting procedure and on the observed spectral range. In case of a wide wavelength coverage it is possible to choose fitting windows free of contributions by other emission components. For spectra with restricted wavelength coverage, the power-law continuum has to be fitted simultaneously with the other components, resulting in a local estimate of the slope that might not be fully representative of the overall continuum shape of the QSO. Decarli et al. (2010) analyzed a sample of 96 QSOs at $z < 3$ by fitting the power-law continuum in 8 different windows free of strong features, and obtained a mean value for the slope of -1.3 with a $1-\sigma$ dispersion of 1.6. Shen et

Table 2.2: Literature sample sorted by redshift.

QSO name ¹	RA ²	DEC ³	z ⁴	Ref. ⁵
BR 1033 – 0327	10 33 51.47	–03 27 45.5	4.51	A
BR 0019 – 1522	00 19 35.90	–15 22 16.0	4.53	A
BR 2237 – 0607	22 37 17.44	–06 07 59.7	4.56	A
SDSS J0310 – 0014	03 10 36.97	–00 14 57.0	4.63	A
SDSS J1021 – 0309	10 21 19.16	–03 09 37.2	4.70	A
SDSS J0210 – 0018	02 10 43.17	–00 18 18.4	4.77	A
SDSS J0211 – 0009	02 11 02.72	–00 09 10.3	4.90	A
PC 1247 + 3406	12 47 17.79	+34 06 12.7	4.90	A
SDSS J0338 + 0021	03 38 29.31	+00 21 56.3	5.00	A
SDSS J1204 – 0021	12 04 41.73	–00 21 49.6	5.03	A
SDSS J0005 – 0006	00 05 52.30	–00 06 56.0	5.85	D
SDSS J1411 + 1217	14 11 11.30	+12 17 37.0	5.93	D, F
SDSS J1306 + 0356	13 06 08.20	+03 56 26.0	5.99	D, F
SDSS J1630 + 4012	16 30 33.90	+40 12 10.0	6.05	B
SDSS J0303 – 0019	03 03 31.40	–00 19 12.9	6.07	E
SDSS J1623 + 3112	16 23 31.80	+31 12 01.0	6.22	F
SDSS J1048 + 4637	10 48 45.05	+46 37 18.3	6.23	B
SDSS J1030 + 0524	10 30 27.10	+05 24 55.0	6.28	D, F
SDSS J1148 + 5251	11 48 16.64	+52 51 50.3	6.43	C

¹ QSO name;

^{2,3} QSO right ascension and declination in J2000.0;

⁴ optical redshift –from discovery paper;

⁵ reference paper for the NIR spectra: A, Iwamuro et al. (2002); B, Iwamuro et al. (2004); C, Barth et al. (2003); D, Kurk et al. (2007); E, Kurk et al. (2009); F, Jiang et al. (2007).

al. (2011) estimated the local slope for a sample of $\sim 100,000$ SDSS QSOs at $z < 4.95$, fitting the the power-law plus an iron template to the wavelength range around four broad emission lines ($H\alpha$, $H\beta$, $Mg\ II$ and $C\ IV$). In particular, from the analysis of the wavelength range adjacent the $Mg\ II$ emission line, they obtained a mean value for the local slope of -1.3 with a $1-\sigma$ dispersion of 0.8 . From our spectral decomposition we find consistent values (see sec. 2.3.4).

2.3.2 BALMER PSEUDO-CONTINUUM

We model the Balmer pseudo-continuum following Dietrich et al. (2003). We assume partially optically thick gas clouds with uniform temperature $T_e = 15000$ K. For wavelengths below the Balmer edge ($\lambda_{BE} = 3646$ Å), the Balmer spectrum can be parametrized as:

$$F_\lambda = F_{norm} B_\lambda(T_e) (1 - e^{-\tau_{BE} (\frac{\lambda}{\lambda_{BE}})^3}), \quad \lambda < \lambda_{BE} \quad (2.2)$$

where $B_\lambda(T_e)$ is the Planck function at the electron temperature T_e , τ_{BE} is the optical depth at the Balmer edge (we assume $\tau_{BE} = 1$ following Kurk et al. 2007), and F_{norm} is the normalized flux density at the Balmer edge (Grandi 1982). The normalization should be determined at $\lambda_{rest} \simeq 3675$ Å, where no Fe II emission is present. Since this wavelength is either not covered or has a very low S/N in our sample, we fix the normalization to a fraction of the continuum strength extrapolated at 3675 Å: $F_{norm} = f_B \cdot F_{power-law}(3675 \text{ Å})$. To define the relative strength of the two components and monitor the effects on the Fe II estimate, we have run various tests with $f_B = 0.1, 0.3, 0.5, 0.8, 1$. Since differences in the Fe II estimates resulting from this test were less than measured errors (with f_B only partially affecting the power-law normalization), we have fixed $f_B = 0.3$ based on the results by Dietrich et al. (2003).

2.3.3 Fe II TEMPLATE

The Fe II ion emits a forest of lines, many of which are blended. We fit the Fe II forest using a modified version of the emission line template by Vestergaard & Wilkes (2001). This template is based on the high resolution spectrum of the narrow line Seyfert 1 galaxy PG0050+124 ($z=0.061$), observed with the Hubble Space Telescope. Vestergaard & Wilkes obtained the emission line template by first fitting and subtracting the power-law continuum and all the absorption/emission features from all the elements but Fe. Afterward an Fe III model was subtracted from the residual to obtain a pure Fe II template. As Kurk et al. (2007) pointed out, no Fe II emission is left in the 2770 - 2820 Å range, due to the $Mg\ II$ line subtraction. We modified the template following Kurk et al. (2007) by adding a constant flux density between 2770 and 2820 Å equal to the 20% of the mean flux density of the template between 2930 and 2970 Å. The justification for

this operation comes from the theoretical Fe II emission line strength by Sigut & Pradhan (2003). Since the Fe II and the Mg II are not distributed in the same way within the BLR, we decided not to fix the Doppler broadening of the Fe II template to the one measured for the Mg II emission line. We have instead run many tests broadening the Fe II template by convolving it with Gaussian profiles with constant FWHM = 7.5, 15, 22.5 Å, corresponding to FWHM \sim 930, 1860, 2800 km/s at 2400 Å respectively. A Gaussian broadening with σ constant in wavelength leads to slightly different velocities over the fitting range, but we have found no significant differences in the measured Fe II normalization in function of the three broadening. Given the typical S/N of our spectra the Fe II normalization is in fact mainly determined by the fit of the template broad bumps rather than individual Fe II features. We finally chose to fix FWHM= 15 Å for the Fe II template.

2.3.4 FITTING PROCEDURE

All the spectra have been shifted to the rest-frame system of reference using optical redshifts. Even if the optical redshifts are slightly different from the NIR ones, this will not affect our results since the Mg II peak wavelength is a free parameter in our fitting procedure. Moreover, the Fe II template is constituted by blended multiplets (broadened by the convolution with Gaussian profiles with FWHM = 15 Å), for which the definition of peak wavelength is not straightforward. We tested a posteriori the difference between the optical redshifts (see tab. 2.1 and tab. 2.2) and the ones measured from the Mg II peak wavelength (see tab. 2.4). The average difference is equal to $\Delta z=0.02$ with a maximum of $\Delta z=0.08$ for J1411 + 1217 (Jiang et al. 2007online Appendix fig. A.12) : in this case the Mg II line profile is severely affected by the atmospheric absorptions. Our spectra have different wavelength coverage because of the different redshifts of the sources and of the various instruments used to collect the data. Similarly the sky contamination varies across the sample with redshift. For these reasons it is impossible to define fixed fitting windows for the entire sample. We focus on the rest-frame region within 2000 and 3500 Å, choosing the fitting windows in a way as homogeneous as possible, as a function of the spectral coverage and of the sky contamination. The fit is performed in two steps. A first set of spectral components is given by the sum of the power-law continuum, the Balmer pseudo-continuum and the Fe II template. The free parameters are the power-law slope (α) and its normalization ($\beta = \log(F_0)$) and the Fe II template normalization (γ). Since the three components are overlapped in this wavelength range, the fit is performed with a χ^2 minimization on a suitable grid in the parameter space. In this way, we minimize the possibility that our solution represents only a local minimum of the χ^2 domain. The errors on the Fe II normalization are computed by marginalizing the probability distributions in 3-D parameter space, selecting all the cases for which $\chi^2 - \chi^2_{min} < 1$ (1σ confidence level). To ensure a reliable estimate of these errors we need enough triplets satisfying the χ^2 condition. After many tests we decided to sample the parameter space with 16

million points: $100 (\alpha) \times 100 (\beta) \times 1600 (\gamma)$. This way, for each fit, we sample the γ parameter space with an average bin of $0.001 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$ and we obtain ~ 100 triplets satisfying the χ^2 , with a minimum of 15 (for the QSO BR 1033 – 0327). This condition is not satisfied for J0303 – 0019 (Kurk et al. 2009, online Appendix fig. A.17): in this case the Fe II major features are severely affected by the telluric absorptions, resulting in a weaker constraint on the Fe II template normalization which is in any case consistent with 0 (see tab. 2.3). In fig. 2.2 we show two examples of the 3D χ^2 -cube projections and of the relative probability distribution for the Fe II template normalization (similar plots for all sources discussed here are shown in the online Appendix). The χ^2 -maps are overall regular (non-patchy), implying the absence of secondary local minima. The degeneracy between power-law slope and intercept is evident from the bottom-right plots in the two panels.

The fitted components are then subtracted and we proceed to fit the Mg II emission line. The Mg II line fit is performed with a least-squares procedure. Since the Mg II doublet is not resolved in the majority of our spectra, we model the emission line as a simple Gaussian (three free parameters: central wavelength, width and normalization). If there was a significant narrow second component that we do not resolve, this would lead to a slight underestimate of the black-hole mass.

Examples of the spectral decomposition are shown in fig. 2.1. The results of the fit and relative χ^2 maps for the literature sample are shown in the online Appendix. The fitted parameters are listed in tab. 2.3. Even if we do not overcome the degeneracy between the power-law slope and intercept, we can compare the distribution of our local slope estimates with those in the literature. We obtain a mean value $\alpha = -1.5 \pm 1.2$ which is in agreement within the uncertainties with both the local slope estimate by Shen et al. (2008) and with the global one by Decarli et al. (2010).

2.4 RESULTS

2.4.1 BLACK HOLE MASS ESTIMATE

We estimate the M_{BH} using scaling relations, calibrated on local AGNs, that are based on broad emission line widths and continuum luminosities. Under the assumption that the dynamics of the broad line region is dominated by the gravity of the central BH, the virial theorem states:

$$M_{\text{BH}} = G^{-1} R_{\text{BLR}} v_{\text{BLR}}^2 \quad (2.3)$$

where M_{BH} is the black hole mass, R_{BLR} is the characteristic radius of the BLR and v_{BLR} is the orbital velocity of the clouds emitting at R_{BLR} . The cloud velocity can be obtained from the width

Figure 2.1: Reduced spectra of the QSOs SDSS J0842+1218 (upper-left panel), SDSS J0353+0104 (upper-right panel) and SDSS J2054-0005 (lower panel). The observed spectra are shown in black continuous line. The modeled components are: power-law continuum (blue dotted line), Balmer pseudo continuum (purple dashed line), Fe II normalized template (light blue dotted line), Mg II emission line (red dotted line). The sum of the first set of components (power-law continuum + Balmer pseudo continuum + Fe II normalized template) is overplotted to the spectrum as a green solid line, while the sum of all the components is over plotted as a red solid line. The spectral decomposition was not performed for the weak line QSO SDSS J2054-0005. Telluric absorption bands are indicated over the spectra with the symbol \oplus : they are extracted from the ESO sky absorption spectrum measured on the Paranal site at a nominal airmass of 1.

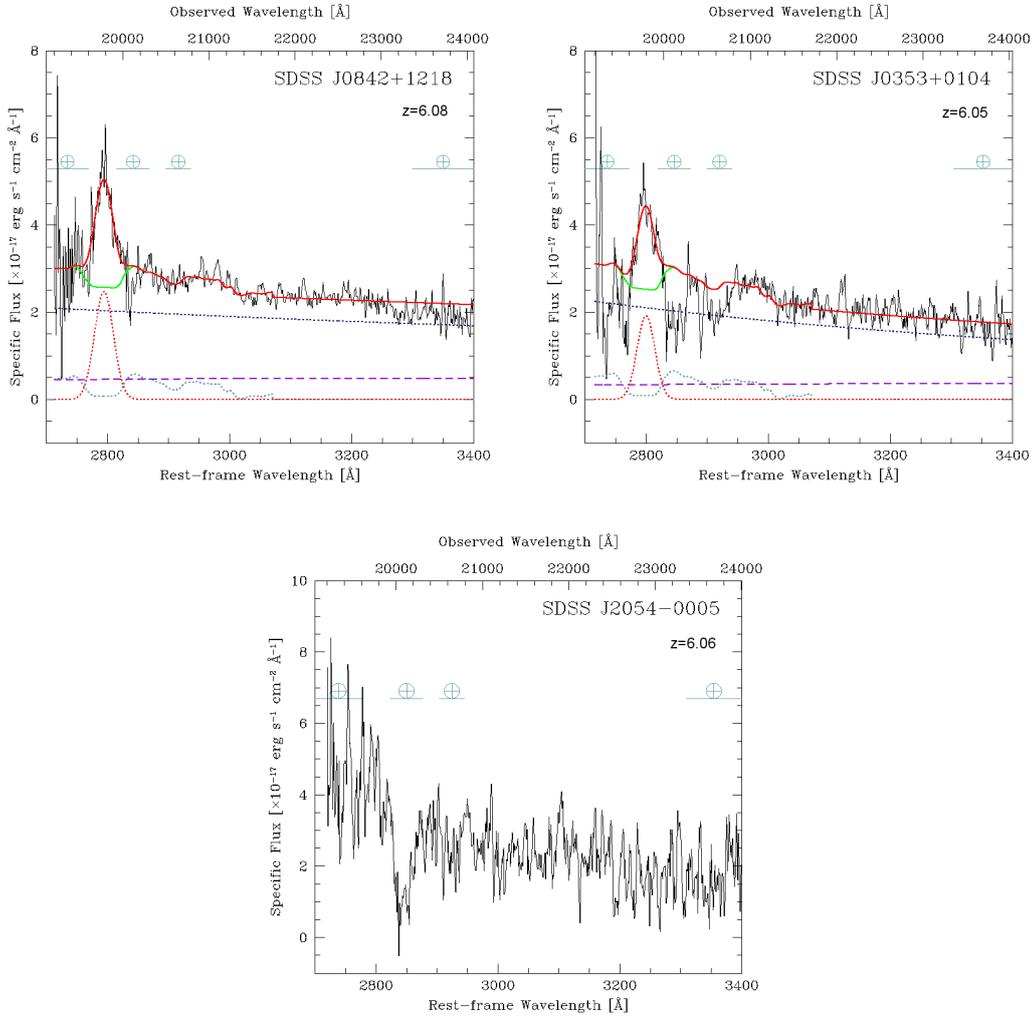
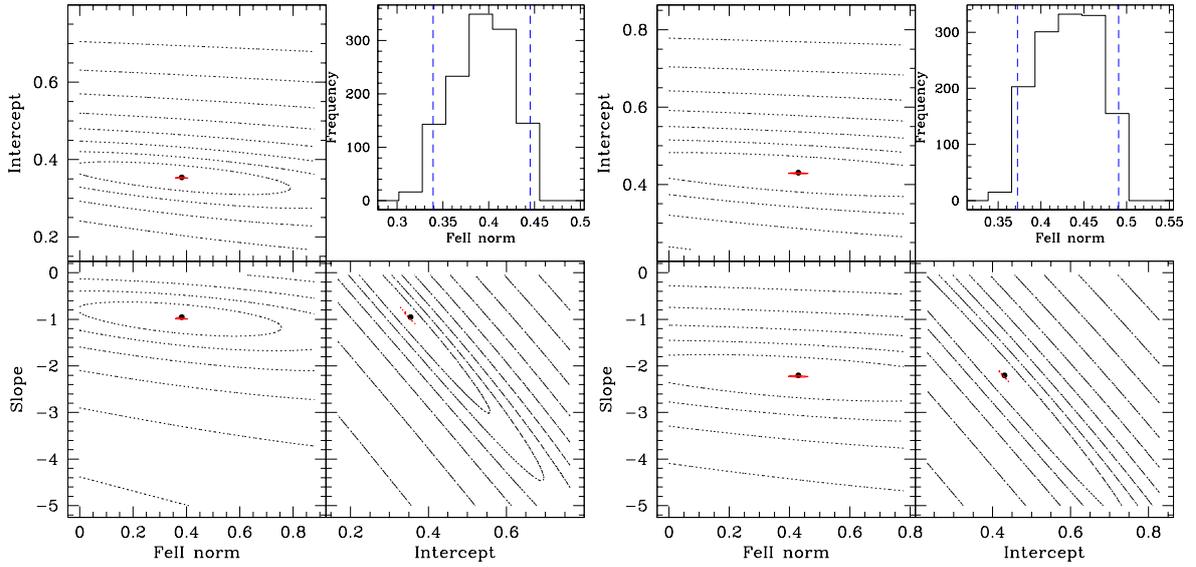


Figure 2.2: The Fe II normalization, obtained from the fit of the first set of components, depends on the power-law slope and its normalization (intercept). SDSS J0842+1218 (left panel) and SDSS J0353+0104 (right panel): χ^2 domain analysis for Fe II error computation. In each panel we show: a) two dimensional projections of the 3D χ^2 -surfaces (Fe II normalization vs Intercept, upper-left plot; Fe II normalization vs Slope, bottom-left plot; Intercept vs slope, bottom-right plot): contours represent iso- χ^2 levels spaced by a factor of 2 while the best fit case is marked with a dot; b) probability distribution for the Fe II template normalization (upper-right plot): the distribution has been obtained by marginalizing the 3-D probability distribution considering only the triplets for which $\chi^2 - \chi^2_{min} < 1$, the dashed vertical lines mark our estimate of the 1- σ confidence level.



of the broad emission lines: $v_{BLR} = f \cdot FWHM$, where f is a geometrical factor that accounts for the de-projection of v_{BLR} from the line of sight (see McGill et al. 2008; Decarli et al. 2008), and FWHM is the full width at half maximum of the line profile. Even if the BLR size cannot be directly measured by single epoch spectra, it can still be evaluated since it is strongly correlated with the continuum luminosity of the AGN (e.g. Kaspi et al. 2000). It is then possible to estimate M_{BH} for high-redshift QSOs using a single spectrum covering the blended Mg II line doublet ($\lambda = 2795, 2803 \text{ \AA}$) and the redward continuum ($\lambda = 3000 \text{ \AA}$). In particular for R_{BLR} we use the relation provided by McLure & Jarvis (2002):

$$\frac{R_{BLR}(\text{Mg II})}{10 \text{ light - days}} = (2.52 \pm 0.03) \left[\frac{\lambda L_{\lambda}(3000 \text{ \AA})}{10^{44} \text{ erg s}^{-1}} \right]^{0.47} \quad (2.4)$$

and for the geometrical factor the value provided by Decarli et al. (2008): $f(\text{Mg II}) = 1.6$, obtained assuming that the Mg II and H β -emitting regions have a similar geometry. All these relations are based on low redshift objects ($z < 0.3$), and the underlying assumption here is that they are still valid at high-redshift. Using relation (2.4) McLure & Jarvis (2002) reproduced M_{BH} obtained with the full reverberation mapping method with an accuracy of 0.4 dex. This intrinsic scatter of the estimator dominates the measurement uncertainties. To compare our results with the ones published in previous works on $z \sim 6$ sources (Barth et al. 2003; Jiang et al. 2007; Kurk et al. 2007, 2009), we also estimate M_{BH} using the relation obtained by McLure & Dunlop (2004):

$$\frac{M_{BH}}{M_{\odot}} = 3.2 \left[\frac{\lambda L_{\lambda}(3000 \text{ \AA})}{10^{37} \text{ W}} \right]^{0.62} \left[\frac{FWHM(\text{Mg II})}{\text{km s}^{-1}} \right]^2 \quad (2.5)$$

from a sample of 17 low redshift AGNs ($z < 0.7$), with luminosities comparable to those of high- z QSOs ($\lambda L_{\lambda} > 10^{44} \text{ erg s}^{-1}$).

From the estimated M_{BH} we compute the QSO Eddington ratios defined as the ratio between the measured bolometric luminosity L_{bol} and the theoretical Eddington luminosity L_{Edd} , as computed from the M_{BH} . We obtain the observed monochromatic luminosity $\lambda L_{\lambda}(3000 \text{ \AA})$ from the continuum component of the fit $f_{\lambda}(3000 \text{ \AA})$ and from the luminosity distance D_L as:

$$\lambda L_{\lambda}(3000 \text{ \AA}) = \lambda F_{\lambda}(3000 \text{ \AA}) 4\pi D_L^2 \quad (2.6)$$

Since $L(3000 \text{ \AA})$ is only a fraction of the total electromagnetic luminosity coming from the QSO, we apply the bolometric correction by Shen et al. (2008) to obtain L_{bol} :

$$L_{bol} = 5.15 \lambda L_{\lambda}(3000 \text{ \AA}) \quad (2.7)$$

The Eddington luminosity is defined as the maximum luminosity attainable, at which the radiation pressure acting on the gas counterbalances the gravitational attraction of the BH:

$$L_{Edd} = 1.3 \cdot 10^{38} \left(\frac{M_{BH}}{M_{\odot}} \right) \text{ erg s}^{-1} \quad (2.8)$$

The obtained M_{BH} and the relative QSO Eddington ratios are listed in tab. 2.4. The two relations, eq. 2.4 and eq. 2.5, lead to a difference in the mass estimates of a factor ~ 1.7 . The results obtained for the sample of published sources via eq. 2.5 are in good agreement with previous estimates, indicating that different fitting procedures imply variations smaller than the errors due to the intrinsic scatter of the estimator. The average difference between the BH masses estimated by us and the ones previously published is equal to 0.1 dex. Only in the case of the QSO SDSS J0005-0006 (Kurk et al. 2007, online Appendix fig. A.11) this difference is higher (0.7 dex): in this case the Mg II line is severely affected by the atmospheric absorption and the line fit becomes significantly procedure-dependent.

2.4.2 DEPENDENCE ON THE S/N AND ON THE SYSTEMATICS IN THE OBSERVATIONS

In fig. 2.3 we show the M_{BH} obtained via eq. 2.4 as a function of the S/N per \AA of the Mg II peak. The noise was derived from regions of the spectra free of line emission. There is no evidence of any dependence of the M_{BH} estimates on the S/N of the Mg II line peak. For SDSS J1411 + 1217, SDSS J1306 + 0356 and SDSS J1030 + 0524 we are in the fortunate situation of having two independent observations by Jiang et al. (2007), with Gemini-GNIRS, and by Kurk et al. (2007), with VLT-ISAAC. We can thus use these to analyze the dependence of the M_{BH} estimates on the systematics in the observations. For these 3 sources Jiang et al. (2007) and Kurk et al. (2007) derived (using eq. 2.5) $0.6, 1.1, 1.0 \times 10^9 M_{\odot}$ and $1.1, 2.4, 1.4 \times 10^9 M_{\odot}$, respectively. Using our fitting results and eq. 2.5, we obtain $0.5, 0.9, 1.1 \times 10^9 M_{\odot}$ and $1.0, 1.7, 1.4 \times 10^9 M_{\odot}$ respectively, in good agreement with the published values. For those objects whose redshifts make their Mg II line fall on top of atmospheric absorption bands, the M_{BH} estimated from the two independent spectra can be significantly different (up to 0.3 dex in the case of SDSS J1306 + 0356): therefore we argue that the atmospheric contamination is a major contributor to systematics.

2.4.3 COMPARISON WITH SDSS QSO SAMPLE AT $0.35 < z < 2.25$

To study the evolution of the BH population with cosmic time, we compare our sources with a sample of SDSS QSOs at lower redshift (SDSS Data Release 7, Shen et al. 2011). For objects with multiple observations we consider the weighted mean of the individual measurements. Since the M_{BH} estimates significantly depend on the adopted estimator (see Peterson 2010; Shen et al. 2008), we select among the $\sim 100,000$ QSOs in the SDSS DR7 sample those with measurements of the Mg II FWHM and of $\lambda L_{\lambda}(3000 \text{\AA})$: i.e., a subset of $\sim 47,000$ sources with $0.35 < z < 2.25$. We also consider a sample of nine $z \sim 6$ QSOs at lower luminosities ($M_{1450} < -24$) from Willott et al. (2010a), with published measurements of the Mg II FWHM and of $\lambda L_{\lambda}(3000 \text{\AA})$. We obtain L_{bol} using the bolometric correction by Shen et al. (2008) (see

Table 2.3: Fitted spectral properties. The spectral decomposition has been performed on the specific fluxes F_λ in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$.

QSO name ¹	slope ²	Fe II norm ³	Mg II λ_{peak} ⁴	Mg II FWHM ⁵	$F_\lambda(3000\text{\AA})$ ⁶
BR 1033 – 0327	-2.4	0.84±0.02	2804.6±2.3	4248±120	8.20±1.59
BR 0019 – 1522	-2.3	1.01±0.03	2802.6±0.6	4317±120	7.78±1.50
BR 2237 – 0607	-1.3	1.57±0.36	2796.3±0.6	3811±163	19.98±0.12
SDSS J0310 – 0014	-1.5	0.24±0.05	2797.3±1.1	4087±260	2.83±0.05
SDSS J1021 – 0309	-1.7	0.26±0.09	2797.4±0.6	3100±162	2.76±0.09
SDSS J0210 – 0018	-2.7	0.11±0.06	2798.4±2.1	6543±803	3.42±0.05
SDSS J0211 – 0009	-2.4	0.38±0.05	2795.7±1.5	5975±468	1.44±0.28
PC 1247 + 3406	-2.2	0.30±0.10	2797.7±0.7	4094±197	3.86±0.08
SDSS J0338 + 0021	-1.8	0.51±0.05	2798.4±0.4	2969±128	3.20±0.05
SDSS J1204 – 0021	-2.3	0.83±0.03	2796.2±3.0	5753±208	4.36±0.04
SDSS J0005 – 0006	0.0	0.18±0.04	2796.2±0.1	1036±65	0.94±0.02
SDSS J1411 + 1217*	-1.7	0.20±0.07	2800.4±1.2	2208±317	2.39±0.06
SDSS J1411 + 1217 ⁺	1.2	0.36±0.04	2787.8±0.4	2824±168	3.03±0.05
SDSS J1306 + 0356*	-2.3	0.21±0.12	2809.5±0.5	3158±145	1.92±0.05
SDSS J1306 + 0356 ⁺	-2.5	0.13±0.13	2809.8±1.1	4134±351	1.97±0.04
SDSS J1630 + 4012	-0.7	0.03±0.03	2798.0±2.0	3366±533	1.46±0.30
SDSS J0303 – 0019	2.0	0.00±0.15	2802.28±0.4	2307±63	0.60±0.01
SDSS J1623 + 3112	0.0	0.31±0.02	2795.0±0.3	3587±118	1.81±0.05
SDSS J1048 + 4637	-1.7	0.56±0.03	2797.9±1.4	3366±532	4.22±0.04
SDSS J1030 + 0524*	-1.9	0.16±0.07	2807.2±0.6	3449±151	1.76±0.05
SDSS J1030 + 0524 ⁺	-1.3	0.17±0.08	2805.9±0.9	3704±231	1.88±0.03
SDSS J1148 + 5251	-2.1	0.91±0.01	2793.7±0.7	5352±289	4.22±0.08
SDSS J0353 + 0104	-2.2	0.43±0.06	2799.6±0.7	3682±281	2.16±0.10
SDSS J0842 + 1218	-0.9	0.38±0.06	2794.2±0.7	3931±257	2.38±0.12

¹ QSO name. For sources with multiple observations: *, Jiang et al. (2007); +, Kurk et al. (2007);

² power-law slope;

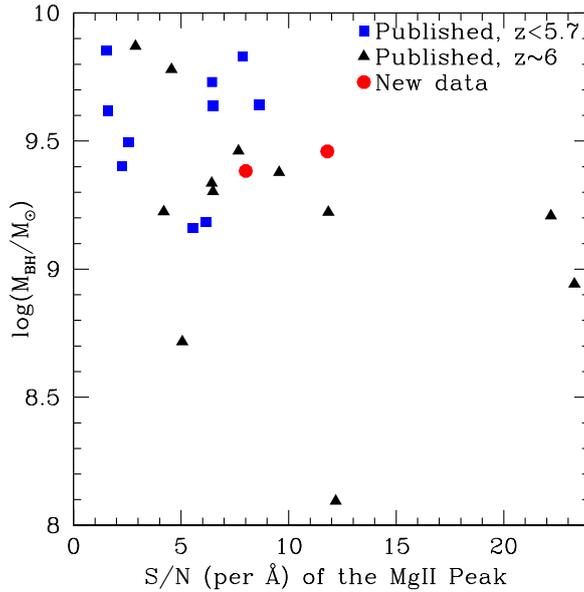
³ Fe II normalization;

⁴ Mg II central wavelength [\AA];

⁵ FWHM [km s^{-1}];

⁶ measured specific flux of the continuum at 3000 \AA , in units of $10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$.

Figure 2.3: M_{BH} estimates as a function of the S/N per \AA at the Mg II peak. Blue squares: literature sample, at $z < 5.7$ (Iwamuro et al. 2002); black triangles: literature sample, $z \sim 6$ (Barth et al. 2003; Iwamuro et al. 2004; Jiang et al. 2007; Kurk et al. 2007, 2009); red circles: this study. Our M_{BH} estimates do not depend on the S/N of the QSO spectra.



eq. 2.7) and we compute M_{BH} using eq. 2.4, in analogy with the analysis performed in our sample. In fig. 2.4 we plot the estimated M_{BH} as a function of the bolometric luminosity for the three samples. The black solid lines indicate regions of the parameter space with constant Eddington ratios: $L_{\text{bol}}/L_{\text{Edd}} = 0.01, 0.1, 1.0$. The $0.35 < z < 2.25$ and our high- z QSO populations occupy different regions of the parameter space: the $0.35 < z < 2.25$ redshift population shows an average $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -1.2$ dex ($L_{\text{bol}}/L_{\text{Edd}} \sim 0.06$) with a dispersion of 0.3 dex, while the typical ratio for the high- z QSOs is $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -0.3$ dex ($L_{\text{bol}}/L_{\text{Edd}} \sim 0.5$) with a dispersion of 0.3 dex. The lower luminosity $z \sim 6$ sample (Willott et al. 2010a) is characterized by lower M_{BH} but comparable Eddington ratios than the most luminous SDSS QSOs at the same redshift. We therefore conclude that the average Eddington ratio for the $z \sim 6$ QSO sample studied to date is significantly higher than the typical SDSS QSOs at $0.35 < z < 2.25$ (in agreement with previous studies, e.g. Trakhtenbrot et al. 2011).

We have to specify that Shen et al. (2011) fit the spectra considering the power-law non-stellar continuum, the Fe II line forest (modeled using the template by Vestergaard & Wilkes 2001), and both the narrow (simple Gaussian) and broad (simple/double Gaussian) components of the Mg II doublet. Since in our sample we are not subtracting the narrow component of the Mg II doublet (not resolved), the mass estimates of the high-redshift QSOs may be systematically biased towards slightly lower values with respect to the ones in the $0.35 < z < 2.25$ redshift sample.

Nevertheless, we note that our fitting procedure yields results consistent with the ones published by Jiang et al. (2007) who, except for the cases in which the line profile is severely affected by the atmospheric absorption, fits the Mg II emission line with a double Gaussian instead of a single one. Therefore, we conclude that the M_{BH} estimates do not strongly depend on the adopted fitting procedure.

2.4.4 COMPARISON WITH LUMINOSITY MATCHED SDSS QSO SAMPLE

In order to perform a statistically robust comparison we now extract only objects with the same (high) bolometric luminosity: $10^{47} < L_{\text{bol}} [\text{erg s}^{-1}] < 3 \times 10^{47}$ from our sample and from the SDSS comparison sample at $0.35 < z < 2.25$. In fig. 2.5 we plot the M_{BH} and the Eddington ratio as a function of redshift. The average M_{BH} at high- z is systematically lower than the typical M_{BH} of the $0.35 < z < 2.25$ population at similar luminosities, while the Eddington ratio is higher. Comparing the Eddington ratio distribution of the two luminosity matched samples (fig. 2.6, $10^{47} < L_{\text{bol}} [\text{erg s}^{-1}] < 3 \times 10^{47}$) we obtain that for the $0.35 < z < 2.25$ population $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -0.80$ dex ($L_{\text{bol}}/L_{\text{Edd}} \sim 0.16$) with a dispersion of 0.24 dex, while for the high- z one $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -0.37$ dex ($L_{\text{bol}}/L_{\text{Edd}} \sim 0.43$) with a dispersion of 0.20 dex. If we perform a Kolmogorov-Smirnov test on the two distributions we obtain that the probability that they come from the same parent distribution is 7×10^{-9} , i.e. negligible. This implies that the high-redshift QSOs in this luminosity bin are building up their mass more rapidly than the ones at $z \sim 2$ (in agreement with what has been found in previous studies, e.g. Kurk et al. 2009; Willott et al. 2010; Trakhtenbrot et al. 2011).

Another bias that could possibly affect our comparison between the high- z Eddington ratio distribution and the lower- z one is the intrinsic luminosity of the QSO broad lines. For instance the equivalent width of high ionization lines in AGN is known to anti-correlate with the continuum luminosity (Baldwin 1977; Baskin & Laor 2004; Shang et al. 2003; Xu et al. 2008). This effect is a function of the line ionization potential and not observed in low ionization lines (e.g. Mg II line, Espey & Andreadis 1999). High- z QSOs are usually selected in function of their photometric colors as drop-out objects. This criterion is to some extent insensitive to the line brightness. For the SDSS QSOs sample with $0.35 < z < 2.25$ the existence of a broad line feature is instead required in addition to the color selection, resulting into a bias against weak line QSOs. Decarli et al. (2011) found that for the DR7 SDSS QSO sample (Shen et al. 2011) the Mg II line luminosity positively correlates with the continuum luminosity. Selecting objects within a given Mg II line luminosity bin would then result in a continuum luminosity cut, like the one previously taken into account.

Figure 2.4: M_{BH} as a function of the bolometric luminosity: comparison between the present dataset and the SDSS QSOs sample at $0.35 < z < 2.25$. Blue squares: literature sample, $z < 5.7$ (Iwamuro et al. 2002); black triangles: literature sample, $z \sim 6$ (Barth et al. 2003; Iwamuro et al. 2004; Jiang et al. 2007; Kurk et al. 2007, 2009); red diamonds: new observations; green pentagons: CFHQS QSO sample (Willott et al. 2010); grey dots: SDSS QSOs sample $0.35 < z < 2.25$ (DR7, Shen et al. 2011); purple lines: isodensity contours. The black solid lines indicate regions of the parameter space with constant Eddington ratios: $L_{\text{bol}}/L_{\text{Edd}} = 0.01, 0.1, 1.0$. The $0.35 < z < 2.25$ and our high- z QSO populations occupy different regions of the parameter space: the $0.35 < z < 2.25$ redshift population shows $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -1.2$ dex with a dispersion of 0.3 dex, while the typical ratio for the high- z QSOs is $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -0.3$ dex with a dispersion of 0.3 dex.

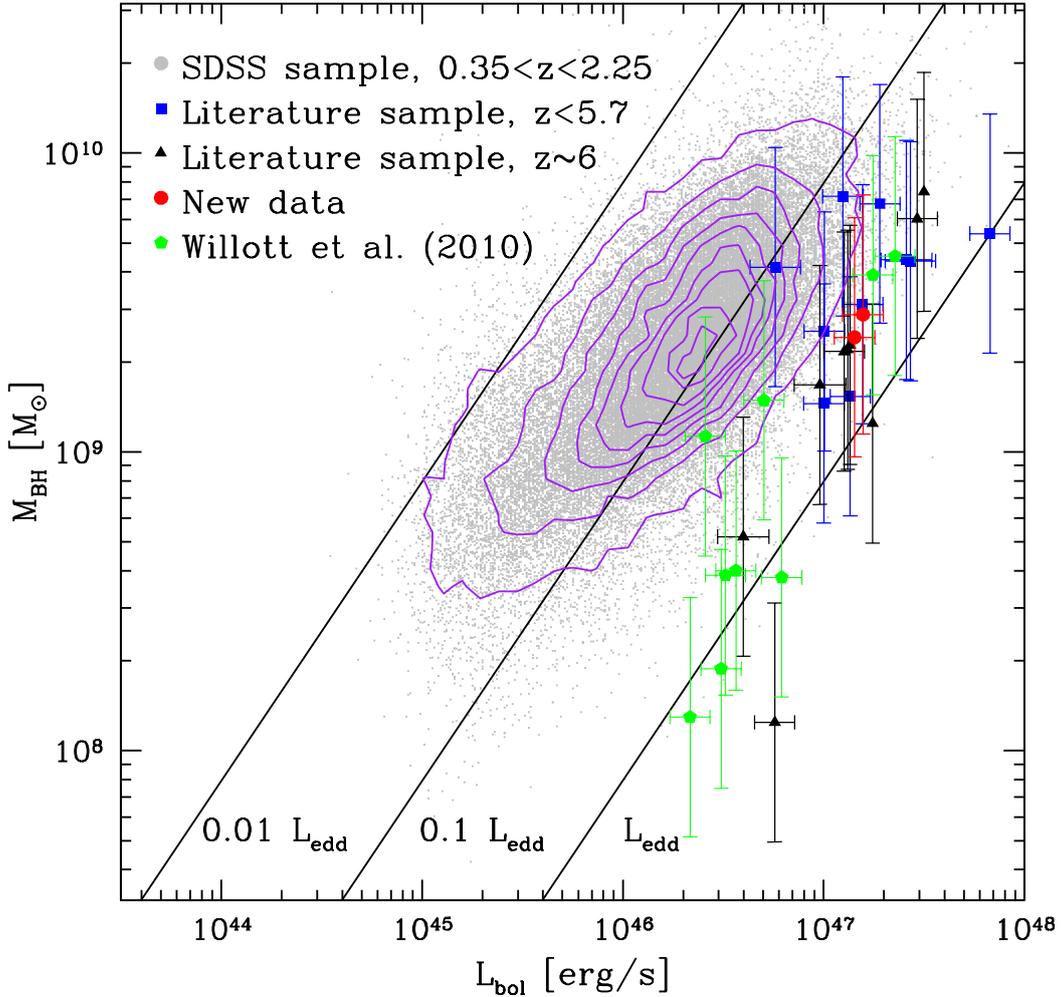


Figure 2.5: Eddington ratio (upper panel) and M_{BH} (lower panel) as a function of redshift for subsamples of QSOs with bolometric luminosities $10^{47} < L_{\text{bol}} [\text{erg s}^{-1}] < 3 \times 10^{47}$ extracted from the $0.35 < z < 2.25$ sample and from our high- z sample. Blue circles: $z < 5.7$ (Iwamuro et al. 2002); black circles: $z \sim 6$ (Barth et al. 2003; Iwamuro et al. 2004; Jiang et al. 2007; Kurk et al. 2007, 2009; this study); grey circles: SDSS QSOs sample $0.35 < z < 2.25$ (DR7, Shen et al. 2011); purple lines: isodensity contours. The average M_{BH} at high- z is lower than the typical M_{BH} of the $0.35 < z < 2.25$ population at similar luminosities, while the Eddington ratio is higher.

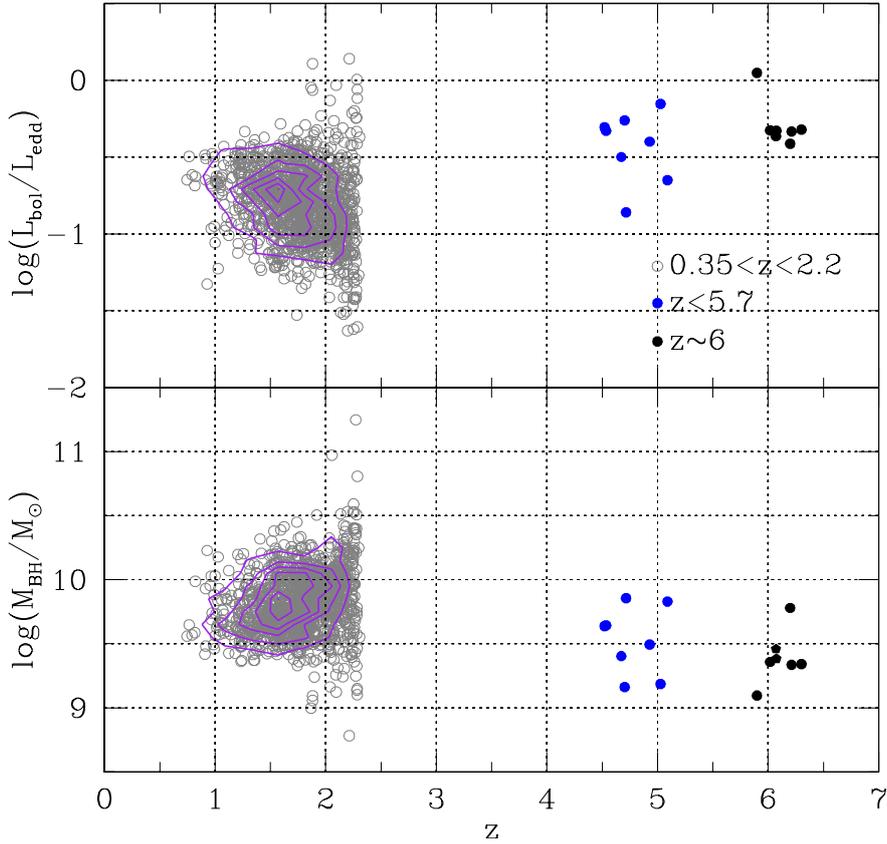
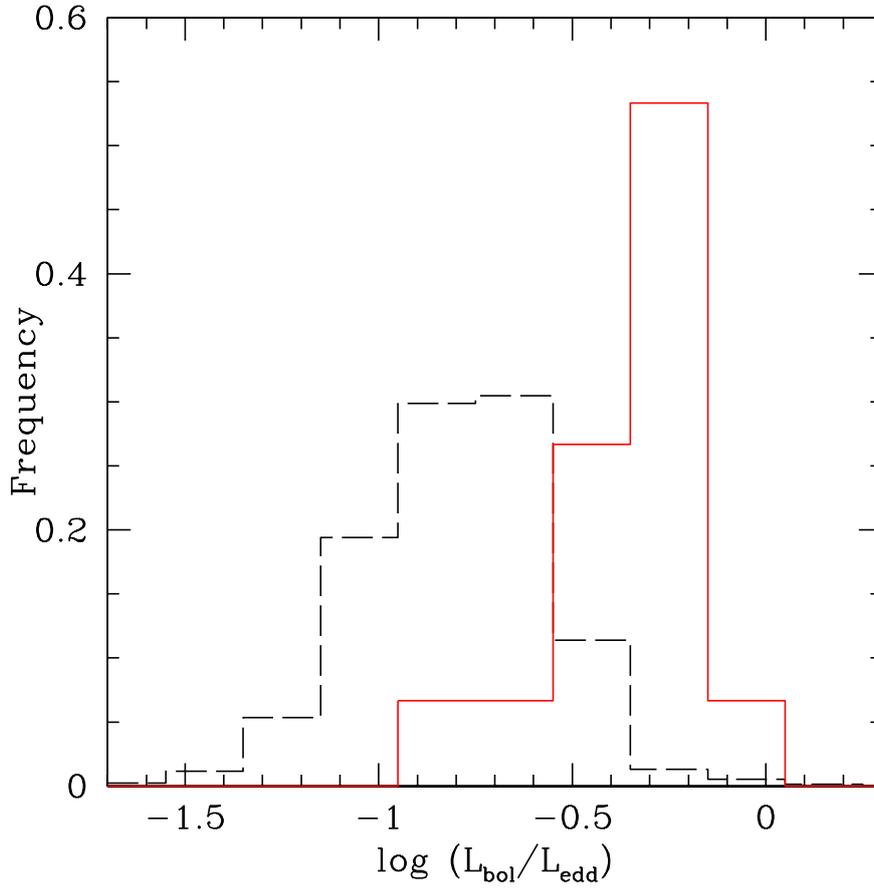


Figure 2.6: Eddington ratio distributions for sub samples of QSOs with bolometric luminosities $10^{47} < L_{bol} [\text{erg s}^{-1}] < 3 \times 10^{47}$ extracted from the $0.35 < z < 2.25$ sample and from our high- z sample: a) black dashed line: SDSS QSOs at $0.3 < z < 2.2$ (SDSS DR7, Shen et al. 2011); b) red solid line: our sample at $4.7 < z < 6.5$. The two distributions are intrinsically different: for the $0.35 < z < 2.25$ population $\langle \log(L_{bol}/L_{Edd}) \rangle \sim -0.80$ dex with a dispersion of 0.24 dex, while for the high- z one $\langle \log(L_{bol}/L_{Edd}) \rangle \sim -0.37$ dex with a dispersion of 0.20 dex. The Kolmogorov-Smirnov probability that the two distributions come from the same parent distribution is negligible.



2.4.5 IMPACT ON BH FORMATION HISTORY

Accurate measurements of BH masses and relative Eddington ratios of $z \sim 6$ QSOs can be used to give constraints on the formation processes of super-massive BHs (SMBHs) in the early universe. The three outstanding theories for SMBH seed formation differ both in the astrophysical processes considered and in the resulting masses of the seeds: (i) BHs seeds of few hundreds of M_{\odot} can be produced by the first generation of stars (Pop III stars), formed out of zero metallicity gas; (ii) BH seeds of $10^2 - 10^4 M_{\odot}$ can be obtained via stellar-dynamical processes amongst Pop II stars; (iii) BH seeds of $10^4 - 10^5 M_{\odot}$ can be produced via gas-dynamical processes (direct collapse of dense gas clouds) (see [Volonteri 2010](#) for a review on the SMBH formation).

We here consider only the 20 QSOs at $z \sim 6$, since they provide tighter constraints on the SMBH formation: 11 sources from our sample ($z > 5.7$) and 9 QSOs by [Willott et al. \(2010a\)](#). The resulting sample includes all the faint $z \sim 6$ QSOs known to date with estimated M_{BH} from the Mg II line. For this $z \sim 6$ population we obtain $\langle \log(M_{\text{BH}}/M_{\odot}) \rangle \sim 9.06$ dex ($M_{\text{BH}} \sim 1.1 \times 10^9 M_{\odot}$) with a dispersion of 0.5 dex and $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -0.23$ dex ($L_{\text{bol}}/L_{\text{Edd}} \sim 0.6$) with a dispersion of 0.29 dex. The time needed by a BH seed to grow at a constant rate from an initial mass M_0 to a final mass M_t is equal to ([Volonteri & Rees 2005](#)):

$$t = 0.45 \text{ Gyr} \times \left[\frac{\epsilon}{1 - \epsilon} \right] \times \frac{L_{\text{edd}}}{L_{\text{bol}}} \times \ln \left(\frac{M_t}{M_0} \right) \quad (2.9)$$

where 0.45 Gyr is the Salpeter time, ϵ is the radiative efficiency ($\epsilon \sim 0.07$, [Volonteri & Rees 2005](#)), and $L_{\text{bol}}/L_{\text{edd}}$ is the Eddington ratio. A BH seed with mass $M_0 = 10^2, 10^5 M_{\odot}$ accreting constantly at the Eddington ratio characteristic of our population ($L_{\text{bol}}/L_{\text{edd}} \sim 0.6$), would then need a time $t \sim 0.91, 0.53$ Gyr, respectively, to obtain a final mass M_t equal to the mean M_{BH} of our sample.

If the characteristic Eddington ratio of the ~ 6 QSO population is well represented by our mean estimate, models with BH seed masses $\sim 10^2 M_{\odot}$ become problematic since the time needed for the seed to grow is comparable to the age of the universe at $z \sim 6$ (~ 0.9 Gyr). Moreover, since our QSOs are selected from flux-limited surveys the resulting distribution of the Eddington ratios should be biased towards higher values: BHs with low accretion rates may not pass the QSO selection magnitude limits. [Kelly et al. \(2010\)](#) and [Willott et al. \(2010a\)](#) have shown that the intrinsic Eddington ratio distribution for a volume limited QSO sample is indeed shifted towards lower values with respect to a luminosity selected population. Our estimate hence represents an upper limit of the characteristic Eddington ratio of the $z \sim 6$ QSO population: lower Eddington ratios, according to eq. 2.9, would yield even longer time for the BH seeds to grow. In the future, with complete samples of QSOs at $z \sim 6$, it will be possible to give better constraints on the super massive BHs formation models.

Table 2.4: Estimated M_{BH} , QSO Eddington ratios and Fe II/Mg II line ratios. The mass estimates have an accuracy of 0.4 dex that dominates the measurement uncertainties.

QSO name ¹	z ²	$M_{\text{BH}}(10^9 M_{\odot})$		$L_{\text{bol}}/L_{\text{edd}}$		Fe II/Mg II ⁷
		eq. (2.4) ³	eq. (2.5) ⁴	eq. (2.4) ⁵	eq. (2.5) ⁶	
BR 1033 – 0327	4.521±0.005	4.3	2.8	0.5	0.8	2.6±0.1
BR 0019 – 1522	4.534±0.001	4.4	2.8	0.5	0.7	2.7±0.1
BR 2237 – 0607	4.561±0.001	5.4	4.0	1.0	1.3	2.1±0.5
SDSS J0310 – 0014	4.672±0.002	2.5	1.4	0.3	0.6	2.7±0.6
SDSS J1021 – 0309	4.703±0.001	1.5	0.8	0.5	1.0	3.3±1.1
SDSS J0210 – 0018	4.715±0.004	7.2	4.1	0.1	0.2	1.5±0.9
SDSS J0211 – 0009	4.894±0.003	4.1	2.1	0.1	0.2	4.2±0.6
PC 1247 + 3406	4.929±0.001	3.1	1.9	0.4	0.7	2.1±0.7
SDSS J0338 + 0021	5.027±0.001	1.5	0.9	0.7	1.2	5.0±0.6
SDSS J1204 – 0021	5.090±0.007	6.8	4.1	0.2	0.4	3.3±0.1
SDSS J0005 – 0006	5.844±0.001	0.1	0.06	3.6	7.1	4.7±1.1
SDSS J1411 + 1217*	5.854±0.003	0.9	0.5	1.3	2.2	2.0±0.7
SDSS J1411 + 1217+	5.903±0.001	1.6	1.0	0.9	1.5	2.9±0.5
SDSS J1306 + 0356*	6.017±0.001	1.7	0.9	0.6	1.0	4.8±2.7
SDSS J1306 + 0356+	6.018±0.003	2.9	1.7	0.3	0.6	1.5 ^{+1.6} _{-1.5}
SDSS J1630 + 4012	6.058±0.005	1.7	0.9	0.5	0.8	0.5±0.5
SDSS J0303 – 0019	6.079±0.001	0.5	0.3	0.6	1.3	0.0±5.8
SDSS J1623 + 3112	6.211±0.001	2.2	1.2	0.5	0.8	4.3±3.3
SDSS J1048 + 4637	6.198±0.004	6.0	3.9	0.4	0.6	4.1±0.3
SDSS J1030 + 0524*	6.302±0.001	2.0	1.1	0.5	0.9	2.9±1.3
SDSS J1030 + 0524+	6.299±0.002	2.4	1.4	0.5	0.8	2.3±1.2
SDSS J1148 + 5251	6.407±0.002	7.4	4.9	0.3	0.5	4.6±0.2
SDSS J0353 + 0104	6.072±0.002	2.4	1.4	0.5	0.8	4.7±0.7
SDSS J0842 + 1218	6.069±0.002	2.9	1.7	0.4	0.7	3.1±0.5

¹ QSO name. For sources with multiple observations: *, Jiang et al. (2007); +, Kurk et al. (2007);

² redshift estimate from Mg II line

³ M_{BH} in units of $10^9 M_{\odot}$ estimated using eq. 2.4;

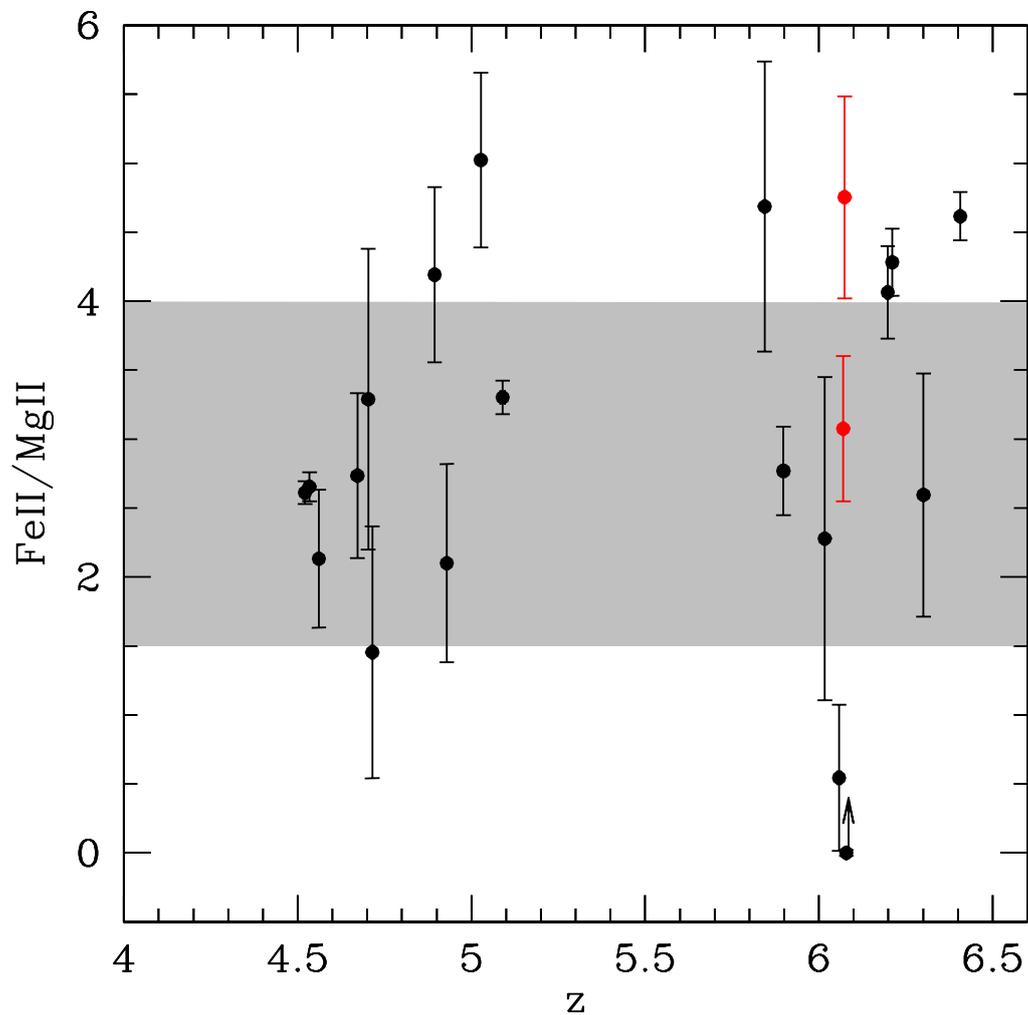
⁴ M_{BH} in units of $10^9 M_{\odot}$ estimated using eq. 2.5;

⁵ Eddington accretion ratio derived from M_{BH} estimated using eq. 2.4;

⁶ Eddington accretion ratio derived from M_{BH} estimated using eq. 2.5;

⁷ Fe II/Mg II ratio.

Figure 2.7: Evolution of the Fe II/Mg II line ratio for $z > 4$. Black points: literature sample, $z > 4.4$; red points: new targets. Grey-shaded area: expected Fe II/Mg II line ratio obtained by modeling the Fe II and Mg II line strength considering gas with cosmic abundance and a range of ionizing continua (Wills et al. 1985). The reported scatter at $z \sim 6$ (Kurk et al. 2007) is now significantly reduced: our measurements of the line ratios span a range $0 < \text{Fe II/Mg II} < 6$, while previous results were distributed up to $\text{Fe II/Mg II} \sim 15$. No redshift evolution of the Fe II/Mg II ratio is observed for $4 < z < 6.5$.



2.4.6 Fe II/Mg II

An estimate of the Fe/Mg abundance ratio at $z \sim 6$ may serve as an indication of the onset of star formation in the highest- z QSOs (see sec. 2.1). The proxy usually used to trace the Fe/Mg abundance ratio at high- z is the Fe II/Mg II line ratio. In the past, numerous NIR-spectroscopy studies have been carried out to analyze the Fe II/Mg II line ratio in high- z QSOs (e.g. Maiolino et al. 2001, 2003; Pentericci et al. 2002; Iwamuro et al. 2002, 2004; Barth et al. 2003; Dietrich et al. 2003; Freudling et al. 2003; Willott et al. 2003; Jiang et al. 2007; Kurk et al. 2007). The combined results revealed an increase in the scatter of the measured Fe II/Mg II line ratios as a function of redshift (see Kurk et al. 2007). A proposed explanation for the increased scatter is that some young objects have been observed such a short time after the initial starburst that the BLR has not been fully enriched with Fe yet.

We compute the Fe II flux by integrating the normalized Fe II template over the rest-frame wavelength range $2200 \text{ \AA} < \lambda_{rest} < 3090 \text{ \AA}$. The Mg II flux is computed by integrating the fitted single Gaussian over the range $\lambda_{peak} \pm 5\sigma$. The resulting Fe II/Mg II flux ratios and relative errors are listed in tab. 2.4. In fig. 2.7 we show the evolution of the Fe II/Mg II as a function of z (for objects with multiple observations we consider the weighted mean of the individual measurements). The reported scatter at $z \sim 6$ (Kurk et al. 2007) is now significantly reduced: our measurements of the line ratios span a range $0 < \text{Fe II/Mg II} < 6$, while previous results were distributed up to $\text{Fe II/Mg II} \sim 15$. The Fe II/Mg II line measurements do not show any clear dependence on the estimated S/N per \AA of the spectral continuum (see fig. 2.8). On the other hand, our flux ratio estimates are well in agreement with previous results by Kurk et al. (2007) who performed an analysis similar to ours. We can conclude that this measurement is significantly dependent on the performed analysis, rather than on the S/N of the spectra. This result immediately implies that we do not need to invoke different chemical evolutionary stages of the BLR gas. We do not see evidence for evolution of the estimated Fe II/Mg II line ratio as a function of cosmic age for $4 < z < 6.5$. Moreover, if we consider the SDSS QSO sample at $0.35 < z < 2.25$ (Shen et al. 2011 see sec. 2.4.3), and we compute the Fe II/Mg II line ratio starting from the published Fe II and Mg II flux measurements, we see no evolution of the line ratio also for $0.35 < z < 2.25$. However it is not possible to perform a direct comparison between the two samples, since these line ratio measurements are highly dependent on the adopted fitting procedure.

To constrain the onset of the star formation in these early QSOs, we would need to derive an estimate of the Fe/Mg abundance ratio from the measurements of the Fe II/Mg II line ratio. Unfortunately, an accurate conversion cannot be performed since the detailed formation of the Fe II bump in QSO spectra has not been fully understood yet. Wills et al. (1985) discussed the suitability of the Fe II/Mg II line ratio as a proxy of the Fe/Mg abundance ratio, as the regions in which the two ions are produced and the radiative transfer of the two lines are both very dif-

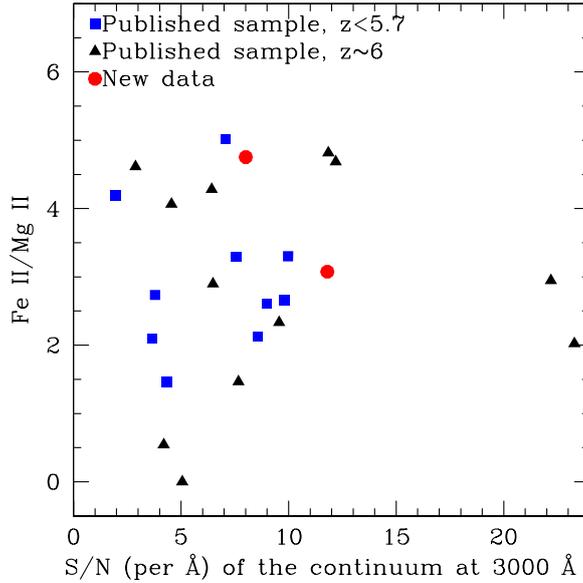
ferent. Computing the Fe II and Mg II line strength for a gas with cosmic abundances and using a wide range of hydrogen densities and ionization parameters, they predicted typical Fe II/Mg II line ratios between 1.5 and 4.0 (this is indicated in fig. 2.7 as a grey-shaded area), and attributed higher Fe II/Mg II values to an overabundance of Fe with respect to Mg. However these computations were based on a limited 70-level model of the Fe atom. Baldwin et al. (2004) used a 371-level Fe⁺ model to reproduce the observed Fe II emission properties in AGNs considering a large set of models of broad emission line region clouds. From their results, all the observed Fe II features can be reproduced only if: (i) the BLR is characterized by the presence of significant microturbulence; (ii) the Fe II emitting gas has different properties (density and/or temperature) with respect to the gas emitting other broad lines. In any case the strength of the Fe II emission relative to the emission line of other ions (e.g the Fe II/Mg II line ratio) depends as much on the Fe abundance as it does on other physical parameters of the BLR (e.g. turbulence velocity), making it difficult to convert the observed line ratios in abundance ratios. Nevertheless, the study of the Fe II/Mg II line ratio as a function of look-back time can be used to give constraints on the BLR chemical enrichment history, under the assumption that physical conditions of the BLR that determine the Fe II emission are not evolving.

The observed lack of evolution in the measured Fe II/Mg II line ratio can then be explained with an early chemical enrichment of the QSO host. I.e. the QSOs in our sample must have undergone a major episode of Fe enrichment in a few hundreds Myr before the cosmic age at which they have been observed (~ 0.9 Gyr). Matteucci & Recchi (2003) showed that for a massive elliptical galaxy characterized by a very intense but short star-formation history, the typical timescale for the maximum SN Ia rate can be as short as 0.3 Gyr. This implies that if the Fe in high- z QSO hosts is mainly produced via SN Ia explosions, it would be possible to observe fully enriched BLR at $z \sim 6$. On the other hand, Venkatesan et al. (2004) pointed out that SNe Ia are not necessarily the main contributors to the Fe enrichment, and that stars with a present-day initial mass function are sufficient to produce the observed Fe II/Mg II line ratios at $z \sim 6$. Fe could also be generated by Pop III stars: these very metal poor stars with typical masses $M \gtrsim 100 M_{\odot}$ might be able to produce large amounts of Fe within a few Myr (Heger & Woosley 2002).

2.5 SUMMARY AND CONCLUSIONS

We presented NIR-spectroscopic observations of three $z \sim 6$ SDSS QSOs. Our NIR spectra cover the Mg II and Fe II emission features, which are powerful probes of M_{BH} and of the chemical enrichment of the BLR. The new data extend the existing SDSS sample towards the faint end of the QSO luminosity function. We have collected 22 literature spectra (19 different sources) of high-redshift ($z > 4$) QSOs covering the rest-frame wavelength range $2700 \text{ \AA} < \lambda < 3200 \text{ \AA}$.

Figure 2.8: Fe II/Mg II line ratio estimates as a function of the S/N per Å of spectral continuum at 3000 Å. Blue squares: literature sample, at $z < 5.7$ (Iwamuro et al. 2002); black triangles: literature sample, $z \sim 6$ (Barth et al. 2003; Iwamuro et al. 2004; Jiang et al. 2007; Kurk et al. 2007, 2009); red circles: new observations. The Fe II/Mg II line measurements do not show a dependence on the S/N per Å of the spectral continuum.



The final sample is composed of 22 sources: our three new sources at $z \sim 6$, 10 spectra from the literature of QSOs with $4 < z < 5.7$ (Iwamuro et al. 2002) and 9 spectra of QSOs with $z \sim 6$ (Barth et al. 2003; Iwamuro et al. 2004; Jiang et al. 2007; Kurk et al. 2007, 2009). In order to see how the derivation of physical parameters depend on the fitting strategy employed, we performed a new consistent analysis of the sample and we gave an estimate of the M_{BH} , of the $L_{\text{bol}}/L_{\text{Edd}}$ and of the Fe II/Mg II line ratio.

Our results and conclusions can be summarized as follows:

1. We have estimated the M_{BH} from the Mg II emission line using empirical mass scaling relations. The QSOs in our sample host BHs with masses of $\sim 10^9 M_{\odot}$. Our results are in good agreement with previous estimates, indicating that different fitting procedures imply variations smaller than the errors due to the intrinsic scatter of the M_{BH} estimator.
2. High-redshift QSOs are accreting close to the Eddington luminosity: $\langle \log(L_{\text{bol}}/L_{\text{Edd}}) \rangle \sim -0.3$ dex ($L_{\text{bol}}/L_{\text{Edd}} \sim 0.5$) with a dispersion of 0.3 dex. The distribution of observed Eddington ratios is significantly different (average Eddington ratio ~ 8 times higher) from that of a comparison sample at $0.35 < z < 2.25$ (SDSS, DR7). This result is not biased

by the luminosity selection of the high- z sample: the difference in the two Eddington ratio distributions persists even considering two luminosity matched subsamples ($z \sim 6$ average Eddington ratio ~ 3 higher than lower redshift one). The high- z sources are accreting faster than the ones at $0.35 < z < 2.25$.

3. We calculated fluxes for the Mg II and Fe II lines. The obtained values are not always in agreement with the published ones, indicating that these measurements are dependent on the performed analysis. The previously observed scatter in the Fe II/Mg II line ratio at $z \sim 6$ is significantly reduced in this work: the BLRs in the highest- z QSOs studied to date all show comparable level of chemical enrichment.
4. No redshift evolution of the Fe II/Mg II ratio is observed for $4 < z < 6.5$. If we consider the Fe II/Mg II line ratio as a second order proxy for Fe/Mg, this indicates that the QSOs in our sample must have undergone a major episode of Fe enrichment in the few hundreds Myr preceding the cosmic time at which the sources are observed.
5. Our results do not show any significant dependence on the S/N of the spectra: the M_{BH} estimates do not correlate with the S/N per \AA of the Mg II line and the Fe II/Mg II line ratios do not show any clear dependence on the S/N per \AA of the spectral continuum.

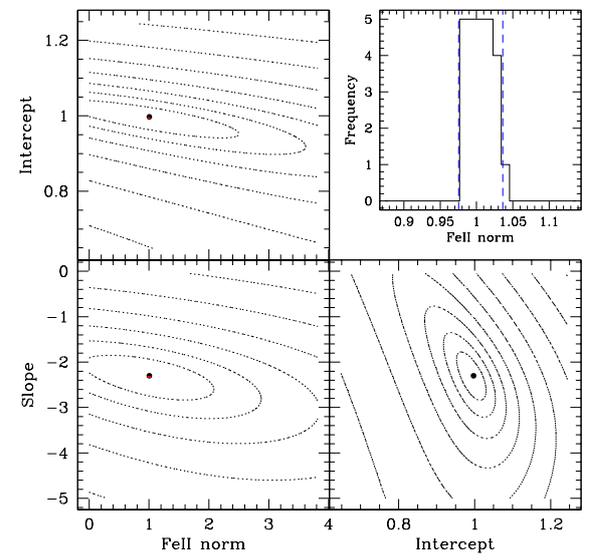
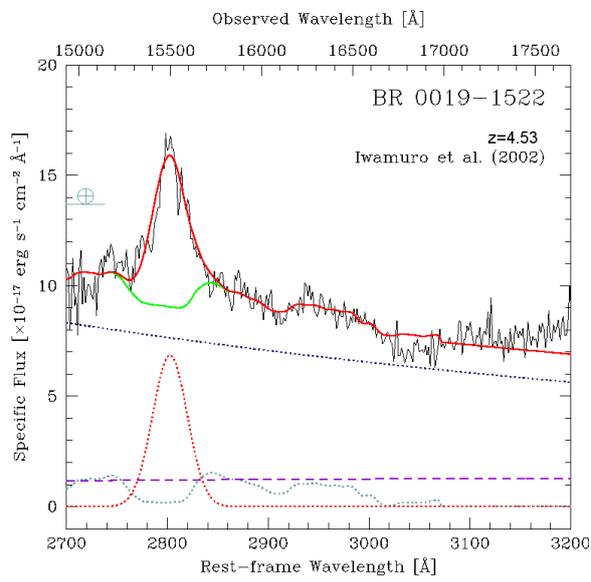
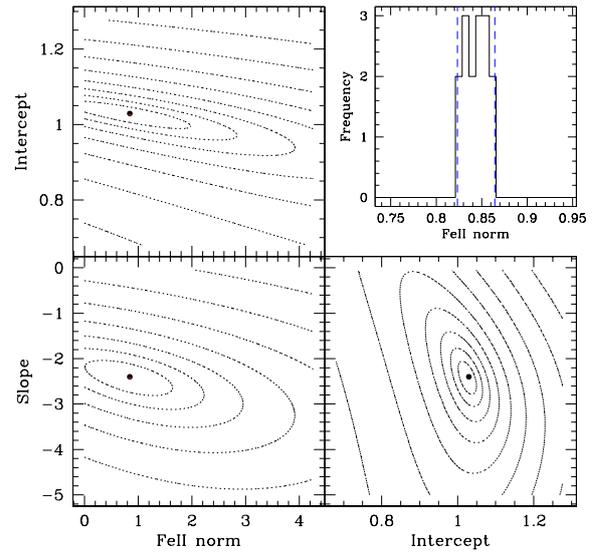
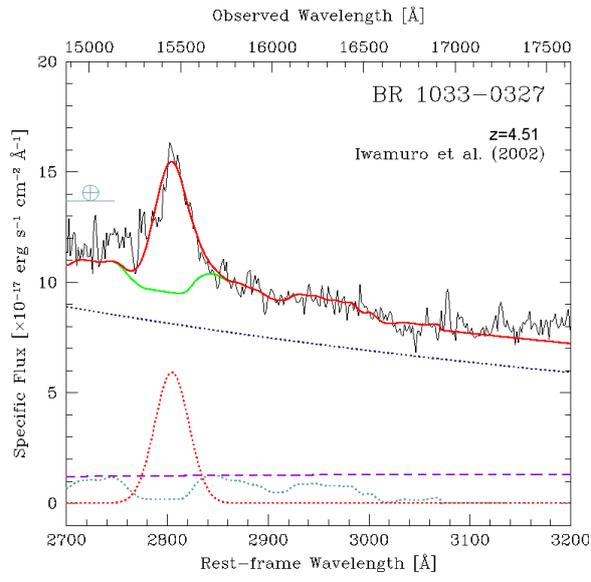
Even though the currently faintest known QSOs are included in the study we presented here (Willott et al. 2010a; Kurk et al. 2009), we are still not probing the bulk of QSOs that should be present at that redshift. Future studies of QSOs that populate the faint end of the QSO luminosity function at $z \sim 6$ are needed to investigate whether or not the results found here are applicable to all $z \sim 6$ QSOs.

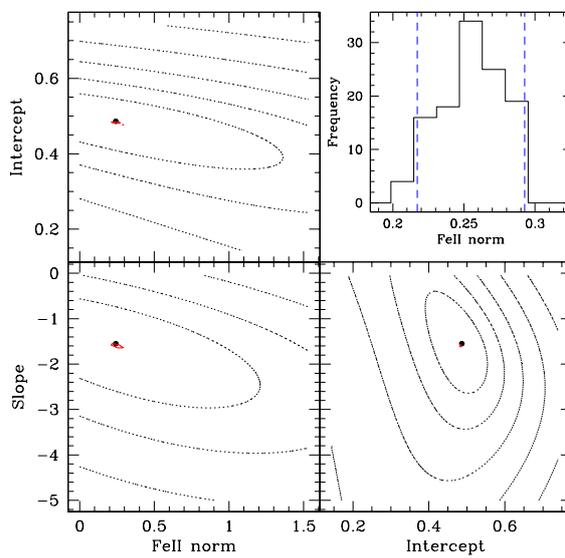
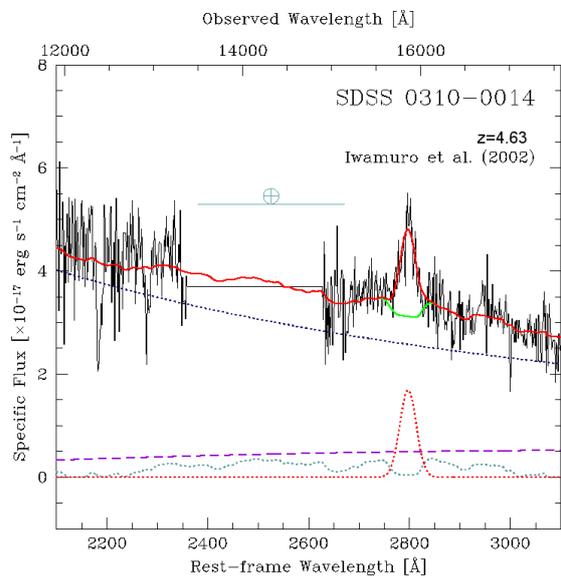
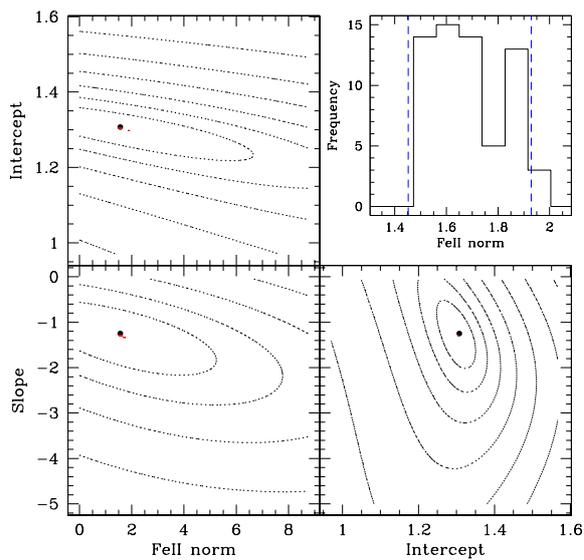
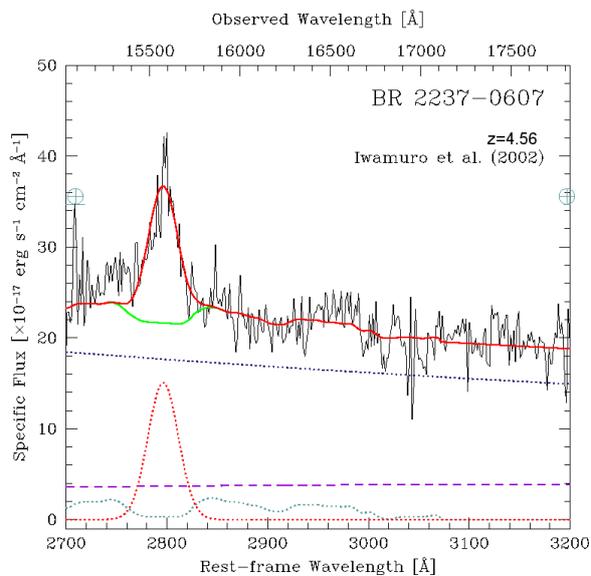
2.A FIT RESULTS: LITERATURE DATA

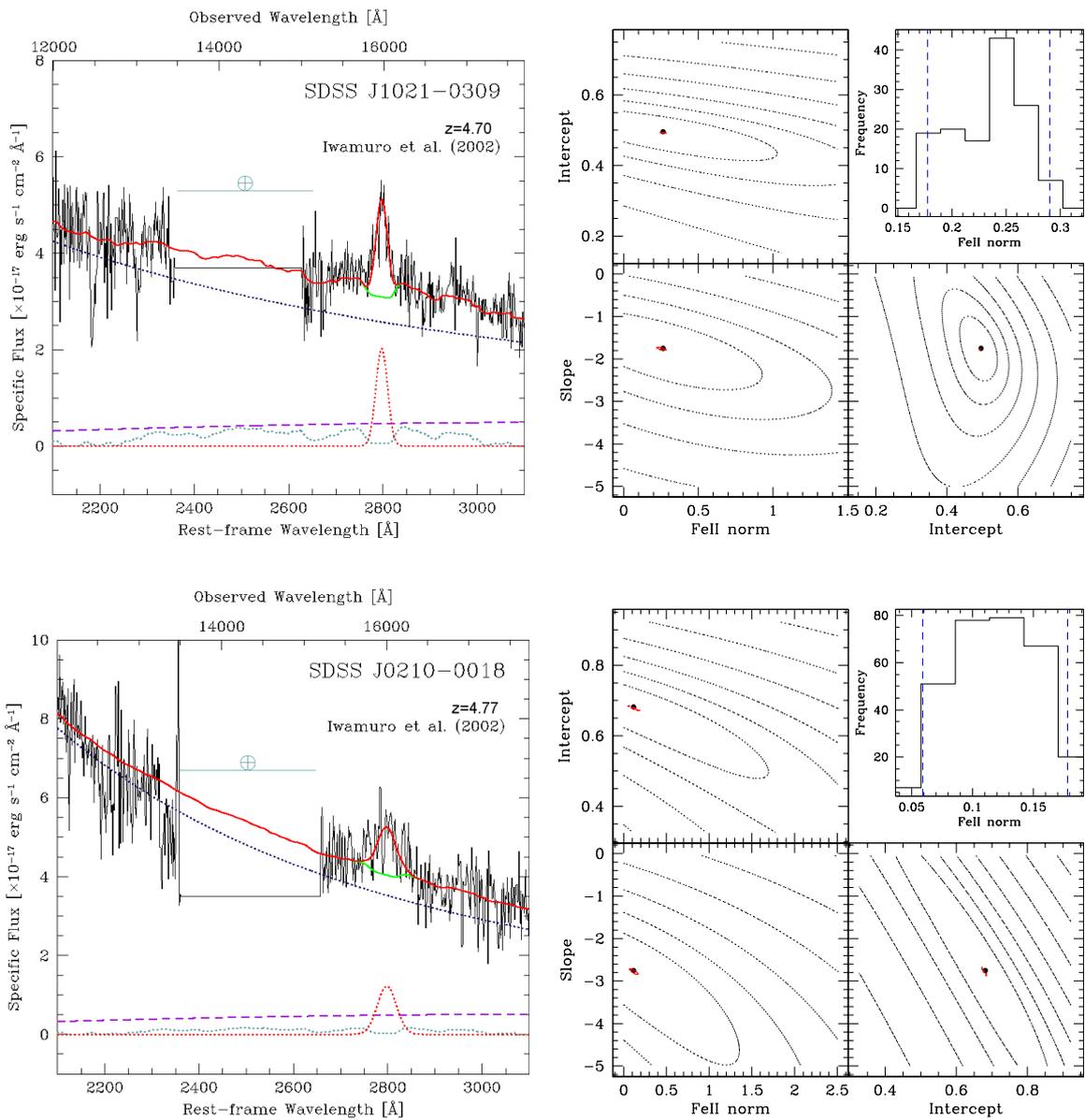
Hereafter we show the fit results for the literature sample and the relative χ^2 maps for Fe II error computation (analogous to fig. 2.2). The sources are sorted in redshift.

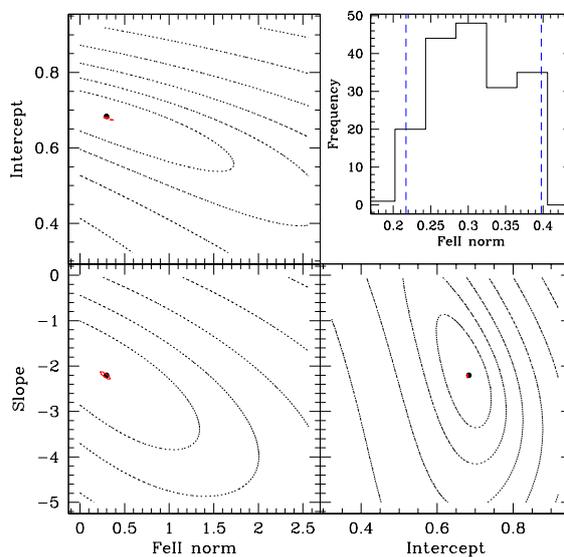
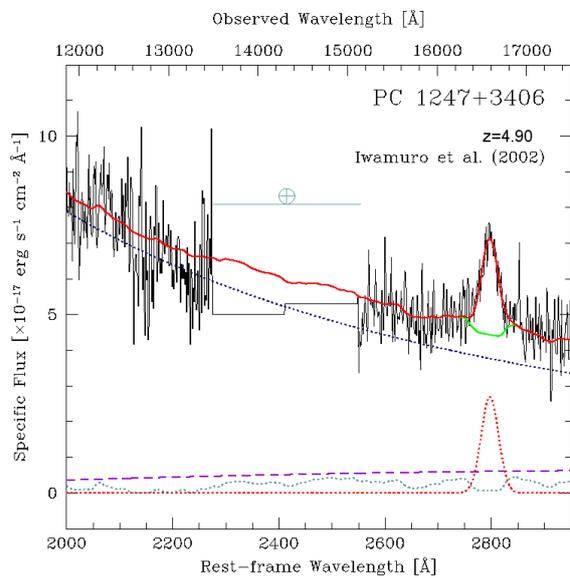
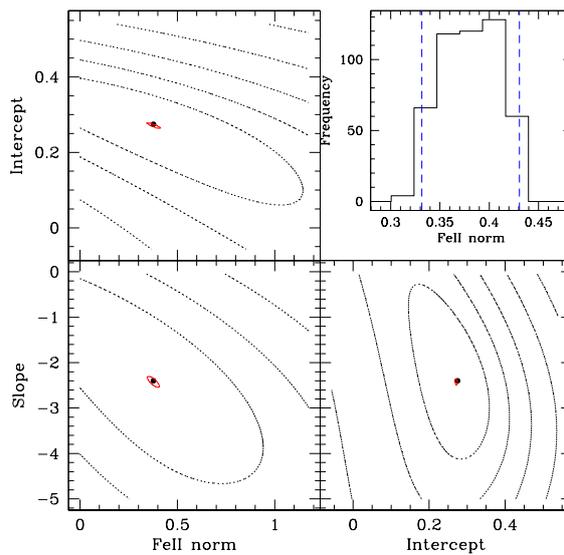
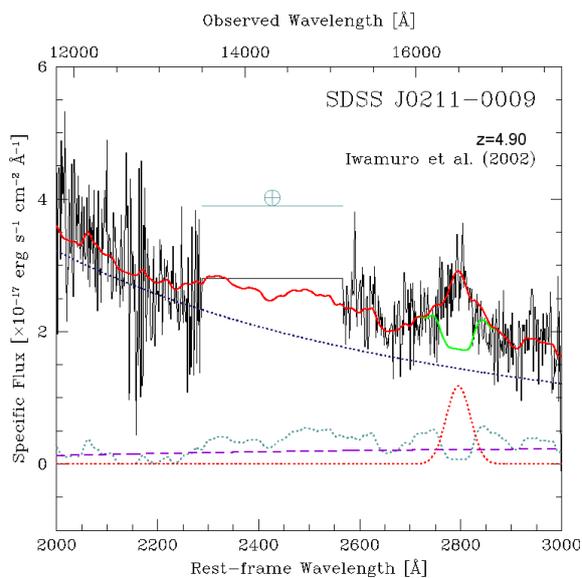
In the left panels we show the spectral decomposition. The observed spectra are shown as a black continuous line. The modeled components are: power-law continuum (blue dotted line), Balmer pseudo continuum (purple dashed line), Fe II normalized template (light blue dotted line), Mg II emission line (red dotted line). The sum of the first set of components (power-law continuum + Balmer pseudo continuum + Fe II normalized template) is overplotted to the spectrum as green solid line, while the sum of all the components is overplotted as a red solid line. Telluric absorption bands are indicated over the spectra with the symbol \oplus : they are extracted from the ESO sky absorption spectrum measured on the Paranal site at a nominal airmass of 1.

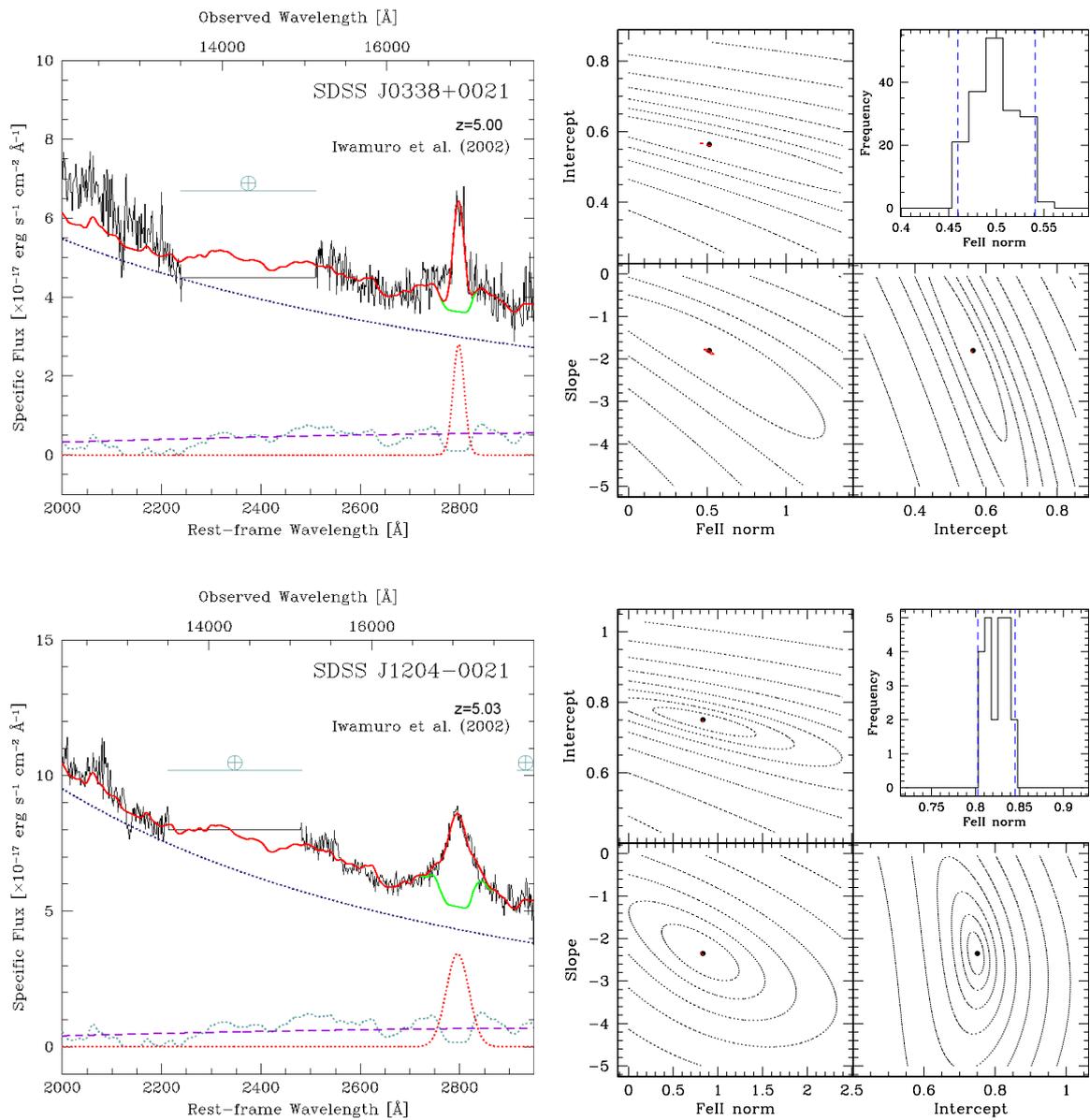
The Fe II normalization, obtained from the fit of the first set of components, depends on the power-law slope and its normalization (intercept). In the right panel we show the χ^2 domain analysis for Fe II error computation: a) two dimensional projections of the 3D χ^2 -surfaces (Fe II normalization vs Intercept, upper-left plot; Fe II normalization vs Slope, bottom-left plot; Intercept vs slope, bottom-right plot): contours represent iso- χ^2 levels spaced by a factor of 2 while the best fit case is marked with a dot; b) probability distribution for the Fe II template normalization (upper-right plot): the distribution has been obtained by marginalizing the 3-D probability distribution considering only the triplets for which $\chi^2 - \chi_{min}^2 < 1$, the dashed vertical lines mark our estimate of the $1 - \sigma$ confidence level.

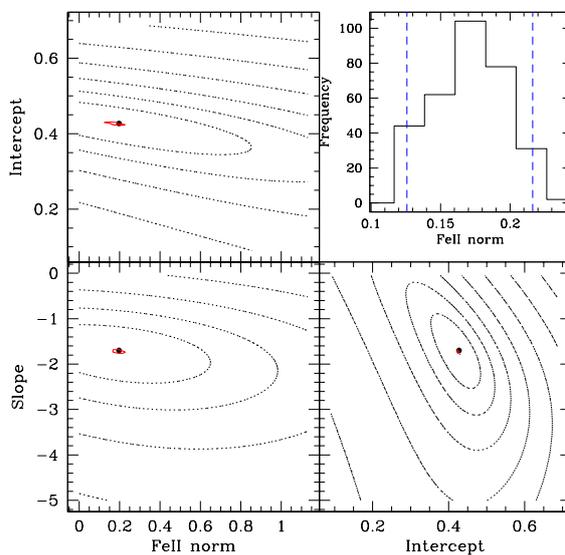
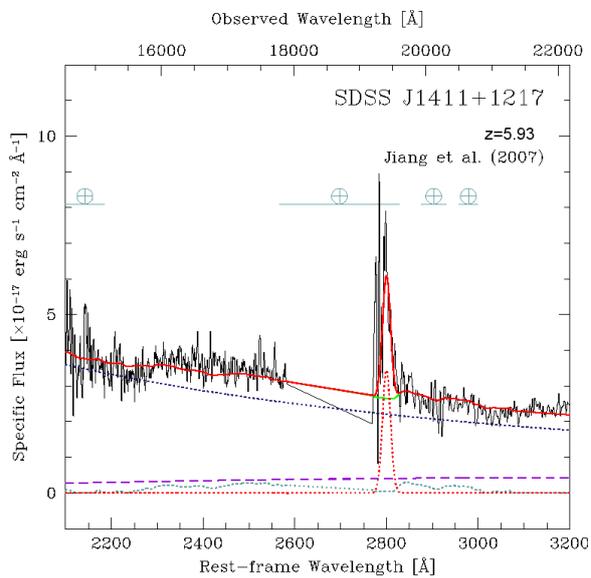
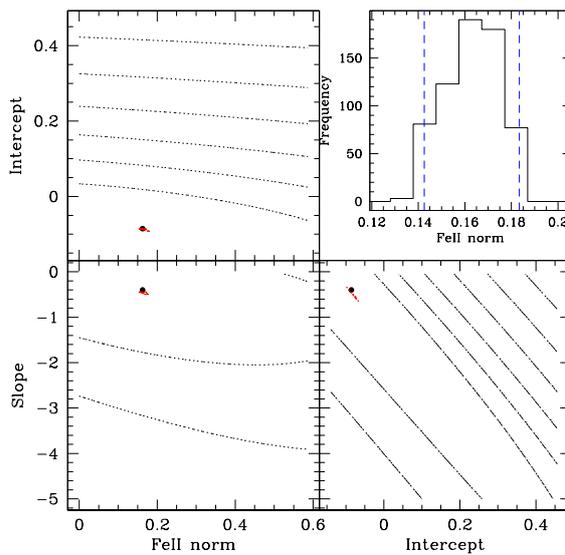
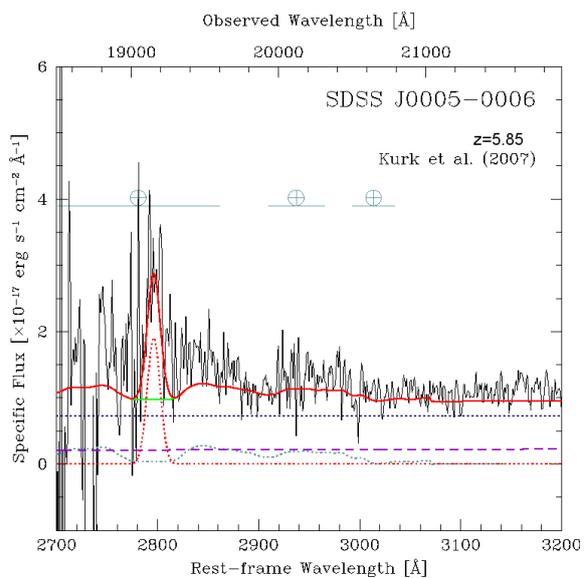


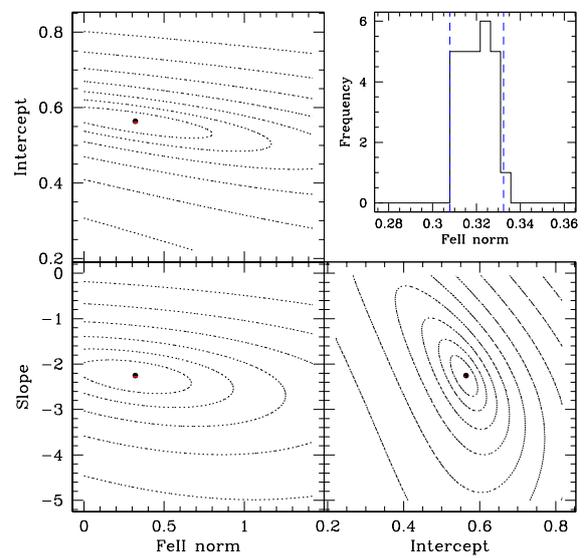
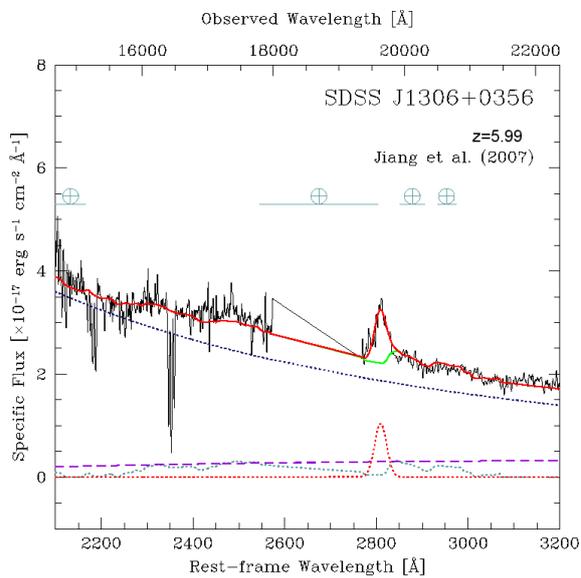
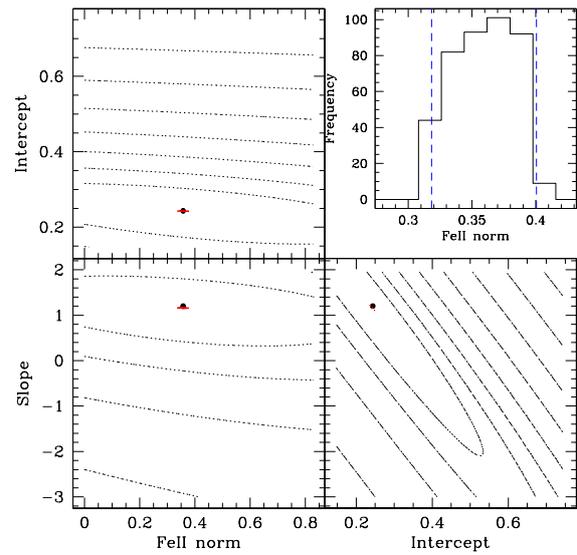
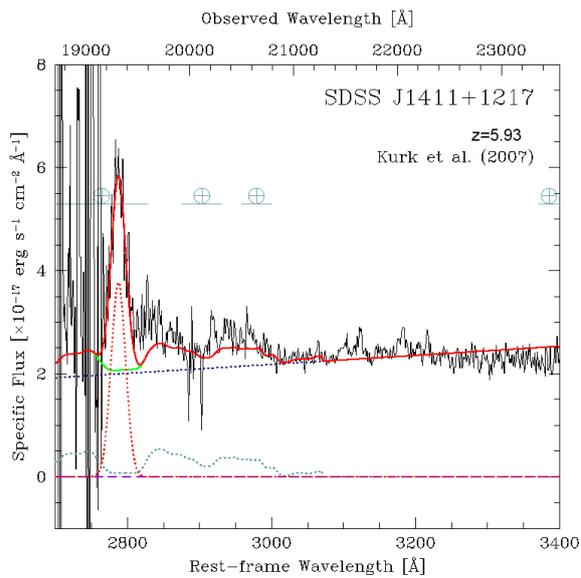


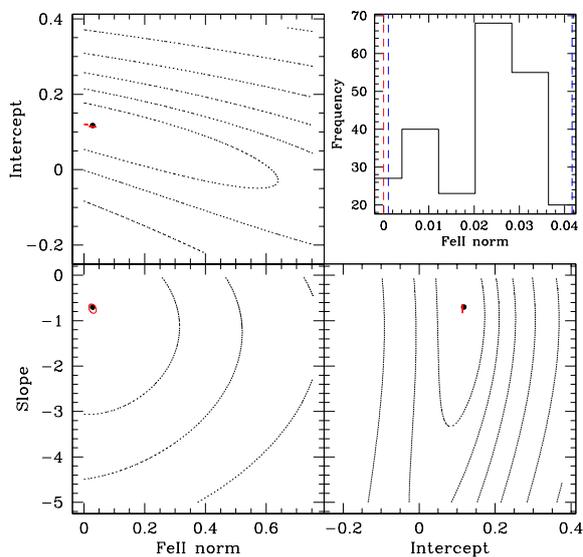
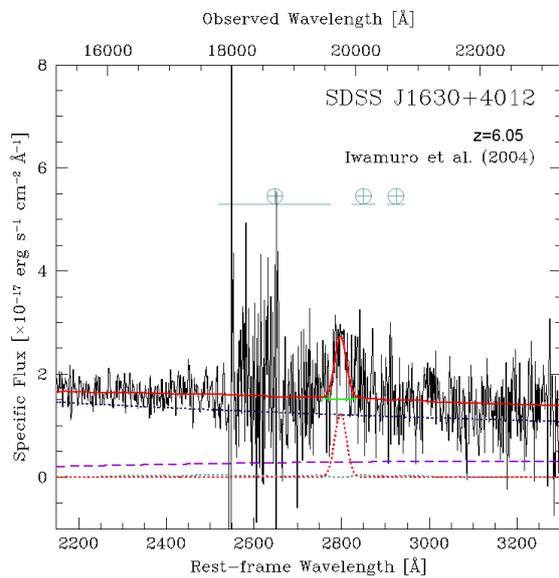
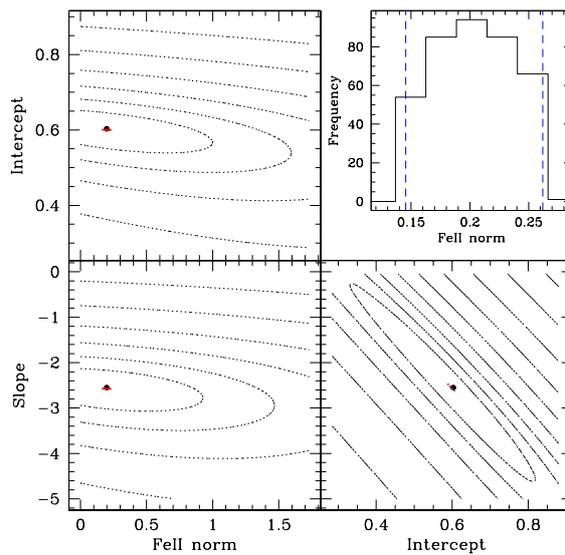
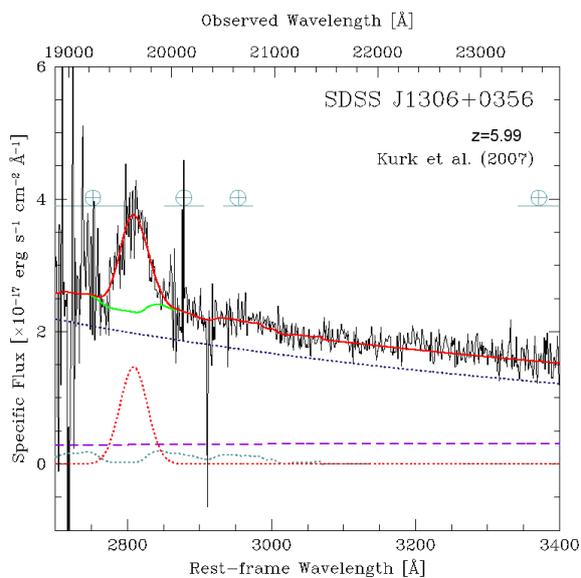


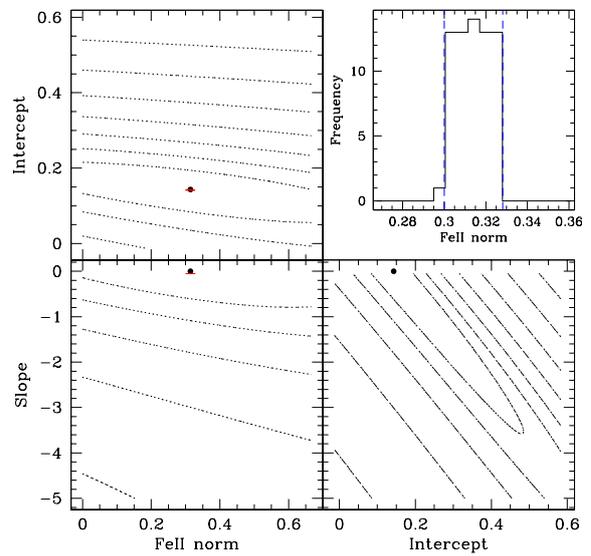
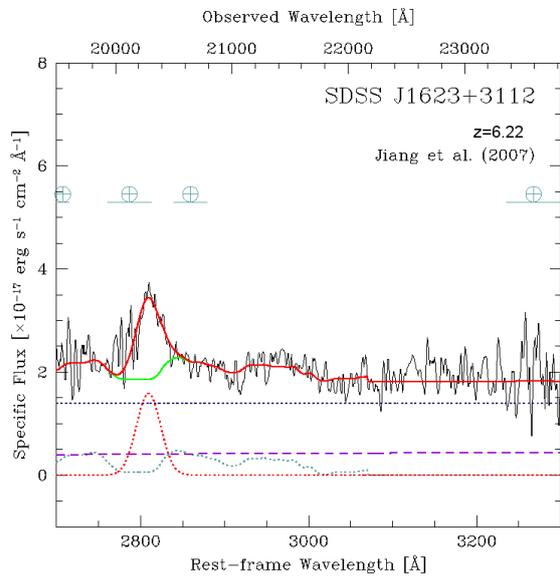
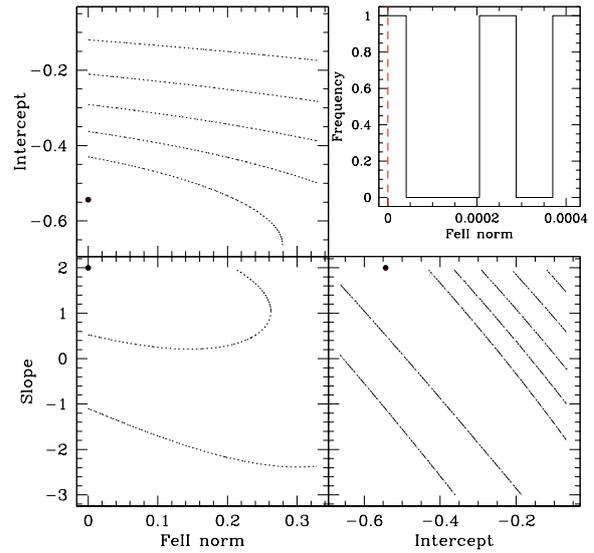
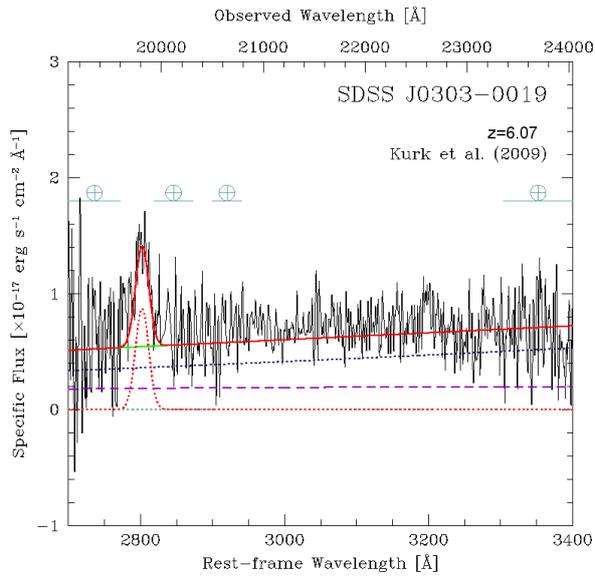


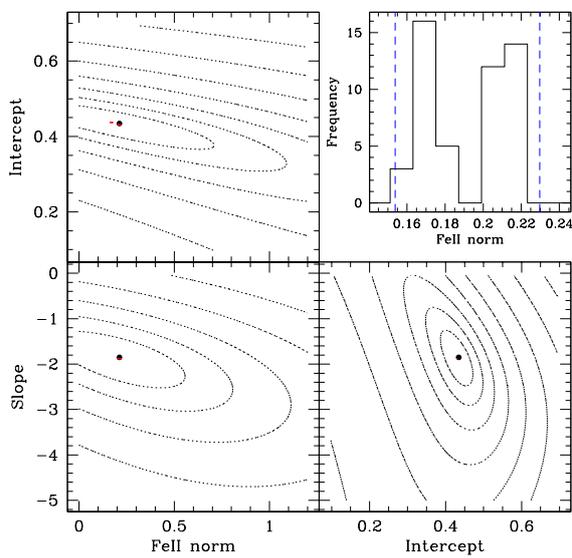
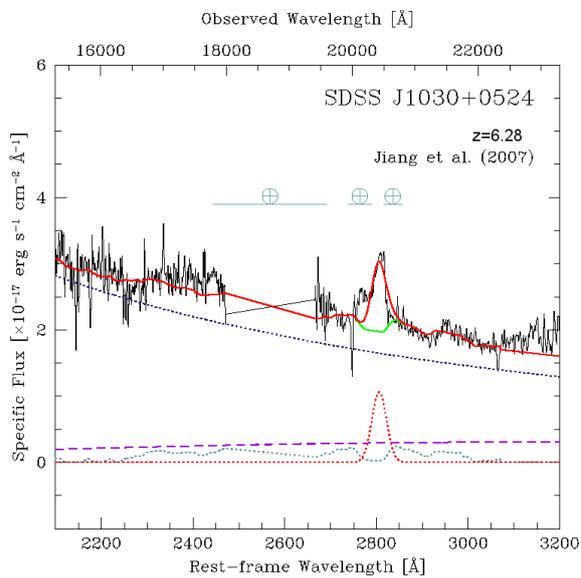
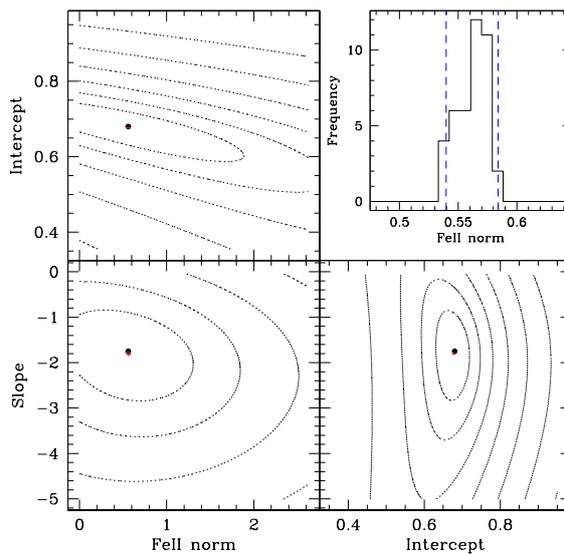
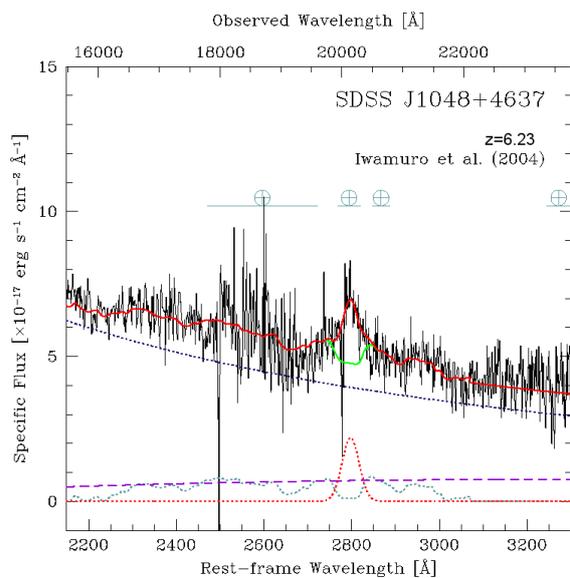


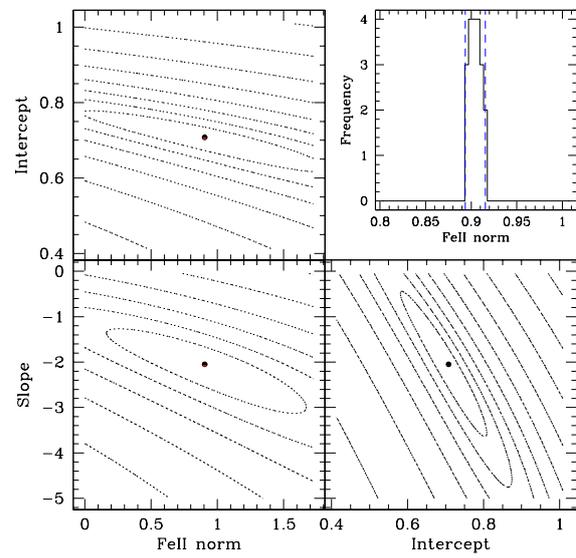
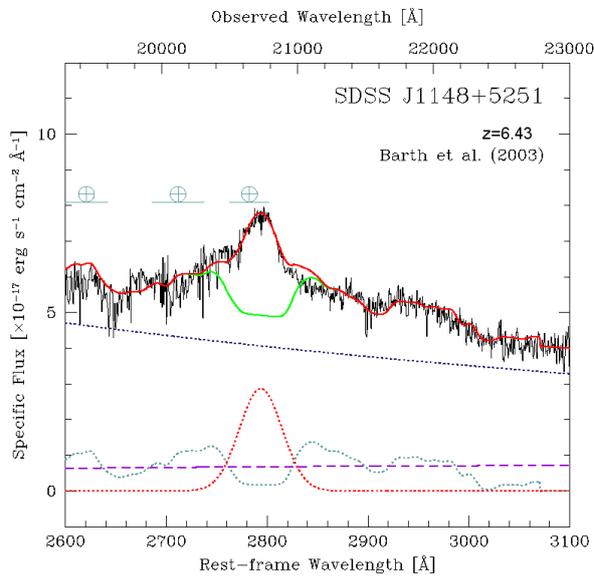
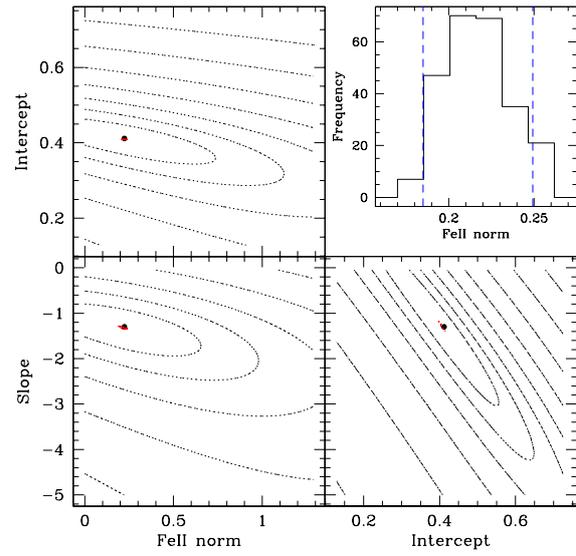
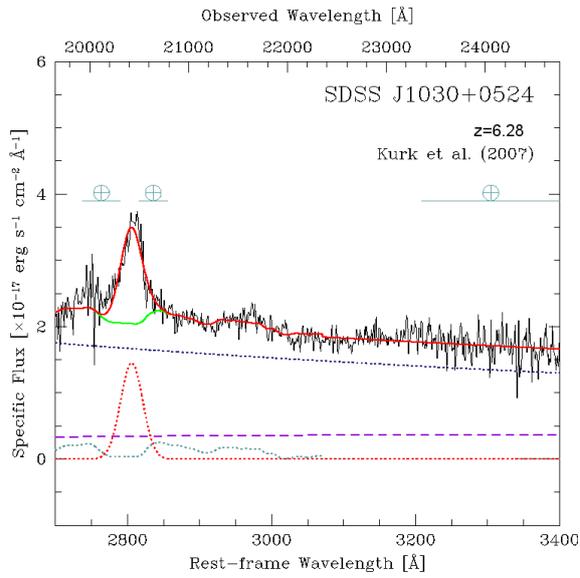












CONTEXT

The broad emission features in QSO spectra are thought to originate predominantly from photoionization of high velocity gas orbiting around the central supermassive black hole (BH). By studying the properties of the emission spectrum we can put observational constraints on the accretion mechanisms and on the dynamics of gas close to the BH, with the ultimate aim of constrain the nature of the link between the evolution of the central BH and that of the host galaxy. To first order the emission properties of UV-optically selected QSOs are surprisingly similar, but, going on higher levels of parametrization, we are able to identify significant variations in their spectral characteristics. At low redshift, the parameters that best characterize the diversity of the broad emission line region (BLR) properties are the FWHM of the $H\beta$ emission line and the relative strength of the Fe II emission with respect to $H\beta$ (see also the discussion about the Mg II/Fe II ratio presented in the previous chapter). The importance of these parameters was first pointed out by Boroson & Green (1992), who used a *principal component analysis* (PCA) in order to identify the features that cause the largest variance in QSO spectra. Since then, other parameters (including the X-ray spectral index α_{OX}) have been added to the set of intrinsic spectral features that accounts for the variance in QSO spectra (e.g. Wills et al. 1999; Laor 2000). Moreover, by using spectral PCA analysis, it has been possible to identify new correlations between emission lines, broad absorption lines and non-stellar continuum (e.g. Shang et al. 2003). In the (rest-frame) UV two well known properties of the C IV emission lines are evident sign-spot of spectral variations. The first one is the anti-correlation between the C IV equivalent width (EW) and the continuum luminosity, generally known as *Baldwin effect* (Baldwin 1977), while the second property is the

*The material presented in this chapter is partly included in the paper *Optical properties of $z \sim 6$ quasars*, De Rosa et al., 2011b, to be submitted

blue-shift of the C IV with respect to the systemic redshift (Gaskell 1982; Wilkes 1984). Finally, there can also be variations in the QSO continua: there is a non linear scaling between the 2 keV X-ray luminosity and the UV luminosity at 2500 Å (Green et al. 1995; Steffen et al. 2006; Just et al. 2007). This $L_{UV} - \alpha_{OX}$ relation is such that QSOs that are brighter in the UV are relatively weaker in the X-ray. In this chapter we will focus on the (rest-frame) UV emission properties of $z \sim 6$ QSOs.

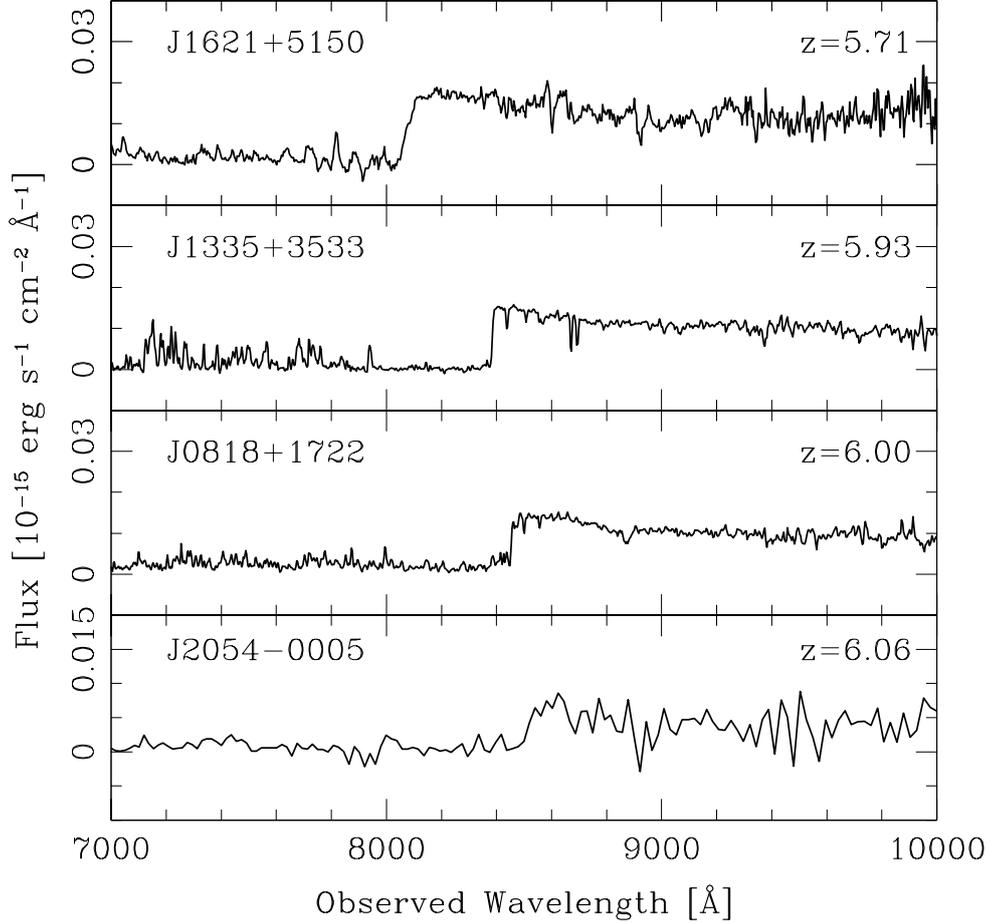
3.1 A SPECTROSCOPIC SAMPLE OF $z \sim 6$ QSOs

In order to perform this analysis, we consider all the $z \sim 6$ quasars for which high quality spectra are available. The sample consists of 26 $z \sim 6$ QSOs (see tab. 3.1), covering the redshift range $z = 5.70 - 6.42$ and a range of absolute magnitudes at the rest-frame wavelength $\lambda_{rest} = 1450$ Å, $M_{1450} = -25.4$ to -27.9 . The characteristics of the QSOs and of the respective spectra are summarized in tab. 3.1. Most of the sources in our sample (24 out of 26) have been discovered in SDSS (Fan et al. 2000, 2001, 2003, 2004, 2006; Jiang et al. 2008, 2009). J0841 + 2905 has been discovered by Goto (2006) with the Subaru telescope, while J1420 – 1602 is the first high-redshift QSO discovered in the Pan-STARRS survey (Morganson, De Rosa et al. 2011). The spectra have been observed either with the Echellette Imaging Spectrograph (ESI) mounted on the Keck II telescope, or with the Red Channel Spectrograph mounted on MMT. They are characterized by different exposure times and spectral resolutions, resulting in a range of signal to noise ratios (S/N).

WEAK LINE QSOs

Four of the sources in our sample are characterized by exceptionally weak emission lines (see fig. 3.1): J1621+5150, J1335+3533, J0818+1722 and J2054-0005. High redshift QSOs are defined as Weak Line QSOs (WLQs) if the rest-frame equivalent width of the Ly α + N V emission line complex is lower than 15.4 Å (Diamond-Stanic et al. 2009). The reason why these QSOs present almost featureless rest-frame UV spectra has not been understood yet. Spectroscopic monitoring of four WLQs suggests that the UV line weakness is intrinsic and cannot be explained by microlensing (Diamond-Stanic et al. 2009). Intrinsically weak emission lines might be explained by extremely fast accretion of the supermassive BH. In this case one would expect a peculiar continuum spectral energy distribution with a relative deficiency in high energy photons, resulting in the preferential suppression of high-ionization emission lines (e.g. Leighly et al. 2007,b; Vasudevan & Fabian 2007). Shemmer et al. (2010), analyzing the NIR and X-ray spectra of two WLQs, found that both high and low ionization lines were weak. At the same time

Figure 3.1: Observed spectra of the four WLQs in our sample (J1621+5150, J1335+3533, J0818+1722 and J2054-0005) sorted by redshift. Fluxes are in units of $10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$.



they found no anomalies in the continuum properties. Their results suggest that the weakness of the emission lines is the consequence of peculiar BLR physical properties (e.g. gas deficit or low covering factor). High S/N spectra covering large spectral regions would be needed in order to monitor the relative weakness of high and low ionization lines. Another possible explanation is that WLQs host BLRs in the early stages of their formation (Hryniewicz et al. 2010). In this case there might be an evolution in WLQs number density as a function of the look-back time. The redshift of a WLQ can often be determined only from the onset of the Ly α forest. For this reason, high redshift sources have been preferentially targeted. Up to date, ~ 80 WLQs have been discovered in SDSS with $z > 2.2$. To quantitatively constrain a possible evolution with the systemic redshift, WLQ samples spanning a wide range of redshifts and corrected for completeness would be needed. Given the intrinsic weakness of the emission lines, we do not include these four QSOs in the following analysis.

Table 3.1: QSO sample sorted by redshift.

QSO name ¹	z ²	M_{1450} ³ [mag]	E(V-B) ⁴ [mag]	A_V ⁵ [mag]	Discovery paper ⁶	Spectrograph ⁷
J162100+515000	5.71	–	0.0597	0.1845	DeRosaip	ESI
J142034-160248	5.73	-27.04	0.0803	0.2501	Morg11	Red
J092721+200123	5.79	-26.78	0.0331	0.1006	Fan06	ESI
J104433-012502	5.800	-27.47	0.0512	0.1592	Fan00	ESI
J000239+255034	5.80	-27.67	0.0367	0.1140	Fan04	Red
J083643+005453	5.82	-27.88	0.0495	0.1523	Fan01	ESI
J143611+500706	5.83	-26.54	0.0176	0.0527	Fan06	Red
J084035+562459	5.85	-26.66	0.0413	0.1276	Fan06	ESI
J000552-000655	5.85	-25.87	0.0335	0.1037	Fan04	ESI
J133550+353315	5.93	-26.68	0.0130	0.0385	Fan06	ESI
J141111+121737	5.93	-26.74	0.0243	0.0762	Fan04	ESI
J084119+290504	5.96	-26.86.	0.0536	0.0538	Goto06	ESI
J130608+035626	5.99	-27.19	0.0280	0.0885	Fan01	ESI
J081827+172251	6.00	-27.40	0.0352	0.1090	Fan06	ESI
J113717+354956	6.01	-27.12	0.0180	0.0568	Fan06	ESI
J035349+010404	6.049	-26.53	0.2934	0.9294	Jiang08	ESI
J205406-000514	6.062	-26.15	0.0964	0.2969	Jiang08	Red
J160254+422822	6.07	-26.83	0.0142	0.0440	Fan04	Red
J030331-001912	6.070	-25.48	0.1091	0.3352	Jiang08	ESI
J084229 + 121850.5	6.08	–	0.0665	0.2045	DeRosaip	ESI
J231546-002358	6.117	-25.43	0.0383	0.1173	Jiang08	ESI
J125051+313021	6.13	-27.14	0.0134	0.0409	Fan06	ESI
J104845+463718	6.18	-27.55	0.0182	0.0570	Fan03	ESI
J162331+311200	6.22	-26.67	0.0224	0.0681	Fan04	ESI
J103027+052455	6.301	-27.16	0.0231	0.0718	Fan01	ESI
J114816+525150	6.43	-27.82	0.0231	0.0697	Fan03	ESI

¹ QSO name, right ascension and declination are in J2000.0;^{2,3} optical redshift and absolute magnitude M_{1450} from discovery paper;^{4,5} Reddening and visual extinction from the map of Schlegel et al. (1998);⁶ discovery paper: DeRosaip De Rosa et al. (2011, in preparation), Fan00 Fan et al. (2000), Fan01 Fan et al. (2001), Fan03 Fan et al. (2003), Fan04 Fan et al. (2004), Fan06 Fan et al. (2006), Jiang08 Jiang et al. (2008), Goto2006 Goto (2006), Morg11 Morganson, De Rosa et al. (2011);⁷ Telescope and spectrograph, Keck ESI; MMT Red Channel;

Table 3.2: Quasar emission lines in the wavelength range $\lambda = [900, 1500]$ Å. Rest-wavelengths are from Vanden Berk et al. (2001) except for N v doublet: in this case we consider the mean of the doublet wavelengths (1238.82, 1242.80) (atomic list, <http://www.pa.uky.edu/peter/atomic>). For blended lines we are quoting the central wavelength measured from the SDSS composite spectrum and the mean ionization potential.

Line	Central wavelength [Å]	Ionization potential [eV]
Ly β + O IV	1033.03	45.51
Ly α	1215.67	13.60
N v	1240.81	97.89
Si II	1262.59	16.35
O I+Si II	1305.42	14.98
C II	1335.30	24.38
Si IV+O IV]	1398.33	61.28

3.2 DATA ANALYSIS

We focus on the rest-frame wavelength range $\lambda_{Range} = [900, 1500]$ Å. Since the QSO redshifts are varying between $z = 5.72 - 6.42$ and the sources have been observed with different instruments, the spectral coverage is not homogeneous all over λ_{Range} . The sample is also characterized by a variety of S/N ratios and spectral resolutions. In tab. 3.2 we list the principal QSO emission lines in λ_{Range} with their rest-frame central wavelengths and ionization potentials. Having different spectral resolutions does not represent a problem for the majority of the emission lines in analysis: the N v, O I+Si II and Si IV+O IV] doublets are not resolved even in the spectra with the highest resolution. The deconvolution of the Ly α emission complex (Ly α + N v+Si II) can instead be a function of the spectral resolution. Possible differences in the spectral decomposition will be discussed individually.

3.2.1 CORRECTING FOR GALACTIC EXTINCTION

In tab. 3.1 we list the reddening E(B-V) and the extinction in the visual band A_V obtained from the map of Schlegel et al. (1998). Our analysis of the emission properties is focused on the observable flux over a short wavelength range: $[7500 - 10500]$ Å. We have computed the extinction A_λ at $\lambda = 7500, 8200, 10000$ Å using the algorithm by Cardelli et al. (1989). Since the mean difference between the total extinction at 7500 and 10000 Å ($\langle A_{7500} - A_{10000} \rangle = 0.03^{+0.06}_{-0.03}$) is less than our average flux error ($\sigma_f \sim 10\%$), we decide to correct each QSO flux for a constant ex-

tion equal to the mean A_λ computed over the wavelength range. For the analysis of the $\text{Ly}\alpha$ and $\text{Ly}\beta$ integrated fluxes (see chap. 4), we correct each value for the extinction computed at the line central wavelength.

3.2.2 SPECTRAL MODEL

We are interested in fitting all the emission lines from $\text{Ly}\alpha$ to $\text{Si IV}+\text{O IV}$. We adopt a commonly used spectral model consisting in a power-law non stellar continuum, a double-Gaussian for the $\text{Ly}\alpha$ line and a single Gaussian for each of the N V , Si II , $\text{O I}+\text{Si II}$, C II and $\text{Si IV}+\text{O IV}$ lines. [Kramer & Haimann \(2009\)](#) analyzed the uncertainties arising in the determination of the $\text{Ly}\alpha$ intrinsic flux when the model is fit to the unabsorbed red wing of the line. Concerning the continuum modeling, they found that when the fitting windows are only red-ward the $\text{Ly}\alpha$ line ($1220 \text{ \AA} < \lambda < 1500 \text{ \AA}$), the extrapolation of the power-law to shorter wavelengths is more reliable if the slope is fixed to -1.3 (median value measured for low redshift QSOs, e.g. [Decarli et al. 2010](#)). Note that this might not be a reliable assumption, but in their analysis it yielded to better extrapolations with respect to the case with free continuum slope. Following [Kramer & Haimann \(2009\)](#) we fix the power-law slope to -1.3. The $\text{Ly}\alpha$ line detected in low redshift QSOs shows two prominent components: a broad component ($\text{Ly}\alpha_B$) and a narrow one ($\text{Ly}\alpha_N$). The two components are thought to arise from gas close to, and relatively further away from the central black hole. Following [Kramer & Haimann \(2009\)](#) we try to fit the $\text{Ly}\alpha$ emission with two independent Gaussians. For the majority of the sources though, a second component is not necessary and the fit is successful with a single Gaussian. Although N V is actually a doublet (1238.81 and 1242.80 \AA) we fit it with a single Gaussian since the two components are broad and difficult to de-blend at these spectral resolutions. The same considerations apply for $\text{O I}+\text{Si II}$ and $\text{Si IV}+\text{O IV}$ lines. Details on the model used for each QSO can be found in section 3.2.6.

The model has 3 free parameters for each Gaussian (central wavelength, width and amplitude) and 1 free parameter for the power-law (normalization). In practice, the power-law is fit first, then the lines are fit to the continuum-subtracted spectrum. Fitting the $\text{Ly}\alpha$ complex ($\text{Ly}\alpha + \text{N V} + \text{Si II}$) by using only the $\text{Ly}\alpha$ red wing can lead to poorly constrained results (there is not enough information to constrain three Gaussian parameters for the three independent components). We then perform the fit twice for each source: the first time all the parameters of the Gaussian function are left free (*M-free*) while the second time we fix the central wavelength of the $\text{Ly}\alpha$ line (*M-fix*) using a reliable estimate of the QSO systemic redshifted (as explained in sec. 3.2.4). **This way we obtain two independent best-fit models that allow us to study how the estimated emission properties depend on the particular spectral model adopted for the $\text{Ly}\alpha$.**

3.2.3 PROCEDURE FOR LINE FITTING

Each fit is performed using the Levenberg-Marquardt χ^2 minimization. Standard fitting windows are defined in the rest-frame system of reference, but they are individually tuned through visual inspection. This is done because we do not know the exact QSO systemic redshift so the fitting windows might result not exactly centered on the feature of interest. Moreover, we want to exclude strong absorption lines due to either the intergalactic medium or the sky telluric. Final fitting windows are shown for each source in the respective plots (see fig. 3.2 to fig.3.7). We perform our fits separately for the continuum and each blended system of lines. First we fit the power-law alone to regions of the spectra chosen to be free of emission line features. The standard power-law fitting windows are: [1280, 1290], [1315, 1320] and [1350, 1360] Å. The non-stellar continuum flux is parametrized as:

$$F_\lambda = a_{cont} * \left(\frac{\lambda}{\lambda_0}\right)^{-1.3} \quad (3.1)$$

where $\lambda_0=1215.67$ Å and a_{cont} is the power-law normalization. The initial value for the normalization is:

$$a_{cont} = \bar{F}_\lambda \quad (3.2)$$

where \bar{F}_λ is the median flux in the continuum fitting regions. Once the normalization of the power-law is determined, we subtract the non-stellar continuum and fit the emission line features to the continuum-subtracted spectrum.

Ly α COMPLEX

Ly α , N v and Si II lines are fit at the same time to the rest-frame spectral region [1220, 1290] Å. The lower limit is chosen in order to avoid all the resonant absorption, including the ones from any foreground gas falling towards the QSO that would absorb light on the red side of the line center (Barkana & Loeb 2003; Mesinger & Haimann 2007; Kramer & Haimann 2009). The higher limit is chosen in order to include a continuum dominated window, to help constraining the wings of significantly broad Gaussian components (Kramer & Haimann 2009). Kramer & Haimann (2009), analyzing a sample of 86 low-redshift unobscured QSOs, found that the Ly α_N , Ly α_B and N v lines show median shifts of -303, -413 and -45 km s⁻¹ with respect to the transition rest-frame wavelengths (corresponding to -1.229, -1.675 and -0.182 Å). They also showed that the extrapolation of the Ly α blue flux from the fit of its red-side alone ($\lambda_{rest} > 1220$ Å) is significantly improved if the central wavelengths of the Gaussian components are fixed considering these median shifts. Therefore we define our set of initial parameters (width, σ , central wavelength, μ , and amplitude, A) considering the same wavelength shifts:

$$\begin{aligned}\sigma_{Ly\alpha_N} &= 4 \text{ \AA}, \quad \mu_{Ly\alpha_N} = 1215.67 - 1.23 \text{ \AA}, \quad A_{Ly\alpha_N} = \bar{F}_G \\ \sigma_{Ly\alpha_B} &= 16 \text{ \AA}, \quad \mu_{Ly\alpha_B} = 1215.67 - 1.67 \text{ \AA}, \quad A_{Ly\alpha_B} = \bar{F}_G\end{aligned}\tag{3.3}$$

$$\begin{aligned}\sigma_{NV} &= 3 \text{ \AA}, \quad \mu_{NV} = 1240.81 - 0.18 \text{ \AA}, \quad A_{NV} = 2\bar{F}_G \\ \sigma_{SiII} &= 6 \text{ \AA}, \quad \mu_{SiII} = 1262.59 \text{ \AA}, \quad A_{SiII} = \bar{F}_G\end{aligned}\tag{3.4}$$

where \bar{F}_G is the median flux in the Gaussian fitting region and the flux of the i -th Gaussian component is parametrized as:

$$F_{\lambda,i} = a_i \exp\left(\frac{-(\lambda - \mu_i)^2}{2\sigma_i^2}\right).\tag{3.5}$$

Note that initial values are not necessarily representative of the final best fit parameters. The subtraction of the measured median shift to the line central wavelength is particularly important for the Ly α components, since in *M-fix* we are fixing their central wavelengths to the initial values. As we have anticipated already, for many of the sources in our sample the Ly α is well fitted by using a single Gaussian model. In case of a single Gaussian Ly α , our set of initial parameters is:

$$\sigma_{Ly\alpha} = 18 \text{ \AA}, \quad \mu_{Ly\alpha} = 1215.67 - 1.45 \text{ \AA}, \quad A_{Ly\alpha} = \bar{F}_G\tag{3.6}$$

where the shift applied to the central wavelength is computed as the mean between the narrow and broad component shifts.

O I+Si II

The O I+Si II system is fit as a single Gaussian to the rest-frame window [1290, 1318] Å. Since this line is a doublet, we cannot use its laboratory transition wavelength as reference wavelength. We then estimate the O I+Si II reference wavelength directly from the SDSS composite spectrum of lower redshift QSOs (Vanden Berk et al. 2001). The underlying assumption is that for $\lambda_{rest} > 1220$ the low-redshift QSO template represents the high redshift sample accurately enough (see also sec. 3.3). We fit the SDSS composite spectrum using our procedure and adopting the same parametrization of the non-stellar continuum. We obtain $\lambda_{O I+Si II} = 1305.317 \text{ \AA}$, which is slightly different with respect to the one measured by Vanden Berk et al. (2001) ($|\Delta\lambda| = 0.1$). In the following, we assume the O I+Si II rest-frame wavelength $\mu_{O I+Si II} = 1305.317 \text{ \AA}$. This reference

wavelength will be a key parameter in the measurement of the QSO systemic redshift. The O I+Si II set of initial parameters is:

$$\sigma_{\text{O I+Si II}} = 6 \text{ \AA}, \mu_{\text{O I+Si II}} = 1305.32 \text{ \AA}, A_{\text{O I+Si II}} = \bar{F}_G \quad (3.7)$$

C II

The C II line is fit as a single Gaussian component to the rest-frame window [1325, 1348] Å. The considered initial parameters are:

$$\sigma_{\text{C II}} = 5 \text{ \AA}, \mu_{\text{C II}} = 1335.30 \text{ \AA}, A_{\text{C II}} = \bar{F}_G \quad (3.8)$$

Si IV+O IV]

The Si IV+O IV] system is fit as a single Gaussian to the rest-frame window [1360, 1446] Å. Following what we have done for the O I+Si II doublet, we measure the rest-frame wavelength directly from the SDSS composite spectrum. The resulting set of initial parameters is the following:

$$\sigma_{\text{Si IV+O IV]}} = 12 \text{ \AA}, \mu_{\text{Si IV+O IV]}} = 1397.73 \text{ \AA}, A_{\text{Si IV+O IV]}} = \bar{F}_G \quad (3.9)$$

3.2.4 SYSTEMIC REDSHIFT

The main caveat to the outlined fitting procedure consists in knowing the source systemic redshift. It is well known (see e.g. Richards et al. 2002, 2011; Shang et al. 2007) that in QSO spectra high-ionization lines (e.g. C IV) can have high velocity offsets, of order 1000 km s^{-1} ($\Delta z \sim 0.02$ at $z \sim 6$), with respect to low-ionization lines (e.g. Mg II). To obtain a reliable redshift measurement, one can assume that the BLR is at the QSO systemic redshift and use low-ionization lines such as Mg II to estimate it (Jiang et al. 2007; Kurk et al. 2007; De Rosa et al. 2011). For $z \sim 6$ QSOs the Mg II line falls in the near-infrared K-band (see chap. 2). An alternative way to directly obtain the measurement of the QSO systemic redshift is to detect the CO molecular gas in the host-galaxy (Wang et al. 2010). $z_{\text{Mg II}}$ is available for a total of 11 sources in our sample while z_{CO} is available for only 5 of them (see tab. 3.4). In order to perform our consistent analysis we need a homogeneous estimate of the systemic redshift. Since the O I+Si II emission complex is visible in most of the optical spectra of $z \sim 6$ QSOs, we want to test whether the redshift inferred from this line can be used as an estimate of the QSO systemic redshift. Considered individually, the two components of this emitting system have very low

ionization potentials (O I = 13.62 eV and Si II = 16.35 eV) which are comparable to the Mg II one (15.03 eV). Therefore, velocity shifts of individual components, if observed, are expected to be small. Unfortunately the O I+Si II doublet is unresolved. We then decide to test the validity of this approach by comparing the measured z_{OI} with the available $z_{\text{Mg II}}$ (De Rosa et al. 2011). We find that the average difference between the two estimates is $\Delta z = 0.004$ ($\sim 200 \text{ km s}^{-1}$) with a $1\text{-}\sigma$ dispersion of 0.003. Since this difference is significantly lower than the typical systematic error in our redshift measurements (0.01), we conclude that z_{OI} represents a good proxy for the QSO systemic redshift. For sources with no O I+Si II line we assume the systemic redshift to be equal to either $z_{\text{Mg II}}$ or z_{CO} . When none of these estimates is available we adopt the discovery paper optical redshift but do not include the sources in the subsequent analysis where z_{sys} is a key parameter. For a more detailed discussion on the redshift estimates see sec. 3.4.

3.2.5 SUMMARY OF THE FITTING PROCEDURE

Our procedure can be summarized as follows:

- We assume the optical redshift from the discovery paper as the first guess for the QSO systemic redshift. After red-shifting the spectrum to the rest-frame system of reference, we perform a preliminary fit of the power-law continuum and of the O I+Si II line. This way we obtain a first estimate of $z_{\text{OI,ini}}$ that allows a better centering of the fitting windows.
- We redshift the spectrum to the rest-frame using $z_{\text{sys}} = z_{\text{OI,ini}}$. We fit the power-law continuum and all the detected emission lines keeping free all the line parameters (*M-free*). We obtain a second estimate of $z_{\text{OI,fin}}$. Generally this value is in good agreement with the first estimate ($\langle \Delta(z_{\text{OI,ini}} - z_{\text{OI,fin}}) \rangle = 0.001$).
- We assume the QSO systemic redshift to be equal to $z_{\text{OI,fin}}$. After red-shifting the spectrum to the rest-frame system of reference, we fit once more the power-law continuum and the Ly α complex, this time fixing the Ly α line central wavelengths (*M-fix*).

3.2.6 FIT RESULTS AND NOTES ON INDIVIDUAL SOURCES

Fit results are shown in fig. 3.2 to fig. 3.7. Sources are sorted by redshift. In black we plot the observed spectrum smoothed over 10 \AA . The red solid line shows the results for *M-free* while the green solid line represents the final model for *M-fix*. If we compare the two models (free vs fixed Ly α central wavelengths), we notice that for 8 sources out of 22 the Ly α line results much broader if we fix its central wavelength considering the O I+Si II redshift. Assuming that our estimate of

the systemic redshift is reliable, we can conclude that for these sources the Ly α line is severely absorbed and the *M-free* model represents only a lower limit to the real Ly α flux. Results for the fit parameters with relative errors are listed in the appendix tables (tab. 3.6-tab. 3.10). We indicate with * the sources for which we use a double Gaussian to model the Ly α line. These sources are usually characterized by extremely narrow Ly α lines. Errors are computed in a standard way from the covariance matrix and rescaled for the χ^2 value, to account for the fit goodness. In some cases the parameters are poorly or barely constrained, especially for the Ly α complex. This can be the result of multiple factors: (1) low S/N; (2) lack of part of the line due to the presence of IGM/sky absorption; (3) overlap of multiple emitting features that are difficult to disentangle at these spectral resolutions, especially if the line peaks are not prominent. In the following we list sources with peculiar properties:

- **J1420-1602**, $z_{\text{sys}} = 5.728$: this QSO has been recently discovered in the Pan-STARRS survey (Morganson, De Rosa et al. 2011). It presents a strong absorption blue-ward the Si IV+O IV] line, which consequently is poorly constrained. The *M-fix* model suggests an almost complete absorption of the Ly α line. The most probable scenario is that this source is a broad absorption line (BAL) QSO. A high S/N near infrared spectrum is needed for confirmation.
- **J1044-0125**, $z_{\text{sys}} = 5.777$: discovered by Fan et al. (2000), this source is a BAL. The *M-fix* model correctly predicts that the Ly α line is almost completely absorbed.
- **J1436+5007**, $z_{\text{Ly}\alpha} = 5.83$: discovered by Fan et al. (2006), this spectrum has a very low S/N. The only lines with enough S/N are the Ly α complex ones. No reliable estimate of z_{sys} exists for this source.
- **J0005-0006**, $z_{\text{sys}} = 5.849$ and **J1411+1217**, $z_{\text{sys}} = 5.901$: this QSOs, discovered by Fan et al. (2004), present very narrow Ly α lines, which have been modeled as double Gaussian. The Si II components result poorly constrained in the *M-free* models, while the situation improves if the Ly α central wavelengths are fixed.
- **J0841+290**, $z_{\text{sys}} = 5.959$: discovered by Goto (2006), this spectrum has a very low S/N. The *M-free* model is consistent with no N V emission. Even if the O I+Si II line is detectable, we do not trust the redshift estimate given the high level of noise.
- **J1306+0356**, $z_{\text{MgII}} = 6.017$: discovered by Fan et al. (2001), this source has a very narrow Ly α line that we model as a double Gaussian. The O I+Si II line is not detected, so we assume $z_{\text{sys}} = z_{\text{MgII}}$. If we fix the Ly α central wavelengths using this redshift estimate, the resulting best fit model for the Ly α line consists in a broad single Gaussian (the resulting Ly α_n component has zero flux).

- **J1137+3549**, $z_{\text{sys}} = 6.014$: discovered by Fan et al. (2006), we model the Ly α line as a double Gaussian. Only in this case the parameters obtained for the *M-fix* model are less constrained than the ones obtained for *M-free*.
- **J0353+0104**, $z_{\text{sys}} = 6.039$: discovered by Jiang et al. (2008), this source is a BAL. Both models are consistent with a completely absorbed Ly α line.
- **J1602+4228**, $z_{\text{sys}} = 6.071$: discovered by Fan et al. (2004), this spectrum has a very low S/N. The Si II line is poorly constrained in both models.
- **J2315-0023**, $z_{\text{MgII}} = 6.117$: discovered by Jiang et al. (2008), this source has a very narrow Ly α line that we model as a double Gaussian. Since the O I+Si II line is not detected, we assume $z_{\text{sys}} = z_{\text{MgII}}$. The Si II emission is poorly constrained in the *M-free* model, but the situation improves if we fix the Ly α central wavelengths.
- **J1048+4637**, $z_{\text{CO}} = 6.2284$: this BAL QSO has been discovered by Fan et al. (2003). Since the O I+Si II is not detected, we assume $z_{\text{sys}} = z_{\text{CO}}$. The *M-free* model is consistent with no N V emission.
- **J1148+5251**, $z_{\text{sys}} = 6.373$, $z_{\text{MgII}} = 6.41$, $z_{\text{CO}} = 6.419$: discovered by Fan et al. (2003), this QSO is the most distant ever found in the SDSS survey and presents the most puzzling emission line properties. Both redshifts inferred from the Mg II and CO lines are consistent with a systemic redshift $z_{\text{sys}} \gtrsim 6.40$. From the analysis of the optical spectrum (O I+Si II line and continuum dominated windows) instead, we obtain a much lower estimate: $z_{\text{sys}} \sim 6.37$. The N V and Si II lines for the *M-free* case are also consistent with z_{OI} (see tab. 3.4). For consistency, we assume the systemic redshift to be equal to the O I+Si II redshift in the analysis that follows. For the *M-fix* case, we fail to fit the N V line (characterized by a weak peak).

3.3 COMPOSITE SPECTRUM

Once we have measured the source systemic redshifts, we can create a rest-frame flux template for the $z \sim 6$ QSOs. Since the objects in our sample were observed with two different spectrographs, they present a variety of spectral resolutions. We select all the QSOs that were observed with ESI, further restricting the sample to those sources for which spectra with the original resolution are available (i.e. before any smoothing/binning was applied). We further remove all the BALs and WLQs. Finally, we exclude from the sample J1148+5251 which presents peculiar characteristics of the BLR (see sec. 3.2.6). We obtain a sub-set of 13 spectra characterized by a spectral resolution $R \sim 3000 - 6000$, depending on the seeing and slit used. We want to compare

Figure 3.2: Fit results for J1420-1602 (top-left panel), J0927+2001 (top-right panel), J1044-0125 (bottom-left panel) and J0002+2550 (bottom right panel). We plot the observed spectrum smoothed over 10 \AA (black solid line), the best-fit models for *M-free* (red solid line), and the best fit model for *M-fix* (green solid line). The red and green horizontal lines indicate the fitting windows considered for *M-free* and *M-fix* respectively. The sky transmittance is plotted below each spectrum.

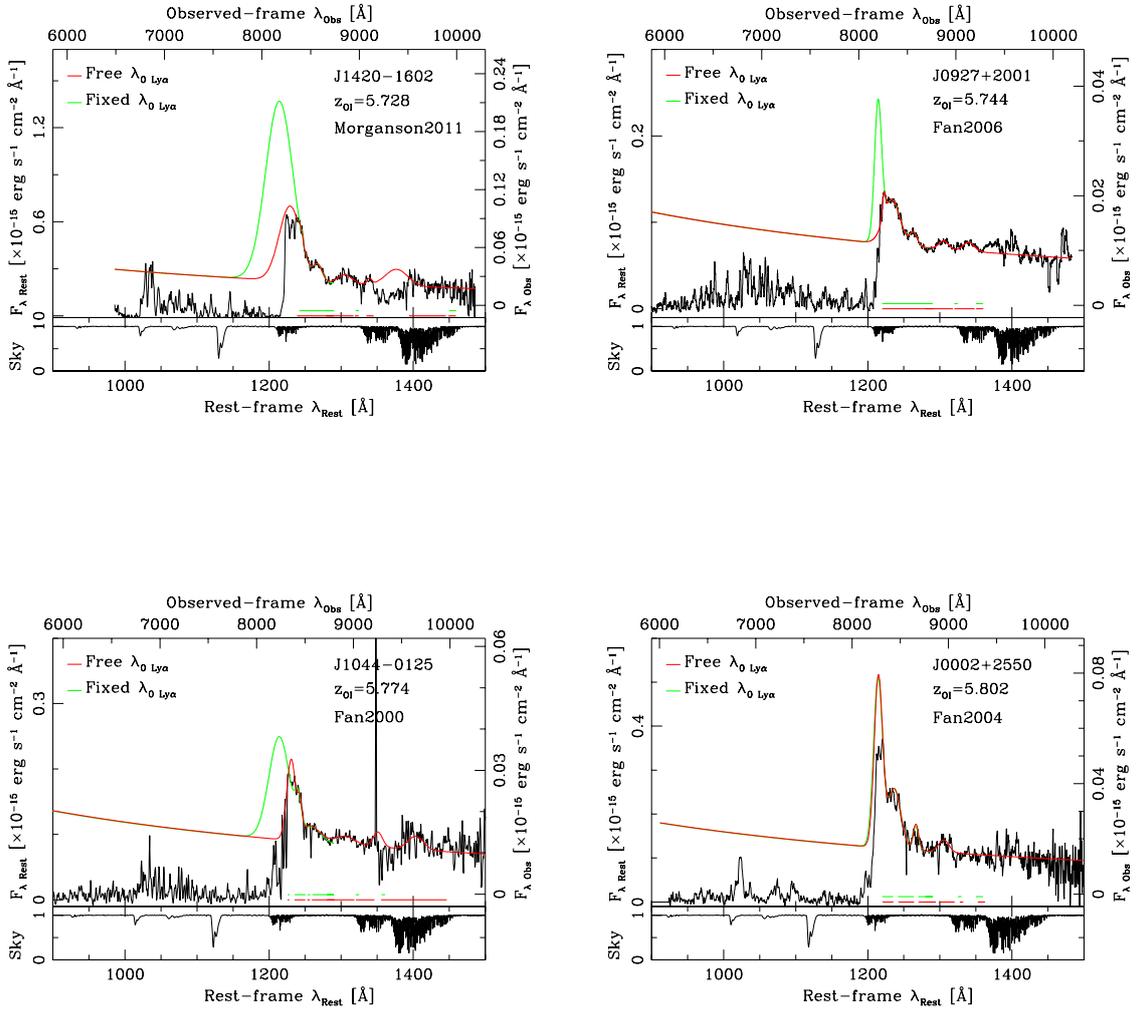


Figure 3.3: Fit results for J0836+0054 (top-left panel), J1436+5007 (top-right panel), J0840+5624 (bottom-left panel) and J0005-0006 (bottom right panel). See fig. 3.2 for a detailed description.

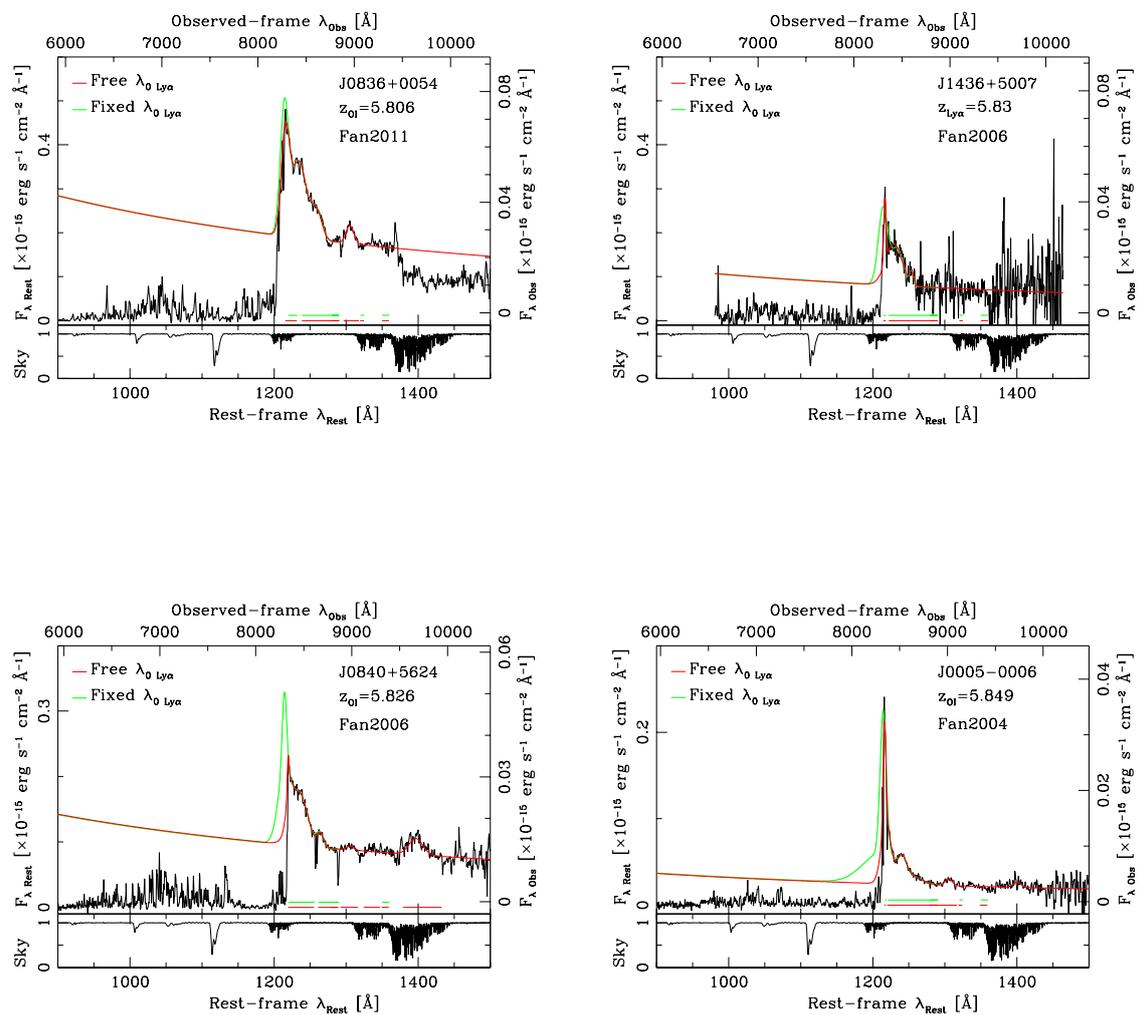


Figure 3.4: Fit results for J1411+1217 (top-left panel), J0841+2905 (top-right panel), J1306+0356 (bottom-left panel) and J1137+3549 (bottom right panel). See fig. 3.2 for a detailed description.

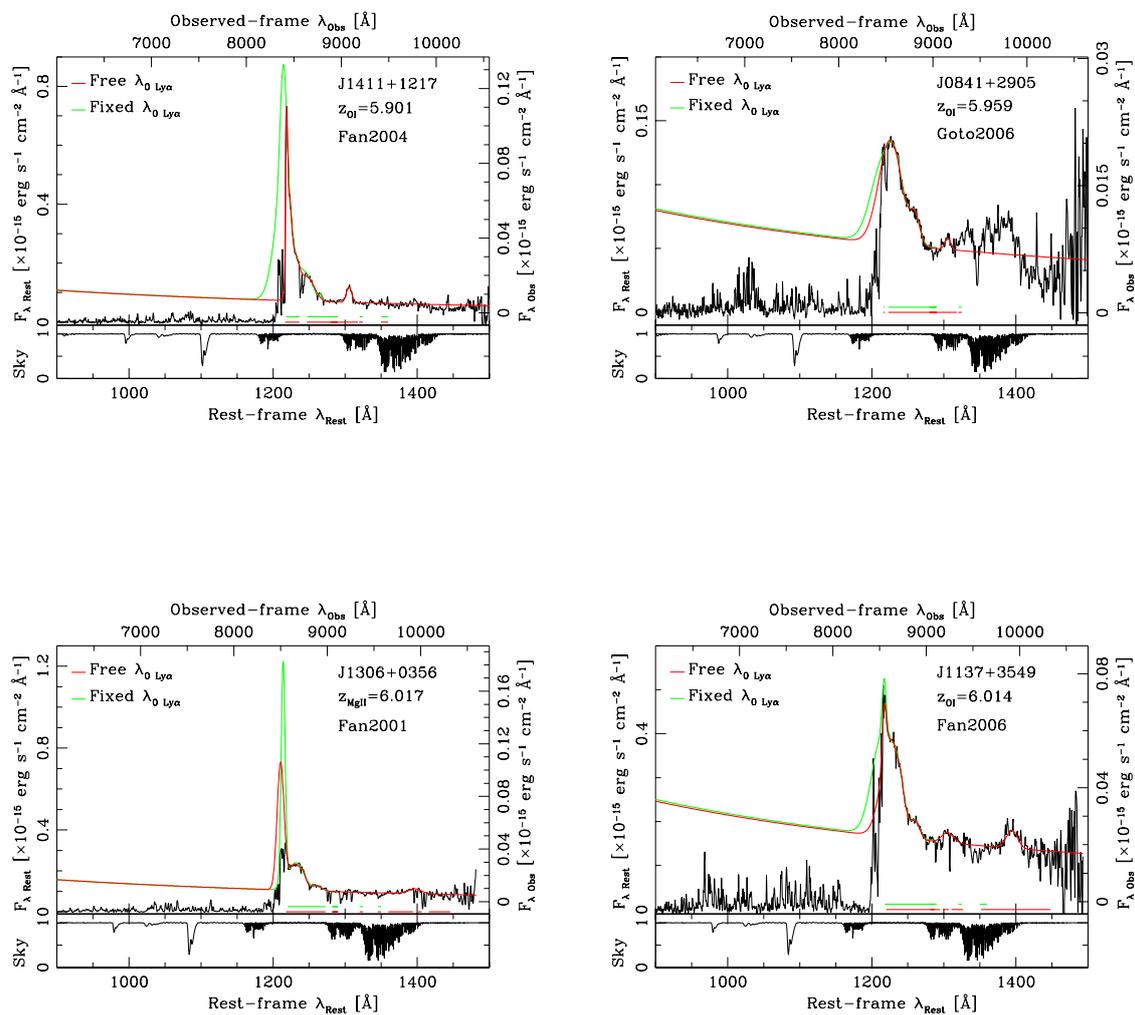


Figure 3.5: Fit results for J0353+0104 (top-left panel), J1602+4228 (top-right panel), J0842+1218 (bottom-left panel) and J0303-0019 (bottom right panel). See fig. 3.2 for a detailed description.

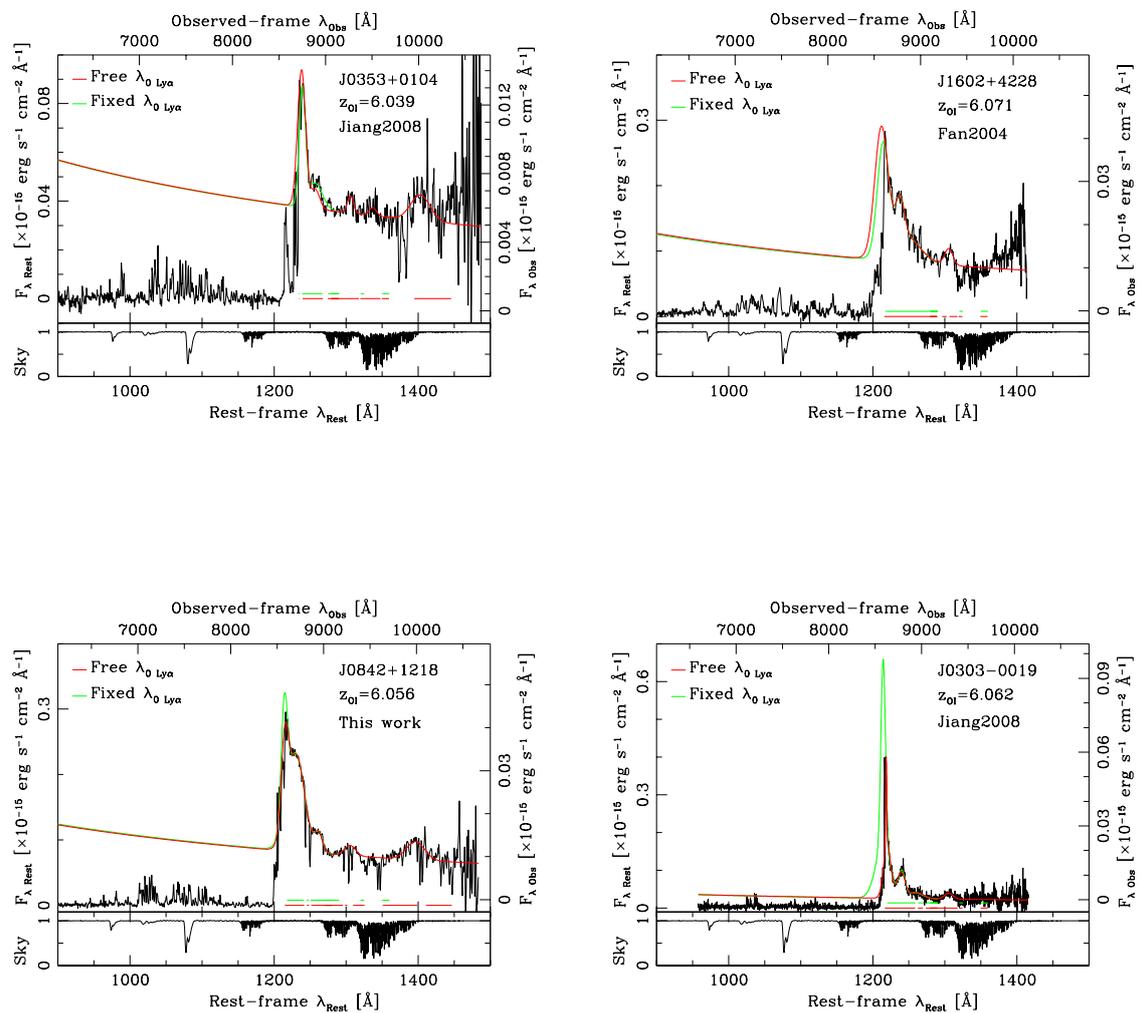


Figure 3.6: Fit results for J2315-0023 (top-left panel), J1250+3130 (top-right panel), J1048+4637 (bottom-left panel) and J1623+3112 (bottom right panel). See fig. 3.2 for a detailed description.

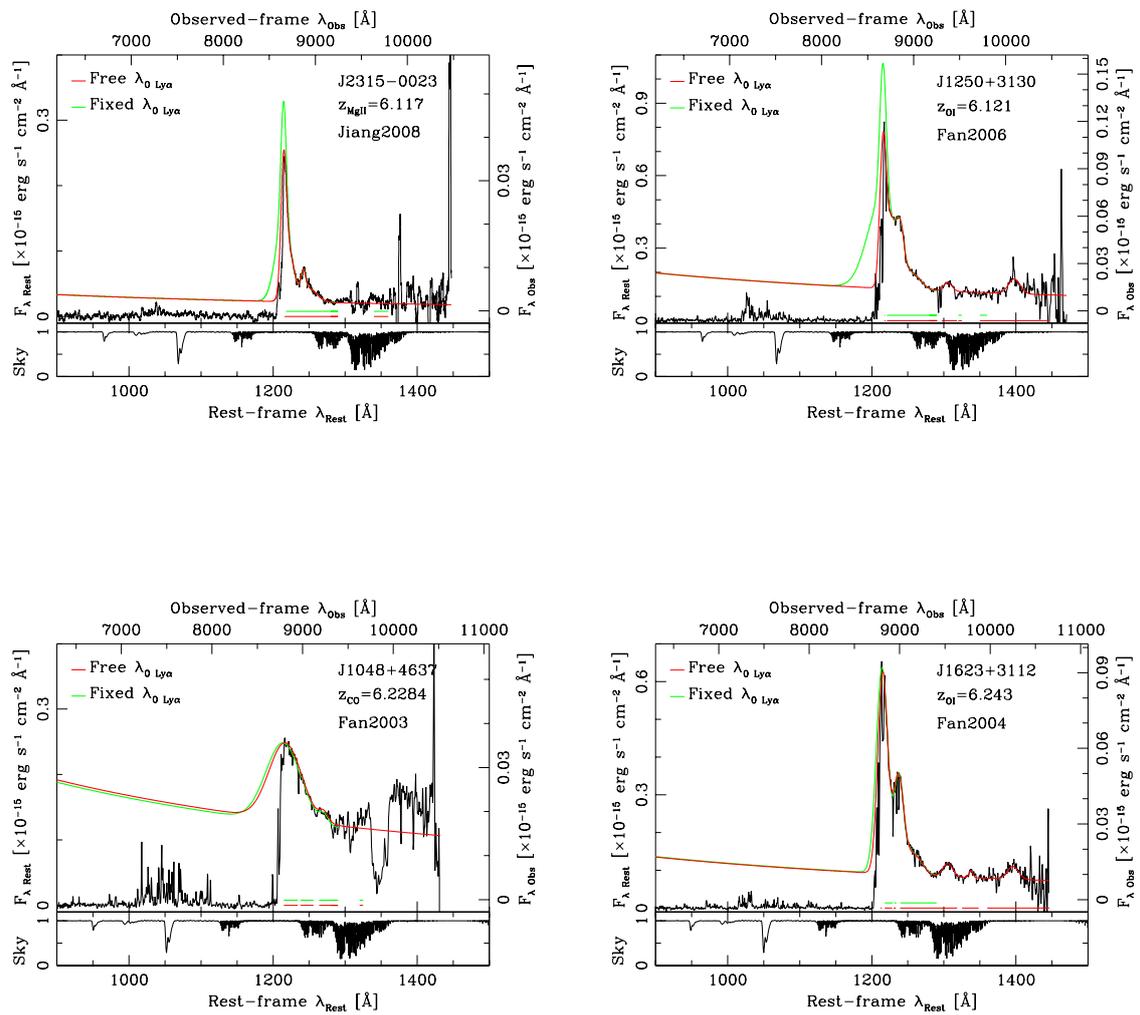


Figure 3.7: Fit results for J1030+0524 (left panel) and J1148+5251 (top right). See fig. 3.2 for a detailed description.

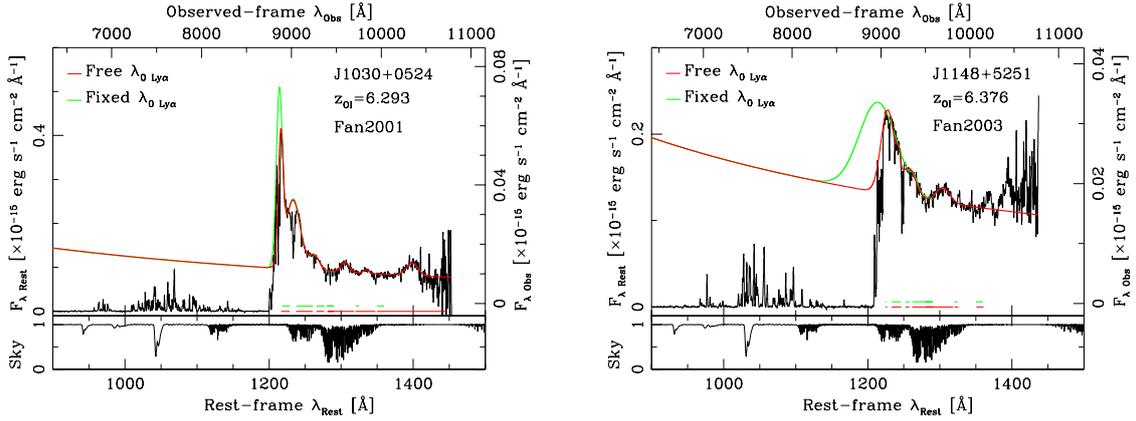
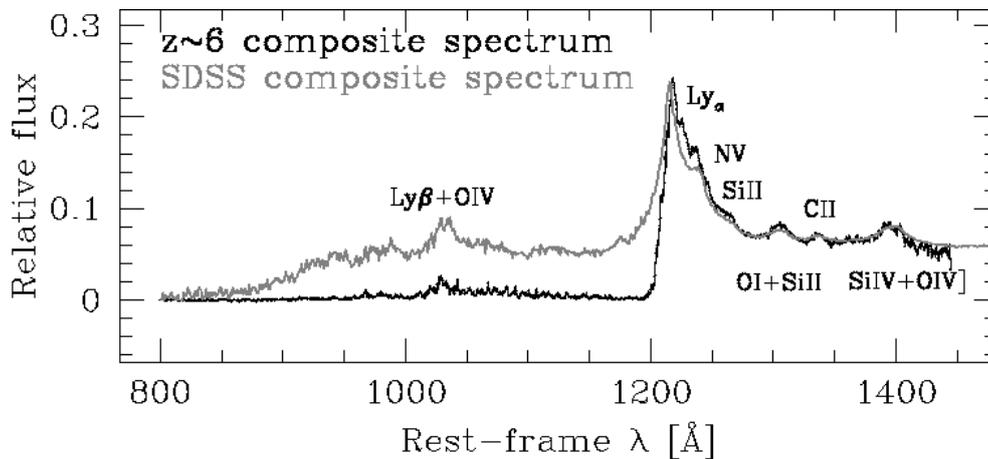


Figure 3.8: Comparison between the $z \sim 6$ QSO composite spectrum (black solid line) and the SDSS composite spectrum (grey solid line). Fluxes are in relative units and normalized to the mean flux in the continuum dominated window [1280, 1290] Å. We indicate the position of the principal emission lines. For $\lambda > 1215.67$ Å the composite spectrum obtained from lower redshift QSOs represents very well the $z \sim 6$ emitting properties.



our composite spectrum with the SDSS one obtained with QSOs at $0.1 < z_{\text{sys}} < 4.8$ (Vanden Berk et al. 2001). This lower redshift composite spectrum is characterized by a resolution of 1 \AA . After redshifting the spectra to the rest-frame, we convolve them with a Gaussian smoothing kernel with $\text{FWHM}=1 \text{ \AA}$. Assuming an average redshift $z_{\text{sys}} \sim 6$, this operation corresponds to smooth our spectra over $\sim 7 \text{ \AA}$ in the observed frame, or, equivalently, to degrade them to a resolution $R \sim 1200$ at $\lambda_{\text{obs}} \sim 8500 \text{ \AA}$. We then order our spectra for increasing redshifts and normalize the fluxes to the continuum flux of J0927+2001 (first QSO in our ordered sample) in the spectral region $\lambda_{\text{rest}} = [1280, 1290]$, which is continuum dominated. Finally, we interpolate each flux over the same wavelengths of the reference spectrum (J0927+2001). Since we are interested in the emission line properties over a relatively short wavelength range, we combine the spectra computing the median of the QSO fluxes. This procedure preserves the relative fluxes of the emission features, but does not preserve the global continuum shape (Vanden Berk et al. 2001). Fig. 3.8 shows the comparison between the SDSS median composite for lower redshift QSOs (grey line) and our $z \sim 6$ composite spectrum (black line). Fluxes have been normalized to the mean flux of the $z \sim 6$ composite in the continuum dominated window $\lambda_{\text{rest}} = [1280, 1290] \text{ \AA}$. We can identify two spectral regions: for $\lambda < 1215.67 \text{ \AA}$, the high redshift composite is significantly affected by neutral hydrogen absorption, while for $\lambda > 1215.67 \text{ \AA}$, the two composite spectra are astonishingly similar. Both central wavelength and intensity perfectly match for all the lines with the exception of the Ly α + N v complex, which is slightly brighter in the $z \sim 6$ composite. Given the small number of objects in our sample, this can be the result of a selection effect: luminous sources with strong Ly α lines have higher probabilities to be identified while looking for drop-out objects. For $\lambda > 1215.67 \text{ \AA}$ the low-redshift composite spectrum represents very well the $z \sim 6$ emitting properties of the BLR. This conclusion validates the assumptions made in sec. 3.2.3 about the direct measurement of the O I+Si II and Si IV+O IV] rest-frame wavelengths from the SDSS composite spectrum of lower redshift QSOs.

3.4 REDSHIFT ESTIMATE

As we have seen in sec. 3.2.4 a significant issue in the analysis of high redshift QSOs is to estimate the source systemic redshift. From studies of low redshift QSOs (e.g Richards et al. 2002, 2011; Shang et al. 2007) we know that high-ionization lines (e.g C IV) can present high velocity offsets with respect to low-ionization lines (e.g. Mg II). Assuming that this is valid also for high-redshift QSOs, one can estimate the systemic redshift from the Mg II emission line (ionization potential of 15.03 eV). This line is thought to reproduce the host galaxy redshift within an error $\Delta z \leq 0.007$. At $z \sim 6$ though, Mg II ($\lambda_{\text{rest}} = 2798.95 \text{ \AA}$) falls in the near-infrared K-band, which is severely affected by the sky telluric absorption. This might complicate the line fit (see chap.2, Jiang et al. 2007; Kurk et al. 2007). A possible way to measure directly the QSO

systemic redshift is through the detection of molecular gas in the host galaxy. CO emission has been detected in 8 SDSS $z \sim 6$ QSOs (Wang et al. 2010). High resolution imaging shows that for one of these sources the CO is centered within $0.2''$ from the QSO (Walter et al. 2004; Riechers et al. 2009). If we assume that this is valid for the entire population of $z \sim 6$ QSOs, then CO redshift measurements will reproduce the systemic redshifts within an error $\Delta z \leq 0.002$. In table 3.3 we list the measurements of Mg II, C IV and CO redshifts available for the sources in our sample. Unfortunately, we have a very small overlap between different estimates: Mg II and C IV measurements for 4 of the sources, while only 2 QSOs have both Mg II and CO redshifts. If we compute the mean difference between the values, we have that $\Delta(z_{\text{MgII}} - z_{\text{CIV}}) = 0.007$ with a standard deviation of 0.001, while $\Delta(z_{\text{MgII}} - z_{\text{CO}}) = -0.015$ with a standard deviation of 0.002. Given the low number of sources in the subsamples, we can easily conclude that these offsets do not represent the general trend of the $z \sim 6$ QSO population.

In tab. 3.4 we list the redshifts inferred from the fitted optical lines with the respective errors. For N V and Si II we have two estimates corresponding to *M-free* and *M-fix* respectively (their central wavelengths are free parameters in both models). The redshifts have been calculated as:

$$z_{\text{line}} + 1 = \frac{\lambda_{0,\text{line}}}{\mu_{\text{line}}}, \quad (3.10)$$

where $\lambda_{0,\text{line}}$ is the central wavelength measured in the observed frame (see tab. 3.6-3.10) and μ_{line} is the reference wavelength for the line transition (see 3.2.3). We have run tests during the fitting analysis and noticed that changes in either the fitting windows or the power-law slope can lead to differences $\Delta z \sim 0.01$ in the O I+Si II and Si IV+O IV] redshift estimates. Therefore, we assume $\sigma \sim 0.01$ as the redshift systematic error due to the fitting procedure.

COMPARISON BETWEEN M-FREE AND M-FIX RESULTS

In the *M-free* case the best fit model for the Ly α line is on average red-shifted with respect to other lines. Moreover, when we model the Ly α line as a double Gaussian, the Ly α_b component is always redshifted with respect to the narrow component, in disagreement with what has been found for lower redshift sources (Kramer & Haimann 2009). The N V estimates resulting from *M-free* and *M-fix* are typically in good agreement, while for the Si II line we can notice some differences. When we are modeling the Ly α line as a double Gaussian, the redshift inferred from the Si II line has smaller errors if we fix the Ly α central wavelengths.

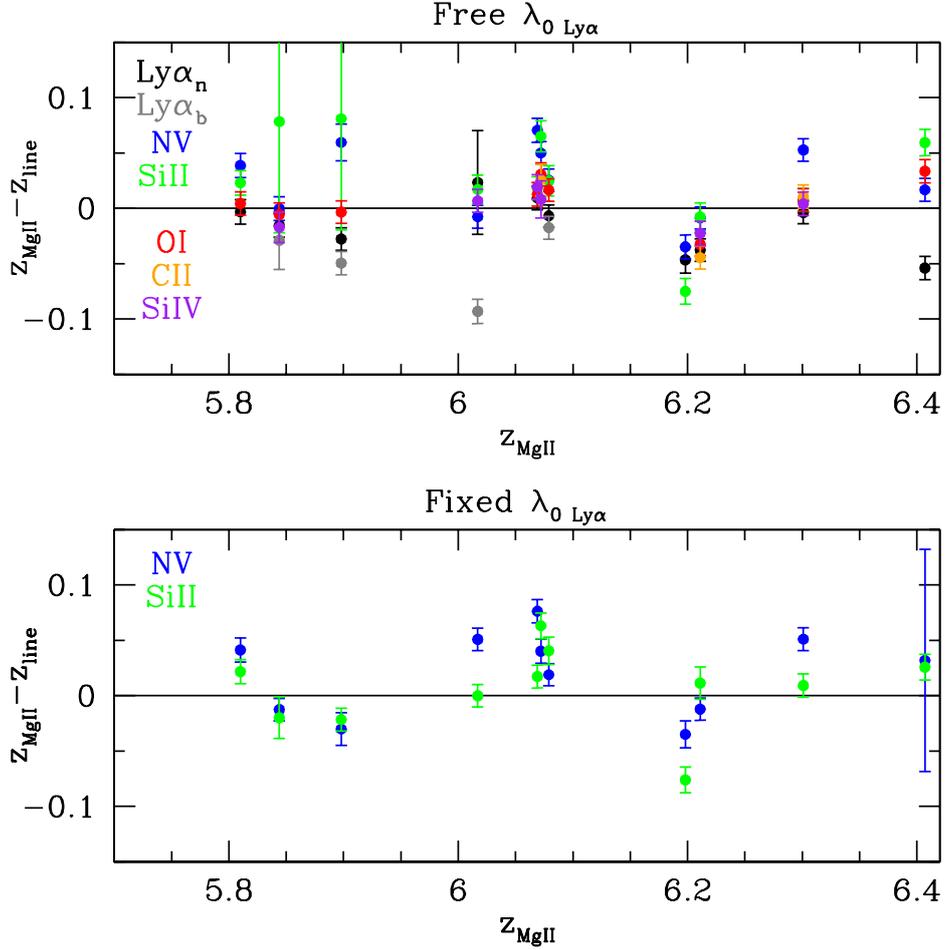
The majority of these variations can be explained through the analysis of the fitting procedure. As we have already noticed, if the Ly α line is significantly absorbed, in the *M-free* case the red-wing of the line is erroneously fitted as a peak (see fig. 3.2-3.7). The resulting model is then centered at higher wavelengths with respect to the *real* emission line. This effect is severe for 8

Table 3.3: Literature redshifts measured from C IV, Mg II and CO emission lines.

QSO name ¹	z_{CIV} ²	z_{MgII} ³	z_{CO} ⁴
J1420-1602	–	–	–
J0927+2001	–	–	5.7722 ± 0.0006^d
J1044-0125	–	–	5.7824 ± 0.0007^d
J0002+2550	–	–	–
J0836+0054	–	5.810 ± 0.003^b	–
J1436+5007	–	–	–
J0840+5624	5.774 ± 0.002^a	–	5.8441 ± 0.0013^d
J0005-0006	5.848 ± 0.001^b	5.844 ± 0.001^c	–
J1411+1217	5.911 ± 0.001^b	5.898 ± 0.002^c	–
J0841+2905	–	–	–
J1306+0356	5.997 ± 0.001^b	6.017 ± 0.001^c	–
J1137+3549	5.962 ± 0.002^a	–	–
J0353+0104	–	6.072 ± 0.002^c	–
J1602+4228	6.051 ± 0.002^a	–	–
J0842+1218	–	6.069 ± 0.002^c	–
J0303-0019	–	6.079 ± 0.001^c	–
J2315-0023	–	–	–
J1250+3130	–	–	–
J1048+4637	–	6.198 ± 0.004^c	6.2284 ± 0.0017^d
J1623+3112	–	6.211 ± 0.001^c	–
J1030+0524	6.262 ± 0.003^b	6.301 ± 0.002^c	–
J1148+5251	–	6.407 ± 0.002^c	6.4189 ± 0.0006^e

¹ QSO name;² C IV, literature: a, Kurk et al. (2007); b, Ryan-Weber et al. (2009);³ Mg II, literature: a, Kurk et al. (2007); c, De Rosa et al. (2011);⁴ CO, literature: d, Wang et al. (2010); e, Bertoldi et al. (2003);

Figure 3.9: Δz_{MgII} as a function of z_{MgII} for *M-free* (top panel) and *M-fix* models (bottom panel). Black points: $\text{Ly}\alpha_n$; grey points: $\text{Ly}\alpha_b$; blue points: N v; green points: Si II; red points: O I+Si II; orange points: C II; purple points: Si IV+O IV]. The solid horizontal lines indicate the level of zero offset.



of the sources in our sample. The differences detected in the Si II redshift estimates simply derive from the fact that the $\text{Ly}\alpha$ broad component can overlap with the Si II emission feature, resulting in a weaker constraint on the Si II central wavelength. This overlap is larger in the *M-free* model, since in this case the resulting $\text{Ly}\alpha_b$ is centered at longer wavelengths compared to the *M-fix* case. Finally, the detected changes in $(\lambda_{0,\text{Ly}\alpha_n} - \lambda_{0,\text{Ly}\alpha_b})$ between the high and low redshift quasars might be due to the fact that we are fitting only the red wing of the $\text{Ly}\alpha$ complex while at low redshifts the $\text{Ly}\alpha$ line is not affected by the neutral hydrogen absorption. Simply, the information contained in our narrow wavelength range might not be enough to properly constrain the 12 Gaussian parameters that we are fitting simultaneously (3 for each emission line). The situation improves if we use the priors on the central wavelengths of the $\text{Ly}\alpha$ components.

COMPARISON WITH LITERATURE ESTIMATES

We measure the differences between the literature Mg II, C IV and CO redshifts and the optical ones:

$$\Delta z_{\text{MgII,CIV,CO}} = z_{\text{MgII,CIV,CO}} - z_{\text{line}}. \quad (3.11)$$

Errors on Δz are computed using standard error propagation and adding our estimate of the systematic error. In fig. 3.9 we show Δz_{MgII} as a function of z_{MgII} , for *M-free* (top panel) and *M-fix* (bottom panel). In the bottom panel we plot only the redshifts estimated from N v and Si II since we are fixing the Ly α central wavelength using z_{OI} . Similar plots have been obtained for Δz_{CIV} and Δz_{CO} . There is no detectable dependence of Δz on the QSO redshift. In order to better quantify the differences in the redshift estimates, we have computed the weighted mean for each Δz (see tab. 3.5). Since the redshift inferred from the Ly α line is strongly dependent on the adopted model, we exclude this line from the comparison. For the N v and the Si II lines we consider the values obtained using the *M-fix* model (see discussion above).

The C IV line is on average blue-shifted with respect to the optical lines, with Δz_{CIV} going from a minimum of -0.041 for the C II line to a maximum of -0.023 for the O I+Si II line. By looking at the differences with z_{MgII} instead, which is supposed to be a good estimator of z_{sys} (see previous discussion), we notice a correlation between the absolute value of the offset and the line ionization potential. In particular, $|\Delta z_{\text{MgII}} = 0.023|$ for the N v line (highest ionization potential, 97.89 eV), while $|\Delta z_{\text{MgII}} = 0.007|$ for the O I+Si II complex (average ionization potential 14.98 eV). If we exclude from the sub-sample J1148+5251, for which we have noticed a significant inconsistency between the optical and the Mg II and CO redshifts (see sec. 3.2.6), the mean Δz_{MgII} drops to 0.004 ± 0.003 for the O I+Si II line while the blue-shift remains unchanged for the N v. Assuming an average redshift $z_{\text{sys}} \sim 6$, these mean offsets correspond to velocity shifts of the order of $\sim 200 \text{ km s}^{-1}$ for the O I+Si II line and $\sim 1000 \text{ km s}^{-1}$ for the N v. The exception to this observed trend is seen in the Si IV+O IV] line. Despite being a high ionization line (average ionization potential equal to 61.28 eV), the redshift inferred from this line is in very good agreement with the Mg II estimate. A peculiar behavior of this emission line has already been noticed in the literature (e.g. Richards et al. 2002). Unfortunately we are not able to compare the redshifts inferred from the optical lines with the ones estimated from the CO lines, since very few z_{CO} measurements are available.

The results for the N v and Si IV+O IV] line shifts can be interpreted through the *disk-wind* model for the broad emission line region (e.g. Murray et al. 1995; Leighly et al. 2004; Richards et al. 2011). In this model the BLR is a combination of a disk and a wind component, the relative importance of which depends on the actual QSO spectral energy distribution. The lines showing blue-shifts are originated in the wind. In particular, the N v has a strong wind component (e.g.

Leighly et al. 2004; Richards et al. 2011), as the C iv line. Therefore, large blue-shifts are expected for the N v line. The Si iv+O iv] complex is instead composed by the Si iv line, which is mainly originated in the disk, and the O iv] line, mainly originated in the wind. Since the Si iv line is dominating the total emission flux, no blue-shift is expected for this complex (see also Richards et al. 2011).

Table 3.4: Source redshifts measured from the optical emission lines.

QSO name ¹	$z_{Ly\alpha}$ ²	$z_{Ly\alpha,b}$ ³	z_{NV} ⁴	z_{NV} ⁵	z_{SiII} ⁶	z_{SiII} ⁷	$z_{OI+SiII}$ ⁸	z_{CI} ⁹	$z_{SiIV+OIV}$ ¹⁰
			Free $\lambda_{0, Ly\alpha}$	Fixed $\lambda_{0, Ly\alpha}$	Free $\lambda_{0, Ly\alpha}$	Fixed $\lambda_{0, Ly\alpha}$			
J1420-1602	5.81±0.13	—	5.746±0.006	5.758±0.002	5.757±0.006	5.756±0.002	5.728±0.006	5.748±0.003	5.6±0.1
J0927+2001	5.789±0.001	—	5.703±0.004	5.705±0.005	5.749±0.002	5.748±0.002	5.744±0.003	5.769±0.004	—
J1044-0125	5.865±0.001	—	5.784±0.002	5.772±0.003	5.749±0.01	5.76±0.02	5.774±0.007	5.853±0.1	5.802±0.005
J0002+2550	5.80±0.05	—	5.779±0.006	5.780±0.005	5.825±0.003	5.823±0.003	5.802±0.006	—	—
J0836+0054	5.813±0.003	—	5.771±0.003	5.769±0.003	5.787±0.004	5.788±0.003	5.806±0.001	—	—
J1436+5007	5.844±0.001	—	5.765±0.008	5.79±0.01	5.784±0.004	5.783±0.004	—	—	—
J0840+5624	5.855±0.005	5.87±0.03	5.79±0.01	5.80±0.03	5.835±0.004	5.831±0.006	5.826±0.005	—	5.822±0.004
J0005-0006	5.859±0.004	5.87±0.02	5.844±0.003	5.857±0.002	5.8±0.1	5.86±0.02	5.849±0.003	—	5.861±0.009
J1411+1217	5.926±0.001	5.947±0.003	5.84±0.01	5.93±0.01	5.8±0.1	5.920±0.001	5.901±0.001	—	—
J0841+2905	6.025±0.003	—	5.9±0.1	5.92±0.03	5.946±0.002	5.940±0.007	5.960±0.003	—	—
J1306+0356	5.99±0.05	6.110±0.004	6.025±0.002	5.966±0.001	6.000±0.008	6.017±0.002	—	—	6.010±0.003
J1137+3549	6.03±0.02	6.087±0.007	6.012±0.003	6.0±0.2	6.008±0.003	6.00±0.01	6.014±0.004	—	5.998±0.003
J0353+0104	—	—	6.022±0.001	6.032±0.003	6.01±0.01	6.009±0.006	6.041±0.003	6.05±0.01	6.06±0.01
J1602+4228	6.06±0.03	—	6.053±0.004	6.052±0.003	5.9±0.1	5.9±0.1	6.071±0.002	—	—
J0842+1218	6.059±0.003	—	5.998±0.003	5.993±0.002	6.051±0.002	6.051±0.002	6.056±0.003	—	6.050±0.004
J0303-0019	6.086±0.001	6.096±0.002	6.054±0.001	6.060±0.001	6.054±0.009	6.038±0.007	6.063±0.002	—	—
J2315-0023	6.118±0.006	6.16±0.03	6.125±0.001	6.126±0.001	6.0±0.1	6.04±0.02	—	—	—
J1250+3130	6.128±0.008	6.20±0.02	6.12±0.01	6.114±0.001	6.06±0.01	6.11±0.02	6.121±0.003	—	6.121±0.003
J1048+4637	6.245±0.005	—	6.2±0.1	6.233±0.006	6.273±0.004	6.274±0.004	—	—	—
J1623+3112	6.249±0.007	—	6.220±0.001	6.223±0.001	6.219±0.008	6.20±0.01	6.243±0.002	6.255±0.002	6.234±0.005
J1030+0524	6.305±0.001	—	6.248±0.001	6.250±0.001	6.293±0.002	6.292±0.002	6.293±0.002	6.290±0.002	6.297±0.003
J1148+5251	6.461±0.002	—	6.390±0.002	6.400±0.100	6.347±0.006	6.381±0.005	6.373±0.002	—	—

¹ QSO name;

² Ly α_n ;

³ Ly α_b ;

^{4,5} NV model A and model B respectively;

^{6,7} Si II for model A and model B respectively;

⁸ O I+Si II

⁹ C II;

¹⁰ Si IV+O IV]

Table 3.5: Weighted mean difference between the C IV, Mg II and CO redshifts and the the optical lines redshifts.

QSO name ¹	C IV, N ²		z _{CIV} - z _{line} ³		Mg II, N ⁵		z _{MgII} - z _{line} ⁶		CO, N ⁸		z _{CO} - z _{line} ⁹		z _{CO} - z _{line} ¹⁰	
	Free	λ _{0, Lyα}	Fixed	λ _{0, Lyα}	Free	λ _{0, Lyα}	Fixed	λ _{0, Lyα}	Free	λ _{0, Lyα}	Fixed	λ _{0, Lyα}	Free	λ _{0, Lyα}
Lyα _n	8	-0.039±0.005	-	-	10	-0.020±0.003	-	-	5	-	-	-	5	-0.035±0.005
Lyα _b	5	-0.084±0.006	-	-	4	-0.051±0.006	-	-	-	-	-	-	-	-
N V	8	-0.033±0.004	-0.023±0.004	-	11	0.022±0.003	0.023±0.003	-	5	0.025±0.005	0.026±0.006	-	5	0.025±0.005
Si II	8	-0.037±0.005	-0.041±0.004	-	11	0.012±0.004	0.007±0.003	-	5	0.017±0.005	0.008±0.005	-	5	0.017±0.005
O I+Si II	6	-0.023±0.004	-	-	9	0.007±0.003	-	-	4	0.027±0.005	-	-	4	0.027±0.005
C II	-	-	-	-	3	-0.009±0.007	-	-	2	0.01±0.01	-	-	2	0.01±0.01
Si IV+O IV]	5	-0.030±0.005	-	-	6	0.001±0.005	-	-	2	0.002±0.008	-	-	2	0.002±0.008

¹ Optical line;^{2,3,4} C IV, number of sources, weighted mean for model A, weighted mean for model B;^{5,6,7} Mg II, number of sources, weighted mean for model A, weighted mean for model B;^{8,9,10} CO, number of sources, weighted mean for model A, weighted mean for model B

3.5 EMISSION PROPERTIES

Through the study of trends and correlations amongst QSO emission features (e.g. line luminosity, EW and FWHM; continuum luminosity) we can try to put constraints on the nature and on the dynamical properties of the central engines that are powering these sources. We have computed the following quantities from the fitted parameters:

- Flux density for the non-stellar continuum at $\lambda_{rest} = 1350 \text{ \AA}$, obtained using the power-law parametrization (see eq. 3.1) and the normalization measured for each QSO (see appendix tab. 3.6).
- Line fluxes, computed analytically as $F = \sqrt{2\pi} A \sigma$, where A and σ are the amplitude and the width of the line, which have been measured from the continuum subtracted spectrum (see the appendix tables 3.6-3.10).
- Line FWHM in km s^{-1} , computed as $\text{FWHM} = c \frac{2.35 \sigma}{\lambda_0}$, where c is the speed of light in km s^{-1} , σ is the line width and λ_0 is its measured line central wavelength (see the appendix tables 3.6-3.10).
- Line EW, estimated as $\text{EW} = \frac{F}{f_{\text{cont}}}$, where F is the line integral flux and f_{cont} is the continuum flux density measured at the line central wavelength.

Errors are computed through standard error propagation and do not include systematics. Using our estimate of the systemic redshift (see sec.3.2.4), we compute the luminosity distance (D_L) for each QSO considering a standard Λ CDM model (Spergel et al. 2007). We then measure the luminosity of each component as:

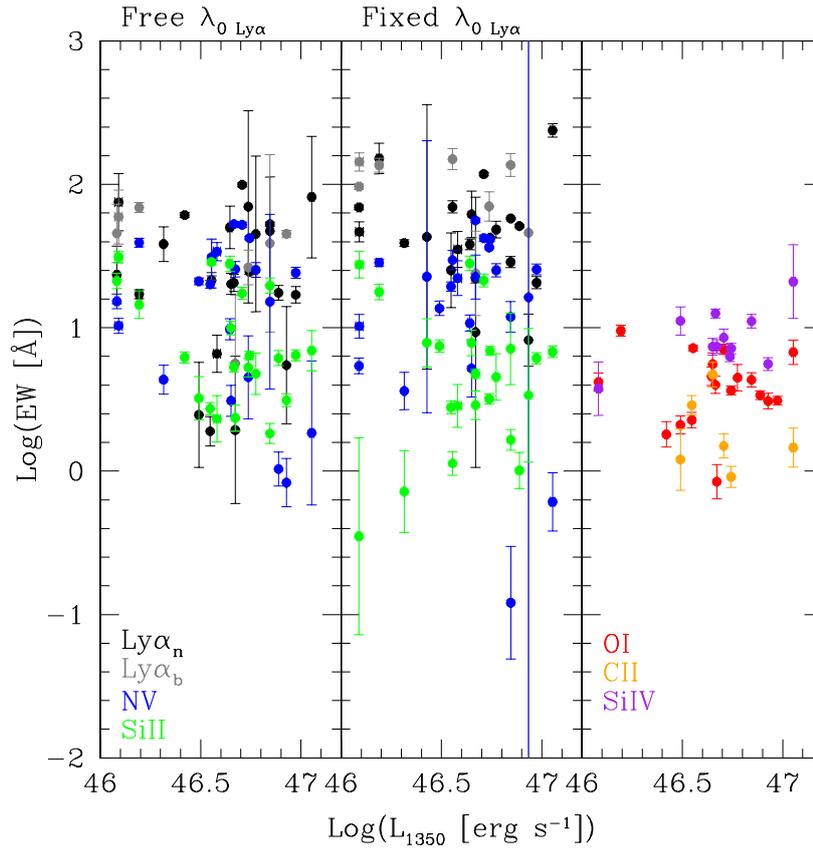
$$L = 4\pi F D_L^2 \quad (3.12)$$

where F is the integral flux (for the power-law continuum $F = \lambda f(\lambda)$).

3.5.1 CONTINUUM LUMINOSITY

For high ionization lines an anti-correlation between the EW and the rest-frame UV luminosity has been observed. This relation, also known as *Baldwin effect*, has been found for the first time by Baldwin (1977) in the study of the C IV emission for a sample of 20 (mostly) radio-loud nearby QSOs. The observed degree of anti-correlation is a function of the line ionization potential (the higher is the line ionization potential the stronger is the anti-correlation). In fig. 3.10 we show

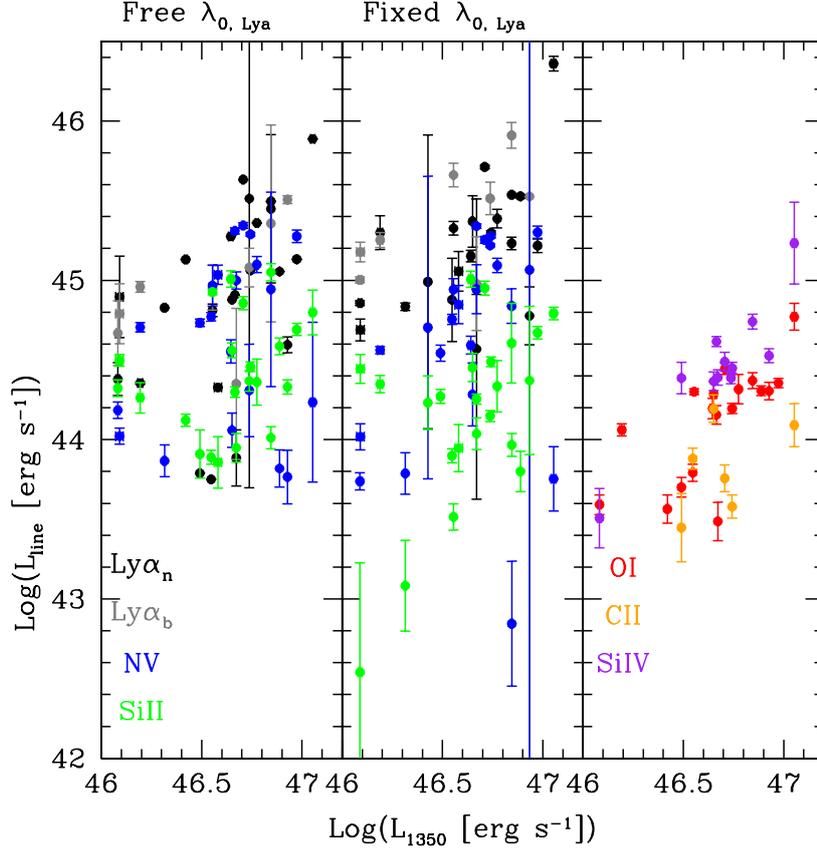
Figure 3.10: EW as a function of $L(1350 \text{ \AA})$: $\text{Ly}\alpha$ complex M -free model, left panel ($\text{Ly}\alpha_n$, black; $\text{Ly}\alpha_b$, grey; N v , blue; Si II green); $\text{Ly}\alpha$ complex M -fix model, central panel; $\text{O I}+\text{Si II}$ (red), C II (orange) and $\text{Si IV}+\text{O IV}$ (purple) lines, right panel. There is no evidence of correlation between these two quantities. While for the low ionization lines the absence of correlation is in agreement with the *Baldwin effect*, we would expect an anti-correlation for the high ionization lines ($\text{Si IV}+\text{O IV}$] and N v).



the line EW as a function of the UV continuum luminosity at $\lambda_{rest} = 1350 \text{ \AA}$ ($\text{Ly}\alpha$ complex M -free model, left panel; $\text{Ly}\alpha$ complex M -fix model, central panel; $\text{O I}+\text{Si II}$, C II and $\text{Si IV}+\text{O IV}$] lines, right panel). There is no evidence of a correlation, as confirmed by the Pearson correlation coefficients. For the low ionization lines the absence of correlation is in agreement with what is expected from the *Baldwin effect*. An anti-correlation for the two lines with higher ionization potential ($\text{Si IV}+\text{O IV}$] and N v) would instead be expected. A lack of anti-correlation for the N v line has also been noted in previous studies (Dietrich et al. 2002).

In addition to the correlation between EW and continuum luminosity we also explore the relation between the line luminosity and continuum luminosity (fig. 3.11: $\text{Ly}\alpha$ complex M -free model, left panel; $\text{Ly}\alpha$ complex M -fix model, central panel; $\text{O I}+\text{Si II}$, C II and $\text{Si IV}+\text{O IV}$] lines, right panel). In this case, a correlation between the two quantities is evident. For the $\text{Ly}\alpha$ complex

Figure 3.11: Line luminosity as a function of $L(1350 \text{ \AA})$: $\text{Ly}\alpha$ complex M -free model, left panel ($\text{Ly}\alpha_n$, black; $\text{Ly}\alpha_b$, grey; N v , blue; Si II green); $\text{Ly}\alpha$ complex M -fix model, central panel; $\text{O I}+\text{Si II}$ (red), C II (orange) and $\text{Si IV}+\text{O IV}$ (purple) lines, right panel. There is a strong correlation between these two quantities for all the emission lines with the exception of N v (high ionization potential).

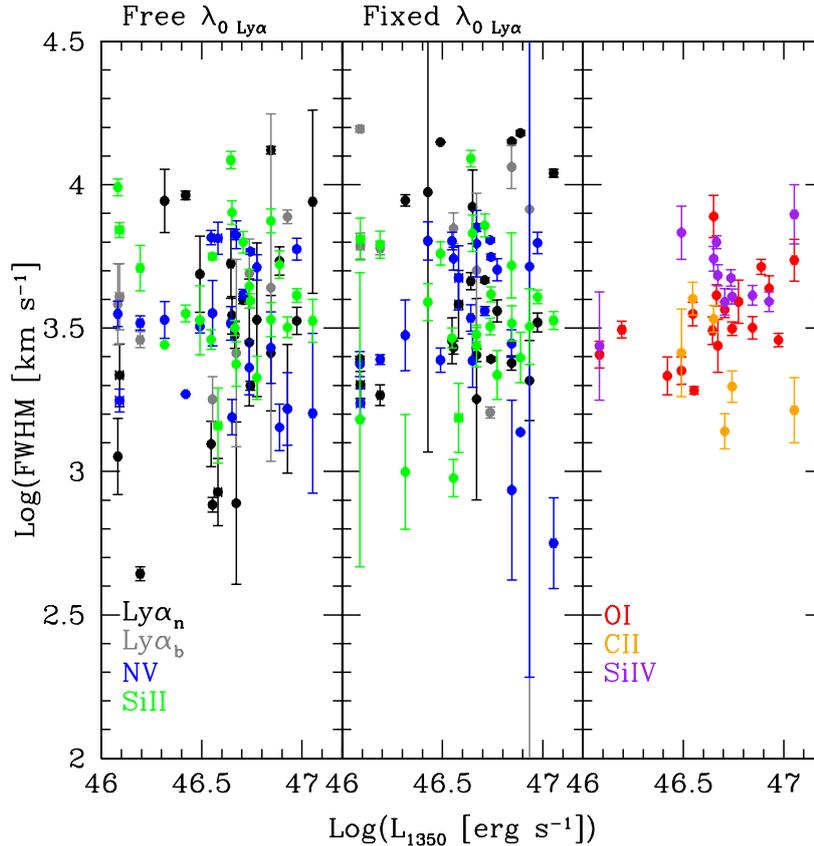


the correlation is stronger for the M -fix case (fixed $\text{Ly}\alpha$ central wavelengths). By measuring the Pearson coefficients, we obtain that the only line for which the hypothesis of correlation can be rejected is the N v line. Since this relation is simply a re-formulation of the EW-continuum luminosity one, the lack of correlation between the line and continuum luminosities might correspond to a weak detection of the anti-correlation between the N v EW and the continuum luminosity. We should keep in mind though, that the N v line can be severely affected by the delicate spectral decomposition of the $\text{Ly}\alpha$ complex. For the $\text{Si IV}+\text{O IV}$ we do not detect the predicted trend.

Finally, we have investigated a possible relation between the line FWHM and the continuum luminosity (see fig. 3.12). We find no significant correlation between the two quantities. This might be due to the small range of $L(1350 \text{ \AA})$ sampled in this study.

The anti-correlation between the line EW and the UV continuum can be explained through the *disk-wind* model. A given BLR is characterized by the presence of a radiation-line-driven wind

Figure 3.12: Line FWHM as a function of $L(1350 \text{ \AA})$: $\text{Ly}\alpha$ complex M -free model, left panel ($\text{Ly}\alpha_n$, black; $\text{Ly}\alpha_b$, grey; N v , blue; Si II green); $\text{Ly}\alpha$ complex M -fix model, central panel; $\text{O I}+\text{Si II}$ (red), C II (orange) and $\text{Si IV}+\text{O IV}$ (purple) lines, right panel. No significant correlation is found. This might be due to the narrow range of $L(1350 \text{ \AA})$ sampled in this study.



if the X-ray ionizing flux (soft X-ray) is not too strong. In this model, high ionization lines can be originated in winds, while intermediate/low ionization lines are always generated in the accretion disk or at the base of the wind. At the same time, we know that QSOs with brighter UV continua are characterized by relatively lower soft X-ray fluxes (e.g. Green et al. 1995; Steffen et al. 2006; Just et al. 2007), therefore they can develop winds in their BLR. Summarizing, if we have a QSO with a bright UV continuum (ionizing fluxes not high enough to generate high ionization lines in the disk), radiation line driven winds can develop in its BLR, hence the high ionization lines will be dominated by their wind component and therefore they will have smaller EWs. As we have seen, the $\text{Si IV}+\text{O IV}$ complex is dominated by the Si IV line, which is mainly originated in the disk. Therefore, for this complex no anti-correlation is expected. However, for the N v line, which is mainly originated in the wind, we expect an anti-correlation with the UV luminosity.

3.5.2 BLACK HOLE MASS

In fig. 3.13 we plot the line FWHM as a function of the line luminosity ($\text{Ly}\alpha$ complex *M-free* model, left panel; $\text{Ly}\alpha$ complex *M-fix* model, central panel; $\text{O I}+\text{Si II}$, C II and $\text{Si IV}+\text{O IV}$] lines, right panel). A strong correlation between the two quantities is evident: brighter lines are broader than weaker lines. The Pearson correlation coefficient is large ($\rho > 0.5$) for all the considered lines except for the C II line ($\rho_{\text{CII}} = 0.18$) and for the $\text{Ly}\alpha_b$ component resulting from the *M-fix* model ($\rho_{\text{Ly}\alpha_b} = -0.02$). While for the other components of the $\text{Ly}\alpha$ complex the correlation strength does not depend significantly on the adopted model, for the $\text{Ly}\alpha_b$ line the degree of correlation changes dramatically ($\rho_{M\text{-free}} = 0.60$, $\rho_{M\text{-fix}} = -0.02$). The strong variations in the $\text{Ly}\alpha_b$ line luminosity and FWHM might arise from the fact that the peak of this line assumes two different positions as a function of the model: while in *M-free* the line is systematically redshifted with respect to the $\text{Ly}\alpha_n$ one, in the *M-fix* model, following the results for low redshift QSOs, we impose a fixed blue-shift with respect to the narrow component (see sec. 3.4).

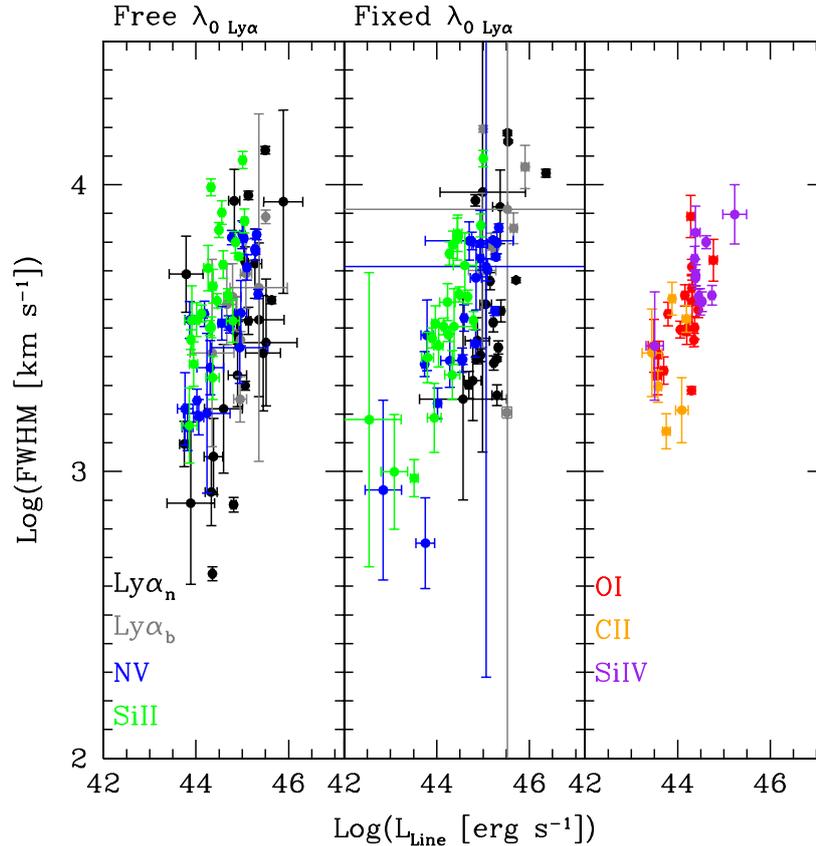
The observed correlation might simply be the result of the intrinsic relation holding between the line luminosity and the line width:

$$L \propto \sigma A \rightarrow L \propto \text{FWHM } \lambda_0 A, \quad (3.13)$$

but it could be witnessing the fact that we are observing QSOs hosting BHs with different masses. If the dynamics of the gas emitting the optical lines is dominated by the gravity of the central BH, we will expect larger FWHM for QSOs hosting more massive BHs but with similar Eddington ratios (Labita et al. 2009). Since the distribution of Eddington ratios for $z \sim 6$ QSOs is relatively narrow (see chap. 2), the correlation between the line luminosity and FWHM might indicate that QSOs with brighter lines are the ones hosting more massive BHs.

To test this statement, we need additional information on M_{BH} . As we have seen in chap. 2, it is possible to estimate M_{BH} using scaling relations, calibrated on local AGNs, that are based on broad emission line widths and continuum luminosities. For $z \sim 6$ QSOs, we can only use the calibrated relations available for Mg II and C IV lines (with few reserves on the latter). In sec. 2.4.1 we have estimated M_{BH} from the Mg II emission line for a set of SDSS QSOs at $z \sim 6$, that partly overlaps with the optical sample in analysis (total of 9 sources out of 22). Therefore, we can use our homogeneous estimates of M_{BH} to test if brighter lines are related to more massive BHs. In fig. 3.14 we plot the line luminosity as a function of M_{BH} (estimated through eq. 2.4). There is a significant correlation between the line luminosity and M_{BH} ($0.60 \leq \rho \leq 0.98$) in all cases except for the N V line ($\rho_{\text{NV}} = 0.11$) (independent on the model adopted for the $\text{Ly}\alpha$ line) and for the Si II line (only for the *M-free* model). As we have already pointed out, the Si II line is better constrained when we fix the central wavelengths of the $\text{Ly}\alpha$ component. Our results support the hypothesis that the relation between the line FWHM and luminosity arises from our sample

Figure 3.13: FWHM as a function of the line luminosity: $\text{Ly}\alpha$ complex M -free model, left panel ($\text{Ly}\alpha_n$, black; $\text{Ly}\alpha_b$, grey; N v , blue; Si II green); $\text{Ly}\alpha$ complex M -fix model, central panel; $\text{O I}+\text{Si II}$ (red), C II (orange) and $\text{Si IV}+\text{O IV}$ (purple) lines, right panel. It is evident a strong correlation between the two quantities.

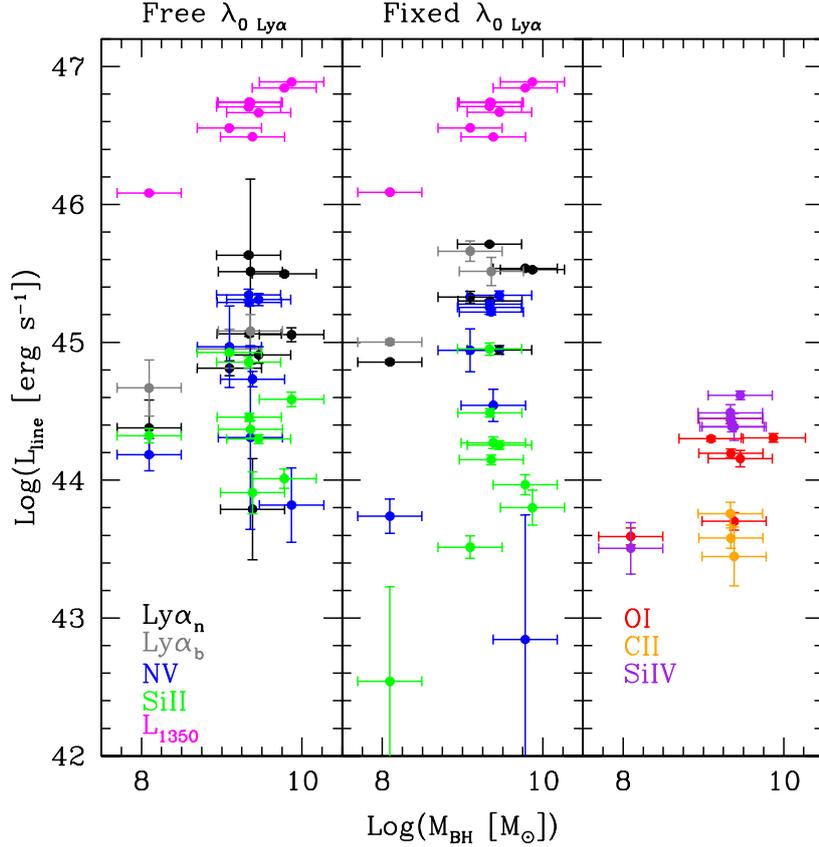


being composed by QSOs with similar Eddington ratios but different M_{BH} , with the line luminosity increasing with increasing M_{BH} . The lack of correlation for the N v line might be a further indication that this line is not emitted in the disk component of the BLR. However, these results can be biased by both the low number of objects and the small range in M_{BH} spanned by our data.

We also plot the continuum luminosity at $\lambda_{\text{rest}} = 1350 \text{ \AA}$ (magenta), which correlates strongly with M_{BH} (Pearson coefficient $\rho > 0.95$). Since the rest-frame UV continuum luminosity is thought to be emitted by gas in the QSO accretion disk which is losing angular momentum before falling on the BH, a strong correlation between the two quantities is expected.

If the gas emitting the lines is orbiting around the BH and the virial theorem is valid, we can use the line FWHM and the continuum luminosity at $\lambda_{\text{rest}} = 1350 \text{ \AA}$ to infer the BH mass (see

Figure 3.14: Line luminosity as a function of the black hole mass: Ly α complex *M-free* model, left panel ($L(1350 \text{ \AA})$, magenta; Ly α_n , black; Ly α_b , grey; N v, blue; Si II green); Ly α complex *M-fix* model, central panel; O I+Si II (red), C II (orange) and Si IV+O IV] (purple) lines, right panel. We find a significant correlation between the two quantities in all cases except for the N v line (independent on the model adopted for the Ly α line) and for the Si II line (only for the *M-free* model).



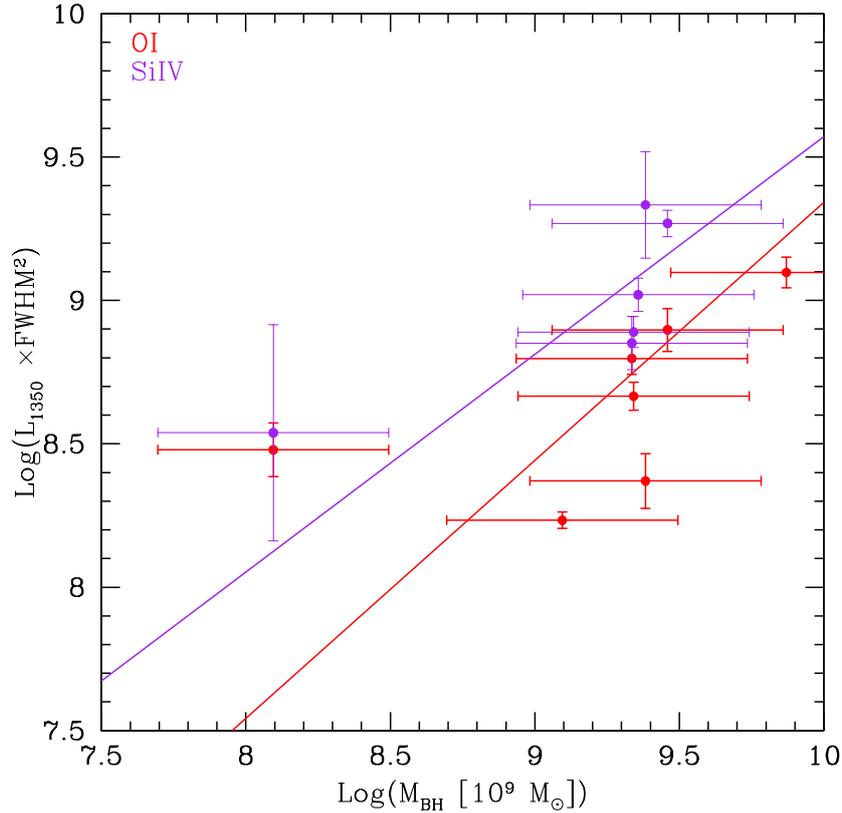
sec. 2.4.1):

$$M_{BH} \propto \text{FWHM}^2 L(1350 \text{ \AA})^{0.5}. \quad (3.14)$$

We focus on O I+Si II and Si IV+O IV] that, being isolated lines, are not significantly affected by the spectral decomposition. We then use M_{BH} estimated from the Mg II line to calibrate a possible scaling relation. In fig.3.15 we plot M_{BH} as a function of the product between the line FWHM and the continuum luminosity. The correlation coefficient is equal to 0.56 for the O I+Si II line and to 0.79 for the Si IV+O IV]. Performing a χ^2 fit of the two quantities we obtain that:

$$\begin{aligned} M_{BH} &= 2.5 (\text{FWHM}^2 L(1350 \text{ \AA})^{0.5})^{1.1}, \quad \text{O I+Si II} \\ M_{BH} &= 0.4 (\text{FWHM}^2 L(1350 \text{ \AA})^{0.5})^{1.3}, \quad \text{Si IV+O IV]} \end{aligned} \quad (3.15)$$

Figure 3.15: M_{BH} estimated from the Mg II line as a function of the product between the FWHM and the square root of the continuum luminosity measured at $\lambda_{\text{rest}} = 1350 \text{ \AA}$ (O I+Si II, red points; Si IV+O IV, purple points). Solid lines represent best fit relations.



Clearly, these relations do not have a strong predictive power at the moment, since they are based on few data-points. Moreover, they have been *calibrated* on the M_{BH} inferred from the Mg II line, which is characterized by a systematic error of 0.4 dex. Still, they open to the possibility of building new M_{BH} estimators for $z \sim 6$ QSOs combining the information gathered from different rest-frame broad lines.

3.6 SUMMARY AND CONCLUSIONS

We have presented the first consistent analysis of the optical emission properties of a large sample of $z \sim 6$ QSOs. We have collected 26 high quality optical spectra of $z \sim 6$ QSOs covering the rest-frame wavelength range $900 \text{ \AA} < \lambda_{\text{rest}} < 1500 \text{ \AA}$. This spectral range is characterized by the non-stellar continuum emission, the Ly α complex (Ly α , N v, Si II), which is strongly affected

by the neutral hydrogen absorption for $\lambda_{rest} < 1215 \text{ \AA}$, and the O I+Si II, C II and Si IV+O IV] broad lines. In order to test how the derived emission properties depend on the fitting model adopted for the Ly α complex, we performed the analysis using two different models for the Ly α line. For the *M-free* model the Ly α central wavelength is a free parameter, while for the *M-fix* model the central wavelength is fixed using a prior on the QSO systemic redshift. After proving that the redshifts inferred from the O I+Si II only show small offsets with respect to the ones inferred from the Mg II lines, we have decided to use z_{OI} as a homogeneous proxy for the QSO systemic redshifts. We have measured the power-law continuum normalization and the central wavelength, width and amplitude for each of the emission lines detected in the QSOs. For the Ly α complex the line properties have been measured twice using the two different models. Our results show that, in case of strong absorption of the Ly α line, the *M-free* model can lead to significant underestimates of the line flux. In general, the properties of the Ly α complex lines are better constrained if the *M-fix* model is adopted.

Using our estimates of the systemic redshifts we have built a median composite spectrum for the $z \sim 6$ QSOs. We have compared the resulting template with the SDSS composite spectrum obtained from QSOs with $0.1 < z < 4.8$. Both central wavelength and intensity match perfectly for all the lines except for the Ly α + N V complex, which is slightly brighter in the $z \sim 6$ composite. Given the small number of objects in our sample, this can be the result of a selection effect: luminous sources with strong Ly α lines have higher probabilities to be identified while looking for drop-out objects.

We have analyzed the emission line shifts through the comparison of the redshifts inferred from the optical lines with the ones obtained from Mg II, C IV and CO lines available from the literature. While C IV is known to present large blue-shifts with respect to its nominal transition wavelength, the redshifts obtained Mg II and CO are thought to be good proxies for the QSO systemic redshift. The C IV line is on average blue-shifted with respect to all the measured optical lines. For z_{MgII} instead, the measured absolute offsets are correlated with the line ionization potentials. The maximum offset (corresponding to $\sim 1000 \text{ km s}^{-1}$ at $z \sim 6$) is detected for the N V line, while the minimum offset ($\sim 200 \text{ km s}^{-1}$) is presented by the O I+Si II. The exception to this trend is represented by the Si IV+O IV] line: even if it is a high ionization line (61.28 eV), the redshift inferred from this line is in exceptional good agreement with the Mg II estimate. Unfortunately we are not able to compare the redshifts inferred from the optical lines with the ones estimated from the CO lines, since very few z_{CO} measurements are available.

From the fit results we have derived the UV continuum flux at $\lambda_{rest} = 1350 \text{ \AA}$, and the EW, FWHM and luminosity for all the lines. We have then tested the existence of possible correlations among these quantities:

- **EW vs continuum luminosity:** we do not find any evidence of correlation between these

two quantities. While for the low ionization lines the absence of correlation is in agreement with the *Baldwin effect*, we would expect an anti-correlation for the high ionization lines (Si IV+O IV] and N V).

- **Line luminosity vs continuum luminosity:** a strong correlation is found for all the emission lines with the exception of N V. Since this relation is simply a re-formulation of the EW-continuum luminosity one, the lack of correlation between the line and continuum luminosities might correspond to a weak detection of the anti-correlation between the N V EW and the continuum luminosity. However, we should point out that the N V line can be severely affected by the delicate spectral decomposition of the Ly α complex.
- **FWHM vs continuum luminosity:** no significant correlation is found, but this might be due to the narrow range of $L(1350 \text{ \AA})$ sampled in this study.
- **FWHM vs line luminosity:** a strong correlation between the two quantities is evident. If the dynamics of the gas emitting the optical lines is dominated by the gravity of the central BH, we will expect larger FWHM for QSOs hosting more massive BHs but with similar Eddington ratios (Labita et al. 2009). Since the distribution of Eddington ratios for $z \sim 6$ QSOs is relatively narrow, the correlation between the line luminosity and FWHM might indicate that QSOs with brighter lines are the ones hosting more massive BHs.
- **Line luminosity vs M_{BH} :** we consider the M_{BH} estimates obtained in chap. 2 from the Mg II line. We have found a significant correlation between the two quantities in all cases except for the N V (independent on the model adopted for the Ly α line) and for the Si II line (only for the *M-free* model). The Si II line is better constrained when we fix the central wavelengths of the Ly α component. Our results support the hypothesis that the relation between the line FWHM and luminosity arises from our sample being composed by QSOs with similar Eddington ratios but different BH masses, with the line luminosity increasing with increasing M_{BH} . The lack of correlation for the N V line might be a further indication that this line is not emitted in the disk. However, these results can be biased by both the low number of objects and the small range in M_{BH} spanned by our data.

Our results can be interpreted through the *disk-wind* model for the BLR (Murray et al. 1995; Leighly et al. 2004; Richards et al. 2011). In this model the BLR is a combination of a disk and a wind component, the relative importance of which depends on the actual QSO spectral energy distribution. High ionization lines can be originated in winds, while intermediate/low ionization lines are always generated in the accretion disk or at the base of the wind. From our analysis, N V is consistent with being mainly originated in the BLR wind component, while all the other emission lines seem to be mainly originated in the disk, in agreement with results for lower redshift QSOs.

The major biases in our analysis are represented by the size of the sample and by the low range of luminosities spanned by our sources. Future studies of larger sample with QSOs that populate the faint end of the QSO luminosity function are needed to investigate whether or not the results found here are applicable to all $z \sim 6$ QSOs.

Table 3.6: Continuum (normalization) and Ly α line fit parameters (central wavelength, width and amplitude) for $M - free$ and $M - fix$ respectively. Results are presented in the rest frame system of reference. Sources with a two Gaussian Ly α model are indicated with *. If a double Gaussian Ly α is considered, we list the properties of the Ly α_n component. We do not write the central wavelength value for $M - fix$ because it is fixed to the nominal one.

QSO	$a_{cont, M - free}$ [10^{-15} erg s $^{-1}$ cm $^{-2}$]	μ [Å]	σ , M-free [Å]	A , M-free [10^{-15} erg s $^{-1}$ cm $^{-2}$]	$a_{cont, M - fix}$ [10^{-15} erg s $^{-1}$ cm $^{-2}$]	σ , M-fix [Å]	A , M-fix [10^{-15} erg s $^{-1}$ cm $^{-2}$]
J1420-1602	0.226±0.005	1230±23	15±11	0.5±0.3	0.227±0.005	18.9±0.6	1.1±0.1
J0927+2001	0.0758±0.0005	1222.2±0.3	2.2±0.4	0.026±0.004	0.0758±0.0005	4.9±0.9	0.15±0.09
J1044-0125	0.0923±0.0008	1230.1±0.2	6.1±0.9	0.120±0.009	0.0919±0.0008	14.4±4.3	0.16±0.03
J0002+2550	0.124±0.002	1212±9	6±3	0.4±0.4	0.124±0.002	6.2±0.5	0.38±0.04
J0836+0054	0.1922±0.0007	1215.9±0.6	5.8±0.6	0.23±0.02	0.1923±0.0007	5.7±0.4	0.28±0.02
J1436+5007	0.081±0.003	1216.8±0.3	1.5±0.4	0.15±0.02	0.081±0.003	7±2	0.17±0.01
J0203+0012	0.0432±0.0005	1215±5	15.1±3.9	0.044±0.005	0.0432±0.0005	15.2±0.7	0.044±0.001
J0840+5624*	0.0963±0.0007	1220.3±0.8	1.3±0.9	0.05±0.05	0.0958±0.0007	3±2	0.1±0.2
J0005-0006*	0.0249±0.0004	1216.6±0.7	1.9±0.6	0.12±0.04	0.0252±0.0004	4.24±0.09	0.164±0.004
J1411+1217*	0.0731±0.0006	1219.1±0.0	1.32±0.08	0.47±0.05	0.0733±0.0006	4.7±0.3	0.44±0.05
J0841+2905	0.0540±0.0007	1225.7±0.4	15.9±0.5	0.0816±0.0008	0.0548±0.0007	16±33	0.06±0.02
J1306+0356*	0.1063±0.0005	1210±8	5±3	0.6±0.9	0.1063±0.0005	—	—
J1137+3549*	0.167±0.001	1216±3	3±1	0.1±0.1	0.169±0.001	3±1	0.15±0.04
J0353+0104	0.0385±0.0003	—	9±3	0.004±0.003	0.0384±0.0003	—	—
J1602+4228	0.086±0.002	1212±4	9±3	0.18±0.04	0.085±0.002	7.9±0.5	0.164±0.008
J0842+1218	0.0833±0.0006	1214.3±0.6	5.1±0.5	0.13±0.01	0.0841±0.0006	4.4±0.2	0.170±0.009
J0303-0019*	0.0263±0.0006	1218.883±0.003	0.76±0.04	0.235±0.009	0.0260±0.0006	3.2±0.3	0.5±0.1
J2315-0023*	0.0227±0.0009	1215±1	3.7±0.9	0.18±0.07	0.0227±0.0009	3.5±0.4	0.12±0.01
J1250+3130*	0.134±0.001	1216±1	4±2	0.6±0.4	0.134±0.001	4.1±0.2	0.37±0.03
J1048+4637	0.127±0.001	1216.9±0.9	22.7±0.7	0.118±0.001	0.127±0.001	24.3±0.3	0.1208±0.0009
J1623+3112	0.0915±0.0009	1214.8±0.1	6.8±0.2	0.532±0.004	0.0924±0.0008	8.0±0.1	0.544±0.007
J1030+0524	0.0975±0.0006	1216.5±0.2	3.4±0.1	0.278±0.007	0.0971±0.0006	4.24±0.07	0.389±0.007
J1148+5251	0.1324±0.0004	1228.2±0.4	9±1	0.099±0.002	0.1324±0.0004	26.0±0.5	0.104±0.002

Table 3.7: Ly α_b line fit parameters (central wavelength, width and amplitude) for $M - free$ and $M - fix$ respectively. Results are presented in the rest frame system of reference. Sources with a two Gaussian Ly α model are indicated with *. We do not write the central wavelength value for $M - fix$ because it is fixed to the nominal one.

QSO	μ [Å]	σ , M-free [Å]	A , M-free [10^{-15} erg s $^{-1}$ cm $^{-2}$]	σ , M-fix [Å]	A , M-fix [10^{-15} erg s $^{-1}$ cm $^{-2}$]
J1420-1602	-	-	-	-	-
J0927+2001	-	-	-	-	-
J1044-0125	-	-	-	-	-
J0002+2550	-	-	-	-	-
J0836+0054	-	-	-	-	-
J1436+5007	-	-	-	-	-
J0203+0012	-	-	-	-	-
J0840+5624*	1222 \pm 5	4 \pm 3	0.05 \pm 0.04	9 \pm 5	0.11 \pm 0.03
J0005-0006*	1219 \pm 4	6 \pm 2	0.07 \pm 0.02	26.9 \pm 0.7	0.036 \pm 0.001
J1411+1217*	1222.5 \pm 0.5	3.1 \pm 0.6	0.28 \pm 0.07	12 \pm 1	0.36 \pm 0.04
J0841+2905	-	-	-	-	-
J1306+0356*	1230.1 \pm 0.8	8 \pm 2	0.13 \pm 0.002	2.8 \pm 0.1	1.1 \pm 0.2
J1137+3549*	1225 \pm 1	13.4 \pm 0.7	0.222 \pm 0.003	14 \pm 20	0.2 \pm 0.4
J0353+0104	-	-	-	-	-
J1602+4228	-	-	-	-	-
J0842+1218	-	-	-	-	-
J0303-0019*	1220.3 \pm 0.4	4.9 \pm 0.3	0.144 \pm 0.006	10.3 \pm 0.5	0.14 \pm 0.01
J2315-0023*	1221 \pm 4	7 \pm 2	0.07 \pm 0.03	11 \pm 1	0.12 \pm 0.01
J1250+3130*	1228 \pm 3	8 \pm 10	0.27 \pm 0.08	20 \pm 3	0.36 \pm 0.02
J1048+4637	-	-	-	-	-
J1623+3112	-	-	-	-	-
J1030+0524	-	-	-	-	-
J1148+5251	-	-	-	-	-

Table 3.8: N v line fit parameters (central wavelength, width and amplitude) for $M - free$ and $M - fix$ respectively. Results are presented in the rest frame system of reference. Sources with a two Gaussian Ly α model are indicated with *.

QSO	μ , M-free [Å]	σ , M-free [Å]	A, M-free [10^{-15} erg s $^{-1}$ cm $^{-2}$]	μ , M-fix [Å]	σ , M-fix [Å]	A, M-fix [10^{-15} erg s $^{-1}$ cm $^{-2}$]
J1420-1602	1245±1	3±2	0.06±0.05	1246.3±0.3	1.0±0.4	0.05±0.01
J0927+2001	1233.1±0.7	11.4±0.7	0.052±0.001	1233.7±0.9	11.1±0.8	0.052±0.001
J1044-0125	1242.2±0.4	2.7±0.4	0.041±0.008	1240.4±0.6	4.3±0.9	0.04±0.02
J0002+2550	1234±1	9.0±0.9	0.137±0.008	1236.8±0.8	8.9±0.8	0.137±0.007
J0836+0054	1234.9±0.6	10.4±0.9	0.174±0.003	1234.0±0.5	10.9±0.9	0.175±0.003
J1436+5007	1229±1	11±1	0.095±0.005	1234±3	8.3±2	0.085±0.008
J0203+0012	1248.9±0.7	6.0±0.9	0.012±0.002	1247±1	5±1	0.011±0.001
J0840+5624*	1235±2	12±1	0.082±0.005	1236±5	11±3	0.08±0.02
J0005-0006*	1240.1±0.6	6.2±0.6	0.024±0.001	1242.0±0.3	4.2±0.4	0.0127±0.0009
J1411+1217*	1230±2	6±2	0.14±0.02	1245±2	9.7±0.9	0.09±0.01
J0841+2905	—	—	—	1234±5	11.1±2	0.044±0.09
J1306+0356*	1242.0±0.4	4.1±0.9	0.05±0.03	1231.6±0.1	11.18±0.2	0.135±0.001
J1137+3549*	1238.7±0.6	2.9±0.8	0.019±0.005	1236±39	9±30	0.1±2
J0353+0104	1237.8±0.2	5.6±0.3	0.0563±0.0009	1239.2±0.6	4.3±0.4	0.048±0.003
J1602+4228	1238.5±0.8	5.8±0.8	0.057±0.006	1237.5±0.5	6.0±0.6	0.059±0.004
J0842+1218	1230.2±0.6	11.7±0.5	0.148±0.002	1229.7±0.4	12.3±0.4	0.150±0.001
J0303-0019*	1239.5±0.3	5.8±0.3	0.069±0.002	1240.3±0.1	4.3±0.2	0.067±0.002
J2315-0023*	1242.0±0.3	3.1±0.3	0.029±0.002	1242.2±0.2	3.0±0.4	0.030±0.004
J1250+3130*	1240±2	5±1	0.2±0.2	1239.5±0.2	4.9±0.5	0.13±0.03
J1048+4637	—	—	—	1242±1	1.5±1	0.004±0.002
J1623+3112	1236.5±0.2	7.3±0.3	0.256±0.005	1237.4±0.1	6.4±0.2	0.238±0.009
J1030+0524	1233.5±0.1	10.2±0.2	0.157±0.002	1233.5±0.1	9.8±0.2	0.160±0.002
J1148+5251	1243.2±0.3	2.5±0.5	0.021±0.004	—	—	—

Table 3.9: Si II line fit parameters (central wavelength, width and amplitude) for $M - free$ and $M - fix$ respectively. Results are presented in the rest frame system of reference. Sources with a two Gaussian Ly α model are indicated with *.

QSO	μ , M-free [Å]	σ , M-free [Å]	A, M-free [10^{-15} erg s $^{-1}$ cm $^{-2}$]	μ , M-fix [Å]	σ , M-fix [Å]	A, M-fix [10^{-15} erg s $^{-1}$ cm $^{-2}$]
J1420-1602	1269±1	6±1	0.10±0.03	1267.9±0.4	6.0±0.4	0.097±0.005
J0927+2001	1263.3±0.4	5.2±0.4	0.0152±0.0009	1263.4±0.4	5.2±0.4	0.0153±0.0009
J1044-0125	1256±2	14±1	0.025±0.001	1260±3	12±2	0.023±0.003
J0002+2550	1264.4±0.6	3.8±0.7	0.06±0.02	1266.5±0.7	3.9±0.8	0.05±0.02
J0836+0054	1259.4±0.7	7.3±0.4	0.064±0.004	1259.3±0.6	7.2±0.4	0.062±0.005
J1436+5007	1254.1±0.7	2.6±0.8	0.028±0.006	1253.9±0.7	2.7±0.8	0.032±0.006
J0203+0012	—	—	—	1254.4±0.5	1.8±0.8	0.007±0.003
J0840+5624*	1265.1±0.8	4.2±0.8	0.020±0.002	1264±1	4.9±0.8	0.021±0.003
J0005-0006*	1248±50	17±1	0.012±0.001	1265±3	3±3	0.001±0.001
J1411+1217*	1248±50	9.9±0.3	0.082±0.002	1266.0±0.2	1.7±0.3	0.019±0.002
J0841+2905	1260.3±0.4	6.3±0.4	0.020±0.001	1259±1	7±1	0.024±0.008
J1306+0356*	1259±1	8±1	0.027±0.002	1262.6±0.4	5.7±0.4	0.022±0.001
J1137+3549*	1259.9±0.5	5.7±0.5	0.035±0.002	1261±2	6±2	0.04±0.04
J0353+0104	1257±2	6±2	0.008±0.002	1257±1	10±1	0.0107±0.0004
J1602+4228	1243±50	21±1	0.043±0.004	1243±50	22±1	0.043±0.004
J0842+1218	1261.1±0.4	5.7±0.3	0.030±0.001	1261.8±0.3	5.4±0.3	0.028±0.001
J0303-0019*	1262±2	9±2	0.016±0.002	1258±1	11±1	0.0160±0.0009
J2315-0023*	1248±50	12.3±0.7	0.022±0.001	1249±3	11±2	0.021±0.003
J1250+3130*	1252±2	13±1	0.076±0.006	1262±4	10±2	0.04±0.02
J1048+4637	1270.4±0.7	6.1±0.8	0.014±0.001	1270.6±0.7	5.9±0.8	0.013±0.001
J1623+3112	1258±1	11±1	0.054±0.002	1255±2	13±1	0.059±0.003
J1030+0524	1263.0±0.3	7.0±0.4	0.034±0.001	1262.4±0.4	7.4±0.5	0.034±0.001
J1148+5251	1258±1	9±1	0.033±0.002	1264±0.9	4.5±0.9	0.011±0.002

Table 3.10: O I+Si II, C II and Si IV+O IV] line fit parameters (central wavelength, width and amplitude). Results are presented in the rest frame system of reference. Sources with a two Gaussian Ly α model are indicated with *.

QSO	$\mu_{OI+SiII}$ [Å]	$\sigma_{OI+SiII}$ [Å]	$A_{OI+SiII}$ [10^{-15} erg s $^{-1}$ cm $^{-2}$]	μ_{CII} [Å]	σ_{CII} [Å]	A_{CII} [10^{-15} erg s $^{-1}$ cm $^{-2}$]	$\mu_{SiIV+OIV}$ [Å]	$\sigma_{SiIV+OIV}$ [Å]	$A_{SiIV+OIV}$ [10^{-15} erg s $^{-1}$ cm $^{-2}$]
J1420-1602	1306±1	10±2	0.055±0.005	1340.1±0.7	3.1±0.8	0.037±0.006	1378±50	15±4	0.10±0.06
J0927+2001	1305.1±0.6	6.5±0.6	0.0096±0.0008	1340.0±0.9	8±1	0.0101±0.0008	—	—	—
J1044-0125	1305±1	14±2	0.0131±0.0008	1350±50	6.5±0.7	0.023±0.004	1403±1	11±1	0.021±0.002
J0002+2550	1303±1	7±1	0.028±0.003	—	—	—	—	—	—
J0836+0054	1305.6±0.3	5.3±0.3	0.041±0.002	—	—	—	—	—	—
J1436+5007	—	—	—	—	—	—	—	—	—
J0203+0012	—	—	—	—	—	—	—	—	—
J0840+5624*	1306±1	5.1±1	0.006±0.001	—	—	—	1397.8±0.8	9.6±0.8	0.025±0.002
J0005-0006*	1305.8±0.5	4.7±0.5	0.0080±0.0007	—	—	—	1401±2	5±2	0.006±0.002
J1411+1217*	1305.8±0.1	3.5±0.1	0.054±0.001	—	—	—	—	—	—
J0841+2905	1305.5±0.6	4.0±0.6	0.009±0.001	—	—	—	—	—	—
J1306+0356*	—	—	—	—	—	—	—	—	—
J1137+3549*	1303.7±0.8	8.0±0.8	0.023±0.002	—	—	—	1396.4±0.5	9.4±0.6	0.024±0.001
J0353+0104	1305.7±0.5	4.2±0.5	0.0071±0.0007	1337±2	5±2	0.0033±0.001	1392.8±0.6	7.7±0.6	0.040±0.003
J1602+4228	1306.1±0.5	5.7±0.6	0.025±0.002	—	—	—	1403±3	13±3	0.0105±0.0008
J0842+1218	1304.8±0.5	7.6±0.7	0.016±0.002	—	—	—	—	—	—
J0303-0019*	1305.8±0.4	5.8±0.4	0.0157±0.0009	—	—	—	1395.9±0.8	12.5±0.7	0.028±0.001
J2315-0023*	—	—	—	—	—	—	—	—	—
J1250+3130*	1305.3±0.5	5.9±0.5	0.036±0.003	—	—	—	1397.7±0.6	8.1±0.6	0.061±0.004
J1048+4637	—	—	—	—	—	—	—	—	—
J1623+3112	1305.1±0.4	6.8±0.4	0.035±0.002	1337.2±0.4	2.6±0.4	0.019±0.002	1395.5±0.9	7.7±0.8	0.034±0.003
J1030+0524	1305.8±0.3	5.8±0.3	0.022±0.001	1335.2±0.4	3.7±0.5	0.0084±0.0009	1398.9±0.6	8.1±0.5	0.029±0.002
J1148+5251	1304.9±0.4	9.5±0.6	0.0171±0.0006	—	—	—	—	—	—

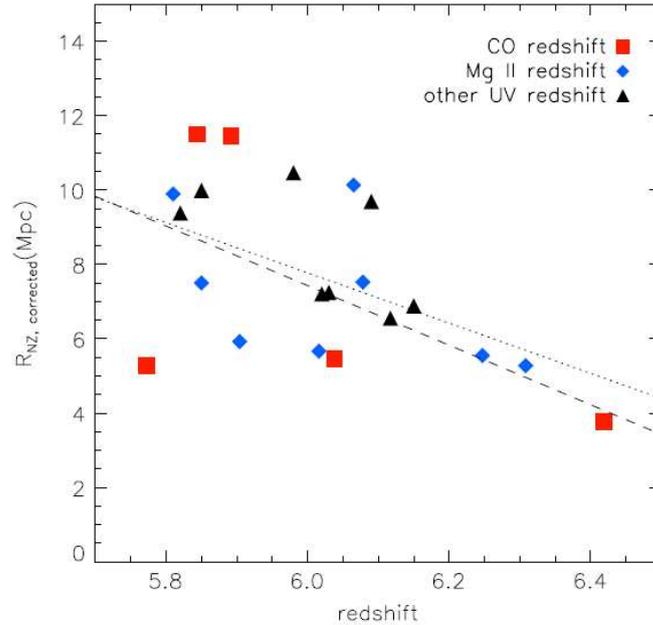
IONIZATION NEAR ZONE AND GUNN PETERSON ABSORPTION*

CONTEXT

After the epoch of recombination, $z \sim 1100$, the universe became mostly neutral, until the first generation of astrophysical sources reionized the inter-galactic medium (IGM) and ended the dark ages. The reionization is expected to be a complex process, extended in time and very inhomogeneous. Observations have set the first constraints to the epoch of reionization. The large scale polarization of the Cosmic Microwave Background (CMB) is consistent with an instantaneous reionization at $z \sim 11 \pm 3$ (Komatsu et al. 2010). Luminous high redshift QSOs have been largely used as indirect probes of the IGM ionization state through Ly α forest studies. The Ly α forest at $z \lesssim 3$ is largely transparent, with occasional absorption associated to rare clouds of neutral hydrogen. At $z \gtrsim 5$, the Ly α forest becomes increasingly opaque (e.g. Becker et al. 2001). Due to the strong absorption cross-section of the Ly α transition, even small amounts of neutral hydrogen (volume densities $n_{\text{HI}} \sim 10^{-5} - 10^{-4}$) are sufficient to render the forest optically thick. Studies of the Gunn-Peterson (GP, Gunn & Peterson 1965) absorption in high redshift QSO spectra qualitatively suggest an end of the re-ionization process at $z \sim 6$, indicated by a rapid increase in the neutral fraction from $n_{\text{HI}} < 10^{-4}$ at $z \leq 5.5$ to $n_{\text{HI}} > 10^{-3}$ at $z \geq 6$. Unfortunately, these measurements have a fundamental limitation as the GP absorption saturates at low neutral fractions. Moreover they are highly dependent on the models adopted for the ionization, density and UV background field. McGreer et al. (2011), showed that performing an (almost) model independent analysis of the covering fraction of dark pixels in the Ly α -Ly β forests of high redshift QSOs, an end of the re-ionization as late as $z = 5$ cannot be ruled out. Another indirect probe of the IGM ionization state can be derived from the study of small windows of Ly α trans-

*The material presented in this chapter is partly included in the paper *Optical properties of $z \sim 6$ quasars*, De Rosa et al., 2011b, to be submitted

Figure 4.1: Adapted from Carilli et al. (2010), fig. 1. Evolution of $R_{NZ,corr}$ as a function of the quasar systemic redshift. Red squares: sources with CO redshift estimates. Blue diamonds: sources with redshifts determined from the Mg II emission line. Black triangles: sources with redshifts estimated from other optical lines (e.g. Ly α + N v). The long-dashed line shows a weighted linear fit to the data. The dotted line shows an unweighted fit.



mission observed in close proximity of the QSOs (see e.g., Mesinger & Haimann 2004; Wyithe & Loeb 2004; Fan et al. 2006b; Bolton & Haenelt 2007a). These highly ionized *near zones* (NZ) lie between the edge of the GP trough and the QSO Ly α emission line, and arise because the hydrogen close to these sources is highly ionized by the QSO itself. Measurements of the NZ sizes cannot be used to give absolute constraints on n_{HI} at a given redshift. However it is possible to give relative constraints on the IGM ionization state studying the evolution of the NZ radii as a function of the QSO systemic redshifts.

4.1 IONIZATION NEAR ZONES

UV photons from luminous QSO are expected to ionize an H II region (Strömgren sphere) in the neutral/partially neutral IGM (Madau & Rees 2000; Cen & Haimann 2000). The observational signature of the QSO ionization is an excess of transmission in the blue wing of the Ly α emission line before the onset of the GP absorption. This excess of radiation is observable in spectra of $z \sim 6$ QSOs. The evolution of a cosmological H II region expanding into a neutral IGM can be

parametrized as (e.g. Shapiro & Giroux 1987; Madau et al. 1999):

$$n_{\text{H}} \left(\frac{dV}{dt} - 3HV \right) = \dot{N}_{\text{Q}} - \frac{n_{\text{H}}V}{t_{\text{rec}}} \quad (4.1)$$

where n_{H} is the average number density of neutral hydrogen, V is the volume of the H II region, H is the Hubble constant at the absorber redshift, \dot{N}_{Q} is the number of ionizing photons per unit time emitted by the QSO, and t_{rec} is the volume averaged gas recombination time. Haimann (2002) showed that when recombination is not important and the QSO life time is much less than the Hubble time at $z = z_{\text{sys}}$ the radius of the H II region (in proper Mpc) can be written as:

$$R_{\text{HII}} = 8.0 f_{\text{H}}^{-1/3} \left(\frac{\dot{N}_{\text{Q}}}{6.5 \times 10^{57} \text{s}^{-1}} \right)^{1/3} \left(\frac{t_{\text{Q}}}{2 \times 10^7 \text{yr}} \right)^{1/3} \left(\frac{1+z_{\text{Q}}}{7} \right)^{-1} \quad (4.2)$$

where z_{Q} is the QSO systemic redshift and t_{Q} is the QSO age. The equation above neglects possible IGM clumpiness. Many authors (Mesinger & Haimann 2004; Wyithe & Loeb 2004; Wyithe et al. 2005; Yu & Lu 2005) have tried to constraint the IGM neutral fraction using the measured sizes of H II regions around $z \sim 6$ QSOs. As Fan et al. (2006b) pointed out though, deriving an absolute measurement of the IGM neutral fraction from eq. 4.2 is complicated by various factors:

- uncertainties on QSO lifetimes and ionizing rates
- background ionizing photons from galaxies in the over-dense environment of luminous $z \sim 6$ QSOs
- large scale clustering of IGM around luminous $z \sim 6$ QSOs
- clumpiness of the IGM (see also Bolton & Haenelt 2007a,b; Maselli et al. 2007, 2009)

Assuming these systematics to be on average the same for QSOs in a narrow redshift range, Fan et al. (2006b) used the evolution in measured H II region radii around $z \sim 6$ QSOs as a relative measurement of the changes in the IGM ionization state.

The usual definition of Strömgen sphere radius is the distance at which the transmitted flux reaches the 0 value for the first time. This definition does not suite this particular scientific case since the IGM transmitted flux is characterized by a mixture of dark gaps and sharp transmission spikes (see Fan et al. 2006b) and using the zero flux as a measure of the H II sphere boundary would be intrinsically dependent on the resolution and smoothing length used. Also, if the IGM surrounding the QSO is highly ionized, the transmitted flux is not going to reach on average the zero value while it is going to show large fluctuations that could lead to an underestimate of the radius. On the contrary, if the IGM is highly ionized reaching the 0 flux does not necessarily

imply reaching the boundary of the QSO ionizing sphere: the IGM could still be ionized by the QSO to a level much higher than the surrounding material but with an optical depth higher than the detection limit. For these reasons Fan et al. (2006b) somewhat arbitrarily defined the NZ radius (R_{NZ}) as the distance at which the ratio between the observed flux and the intrinsic emitted flux F_{obs}/F_{em} (transmission, T) drops for the first time to 0.10 after smoothing the spectrum to a resolution of 20 Å. Since then, this definition has been used by many authors in their analysis.

Fan et al. (2006b) measured R_{NZ} for a sample of 16 ESI SDSS $z \sim 6$ QSOs. After correcting the values for the QSO intrinsic brightness by scaling R_{NZ} to a common absolute magnitude (assumed to be proportional to \dot{N}_Q):

$$R_{NZ,corr} = R_{NZ} \times 10^{0.4(27+M_{1450})/3}, \quad (4.3)$$

they found a strong correlation between the corrected H II region size and the systemic redshift (Pearson coefficient of 0.72):

$$R_{NZ,corr} = 6.7 - 8.1(z - 6) \text{ Mpc}. \quad (4.4)$$

Assuming the simplest possible physical picture (see eq. 4.2, $R_{NZ} \propto f_{HI}^{-1/3} (1 + z_Q)^{-1}$), Fan et al. (2006b) interpreted the observed decrease in R_{NZ} as an increase in the volume-averaged neutral fraction by a factor of ~ 10 . Wyithe et al. (2008) and Bolton et al. (2010) have proposed a more complex interpretation of the relation between the H II radii and the redshift. During the percolation phase (end of the reionization), the increase in the mean free path allows the background ionizing photons to contribute significantly to the ionization around the edge of the NZ region, effectively increasing the observed sizes of the regions. Observationally, reaching the final stages of cosmic reionization would yield a decrease of R_{NZ} with redshift. Mortlock et al. (2011) have recently published the highest redshift quasar found to date, ULAS J1120 + 0641 at $z = 7.085$, discovered in the UKIRT Infrared Deep Sky Survey (UKIDSS, Lawrence et al. 2007). In their analysis of the NZ region, Mortlock et al. (2011) found a significantly smaller region with respect to comparably luminous $z \sim 6$ QSOs. They also found that the NZ transmission profile of the ~ 7 QSO is smoother as compared to the transmission profiles of J1030+0524 ($z = 6.30$) and J1148+5251 ($z = 6.42$). Assuming that the quasar is embedded in a homogeneous IGM, they concluded that the observed transmission profile is consistent with a surrounding IGM with $\langle n_H \rangle > 0.1$ at $z \sim 7$.

One of the major uncertainties in R_{NZ} measurements is the definition of the QSO **systemic redshift**. Fan et al. (2006b) had available redshift measurements from Mg II or CO for only 5 objects, while for the remaining sources they converted redshifts measured from higher ionization lines (typically C IV, Si IV+O IV) or N V) into z_{MgII} following Richards et al. (2002): $z_{MgII} = z_{High-ionization} + 0.02$ with $\sigma_z = 0.02$ (corresponding to $\sigma_{R_{NZ}} \sim 1.2$ Mpc). Carilli et al. (2010) re-proposed the R_{NZ} measurement (defined as in Fan et al. 2006b) for a sample of 23 SDSS QSOs

at $z \sim 6$, with reduced systematics on the QSO redshift estimates. They used CO redshifts for 5 sources, Mg II redshifts for 9 objects and optical redshifts for the remaining QSOs. We show their results for $R_{NZ,corr}$ as a function of the QSO systemic redshifts in fig. 4.1 (adapted from Carilli et al. 2010, fig. 1). $R_{NZ,corr}$ was computed by applying the same correction for the QSO ionizing flux as in Fan et al. (2006b) (eq. 4.3). Carilli et al. (2010) found a similar trend in redshift with respect to the one found by Fan et al. (2006b):

$$R_{NZ,corr} = (7.7 \pm 0.4) - (6.8 \pm 2.1)(z - 6) \text{ Mpc}, \quad (4.5)$$

but a lower correlation between the two variables (Spearman coefficient $\rho = 0.51$).

A fundamental step in this measurement is the definition of the intrinsic emitted flux. In previous chapters we have seen how models can be significantly different as a function of the priors we set during the fitting analysis. In order to assess the significance of the observed trend, we here measure $R_{NZ,corr}$ for the sample of optical spectra described in sec. 3.1 while using the two different models for the emitted flux outlined in sec. 3.2.3 (*M-free*: the Ly α central wavelength is a free parameter; *M-fix*: the Ly α central wavelength is fixed using the QSO systemic redshift). We remove from our sample the WLQSOs, the BAL QSOs and those sources for which we do not have a reliable measurement of the systemic redshift (J1436+5007, J0841+2905, J0203+0012). This leaves us with a sample of 16 objects. For 13 QSOs in the sub-sample we are using our estimates of the systemic redshift from the O I+Si II line (see sec. 3.2.4). For the 3 remaining sources we are using systemic redshifts estimated from the Mg II line. In sec. 3.4 we have seen that the two redshift estimates are well in agreement, with an average difference $\langle z_{MgII} - z_{OI} \rangle \sim 0.004$ and a 1σ standard deviation of 0.001. This way we should be able to significantly reduce the systematics due to redshift measurements from different emission lines, assuring a consistent comparison amongst individual $R_{NZ,corr}$ estimates for the first time.

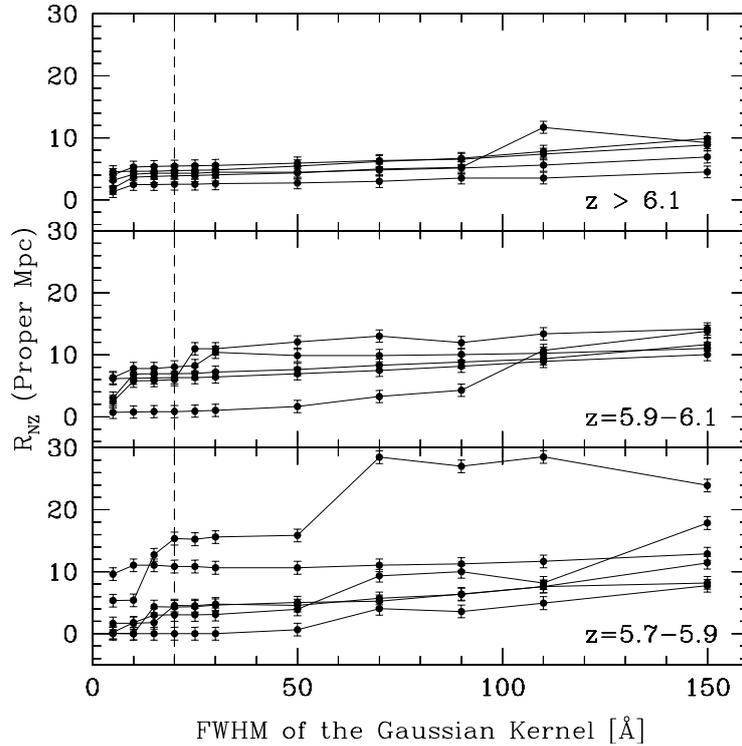
4.1.1 TESTING THE DEFINITION OF RNZ

First of all we test the robustness of the NZ radius definition, by analyzing its dependence on the choice of the Gaussian smoothing kernel. We smooth our spectra using 11 Gaussian kernels with increasing FWHM (5, 10, 15, 20, 25, 30, 50, 70, 90, 110, 150 Å) and for each case measure the redshift $z_{NZ} = \frac{\lambda_{NZ}}{1215.67} - 1$ at which $f_{obs} = 0.10 f_{em}$. For each z_{NZ} we estimate R_{NZ} as:

$$R_{NZ} = (D_{sys} - D_{NZ}) \frac{1}{1 + z_{sys}} \quad (4.6)$$

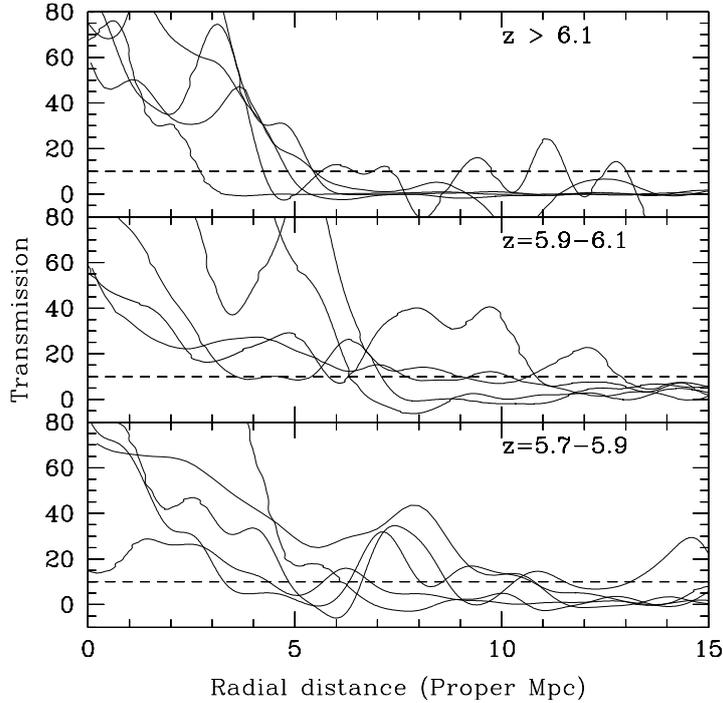
where D_{sys} and D_{NZ} are the comoving distances implied by z_{sys} and z_{NZ} respectively. Assuming a systematic error on both z_{NR} and z_{sys} of ~ 0.01 , we obtain a typical error $\sigma_{R_{NZ}} \sim 1$ proper Mpc. Fig. 4.4 shows the evolution of R_{NZ} as a function of different smoothing kernels for the *M-fix*

Figure 4.2: *M-fix*. Evolution of R_{NZ} as a function of different smoothing kernels ($z > 6.1$, top panel; $5.9 \leq z < 6.1$, central panel; $z \leq 5.9$, bottom panel). The vertical dashed line indicates the smoothing kernel used by Fan et al. (2006b). At $z > 6.1$ the estimate of R_{NZ} does not depend significantly on the choice of the smoothing kernel. At lower redshifts a higher scatter is visible both in the R_{NZ} measurement at a fixed smoothing resolution and in the evolution of R_{NZ} as a function of different smoothing kernels.



model ($z > 6.1$, top panel; $5.9 \leq z < 6.1$, central panel; $z \leq 5.9$ bottom panel). The vertical dashed line indicates the smoothing kernel used by Fan et al. (2006b). We notice a trend in the evolution of R_{NZ} as a function of the systemic redshift. At $z > 6.1$ the estimate of R_{NZ} is almost independent on the choice of the smoothing kernel. For progressively decreasing redshifts a higher scatter in R_{NZ} at a fixed smoothing resolution is evident. This can be the result of different QSO ionizing fluxes. At lower redshifts it is also visible an increase of the scatter in the evolution of R_{NZ} as a function of different smoothing kernels. By looking at the QSO spectra (see Fig. 3.2-3.7) we notice that at lower redshifts there is an increase in the fluctuations of the transmitted flux. For large enough smoothing kernels then, it is possible to wash out the signature of the H II region by convolving the excess of flux in the NZ region with a transmission spike originated in the surrounding IGM. For J0840+5624 this effect is particularly strong (fig. 4.4, bottom panel, profile with the highest values of R_{NZ}). This source shows an almost completely absorbed Ly α line, characterized by a sharp absorption profile and a broad transmission spike at $z_{spike} \sim z_{sys}$, which results in an ambiguous definition of R_{NZ} also at small smoothing lengths. The length of the smoothing kernel at which the convolution with the IGM spikes starts, depends on the IGM

Figure 4.3: *M-fix*. Transmitted flux smoothed at 20 \AA within the H II regions and in the surrounding IGM ($z > 6.1$, top panel; $5.9 \leq z < 6.1$, central panel; $z \leq 5.9$ bottom panel). The horizontal dashed line indicates our definition of R_{NZ} ($f_{obs} = 0.10 f_{em}$). The $z \sim 6$ QSOs show a steeper decline towards the IGM value and cross the dashed line on average at lower radii with respect to the lower redshift sources.

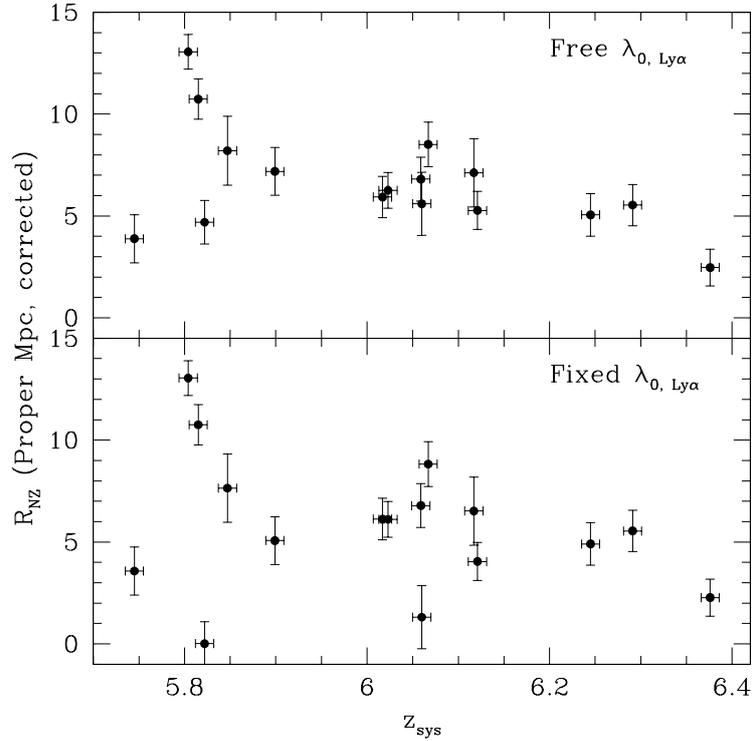


structure surrounding each QSO. Hence, the definition of a common global smoothing kernel does not appear to be the best strategy. For comparison with previous results (e.g. Fan et al. 2006b) we adopt a Gaussian smoothing kernel with a FWHM= 20 \AA for the analysis that follows. In tab. 4.1 we list z_{sys} , z_{NZ} , D_{sys} , D_{NZ} and the corresponding R_{NZ} measurement for the *M-fix* model.

4.1.2 RNZ AS A FUNCTION OF THE QSO SYSTEMIC REDSHIFT

Fig. 4.3 shows the transmitted flux smoothed over 20 \AA within the H II regions and in the surrounding IGM for the *M-fix* model. The horizontal dashed line indicates our adopted definition of R_{NZ} ($f_{obs} = 0.10 f_{em}$). The $z \sim 6$ QSOs seem to have a steeper decline towards the IGM value, which only at the highest redshifts reaches the zero transmitted flux. They also cross the dashed line on average at smaller radii with respect to the lower redshift QSOs. Even though a trend as a function of the QSO systemic redshifts is evident, it is difficult to identify a global behavior at small radial distances, given the high level of scatter and the low number of sources for each red-

Figure 4.4: Evolution of $R_{NZ,corr}$ as a function of the systemic redshift for the considered models (M -free top panel, M -fix bottom panel). The degree of correlation depends on the model adopted. If we compare our measurements with previous results (see fig. 4.1), we notice that the degree of correlation is significantly reduced.



shift bin. Moreover, as we have pointed out already, significant variations in the transmitted flux could also arise because of different QSO ionizing fluxes. If we want to compare individual measurements of QSOs of various luminosities we need to rescale R_{NZ} to a common ionizing flux. Following Fan et al. (2006b), we then correct our R_{NZ} estimates using eq. 4.3, where we measure the absolute magnitude M_{1450} extrapolating our power-law continuum model at $\lambda_{rest} = 1450 \text{ \AA}$ (the extinction correction is included, see sec. 3.2.1). Our M_{1450} estimates differ in some cases from literature values. This can be the result of the different measurement procedure: Fan et al. (2006b) extrapolated the apparent magnitude at $\lambda_{rest} = 1280 \text{ \AA}$ (corrected for galactic extinction) to $\lambda_{rest} = 1450 \text{ \AA}$ assuming a trend $f_\nu \propto \nu^{0.5}$. In tab. 4.2 we list the measured absolute magnitudes M_{1450} and R_{NZ} corrected for the QSO intrinsic luminosity (M -free, col. 2; M -fix, col. 3). Errors are obtained through standard error propagation and are dominated by $\sigma_{R_{NZ}} \sim 1$ proper Mpc. For 4 of the sources we obtain results for the two models that are significantly different (J0303-0019, J0840+5624, J2315-0023 and J1250+3130). For these 4 QSOs the M -fix model suggests a severe absorption of the Ly α line and a corresponding underestimate of the intrinsic profile by the M -free model. Moreover J0303-0019, J2315-0023 and J1250+3130 present a noisy Ly α absorption profile characterized by sharp transition spikes. The combination of these two factors leads

to the high differences between the results obtained assuming respectively *M-free* and *M-fix* as models for the emitted flux. For J0840+5624 instead the situation is extreme: the R_{NZ} estimate we obtain for the *M-fix* model is clearly unphysical. As we have already seen in sec. 4.1.1, this source is characterized by a peculiar absorption profile that complicates the definition of the NZ radius.

Fig. 4.4 shows the evolution of $R_{NZ,corr}$ as a function of the systemic redshift for the two models (*M-free* top panel, *M-fix* bottom panel). It is evident that the degree of correlation depends on the model adopted for the emitted flux. If a decreasing trend is still visible for the *M-free* model, the correlation appears to be significantly less strong for the *M-fix* model. If we quantify the degree of correlation using the Spearman correlation coefficient ρ , we obtain $\rho = -0.37$ for *M-free*, implying a probability of 15% that there is no correlation between $R_{NZ,corr}$ and z_{sys} , and $\rho = -0.27$ for *M-fix*, with a probability of 32% of no correlation. Performing a χ^2 fit to the data assuming a linear correlation between the two variables, we obtain relations that are similar to the ones obtained by Fan et al. (2006b) and Carilli et al. (2010):

$$R_{NZ,corr} = 7.0 - 9.1(z - 6), \chi_v^2 = 5, \text{ } M - free \quad (4.7)$$

$$R_{NZ,corr} = 6.3 - 7.5(z - 6), \chi_v^2 = 10, \text{ } M - fix \quad (4.8)$$

but in both cases, the reduced χ^2 largely exclude the goodness of the model. If we assume that the R_{NZ} definition is ambiguous for J0840+5624, we can remove the corresponding point from the *M-fix* data set and repeat the analysis. The Spearman correlation coefficient for the resulting sub-sample is equal to -0.42, implying a probability of no correlation of 12%, while from the χ^2 linear fit we obtain:

$$R_{NZ,corr} = 6.9 - 10.6(z - 6), \chi_v^2 = 6, \text{ } M - fix, \text{ } sub - sample \quad (4.9)$$

Given the low number of sources in our sample, the correlation is significantly improved, but the linear relation can still be ruled out. We need to specify that J0840+5624 is included in both Fan et al. (2006b) and Carilli et al. (2010) samples and it does not result to be an outlier in their studies. With the exclusion of J0840+5624 from the data-set, the agreement between *M-free* and *M-fix* is highly improved. We can conclude that while performing a consistent analysis, choices of different emitted flux models cause on average variations on the $R_{NZ,corr}$ estimates that are lower than the typical errors. If we now compare our measurements with previous results, we notice that the degree of correlation between $R_{NZ,corr}$ and z_{sys} is significantly reduced. Since the analysis has been performed mainly on the same spectra, this indicates that the $R_{NZ,corr}$ measurements are strongly dependent on the adopted procedure. Summarizing, it does not appear to be straightforward to give relative constraints on the IGM ionization state through the analysis of the relation between the $R_{NZ,corr}$ and the redshift, once the spectra are analyzed and fitted consistently.

Table 4.1: R_{NZ} estimates for the M -fix model.

QSO name ¹	z_{sys} ²	D_{sys} (Mpc) ³	z_{NZ} ⁴	D_{NZ} (Mpc) ⁵	R_{NZ} (Mpc) ⁶
J0927 + 2001	5.75	8307.7	5.70	8287.2	3.0
J0002 + 2550	5.82	8340.2	5.66	8266.3	10.8
J0836 + 0054	5.80	8334.7	5.58	8230.3	15.3
J0840 + 5624	5.82	8342.9	5.82	8342.9	0.01
J0005 – 0006	5.85	8354.3	5.78	8323.1	4.5
J1411 + 1217	5.90	8377.6	5.83	8347.8	4.3
J1306 + 0356	6.02	8429.7	5.92	8387.5	6.0
J1137 + 3549	6.02	8432.3	5.91	8383.3	7.0
J1602 + 4228	6.07	8451.3	5.93	8394.5	8.0
J0842 + 1218	6.06	8447.9	5.95	8403.3	6.3
J0303 – 0019	6.06	8448.3	6.05	8442.3	0.8
J2315 – 0023	6.12	8472.8	6.05	8445.1	3.9
J1250 + 3130	6.12	8474.5	6.05	8443.8	4.3
J1623 + 3112	6.25	8526.6	6.16	8492.5	4.7
J1030 + 0524	6.29	8545.6	6.19	8505.9	5.5
J1148 + 5251	6.38	8580.7	6.33	8562.3	2.5

¹ QSO name;

^{2,3} systemic redshift and corresponding comoving distance;

^{4,5} z_{NZ} and corresponding comoving distance

⁶ NZ radius.

Table 4.2: $R_{NZ,corr}$: comparison between the two models.

QSO name ¹	M_{1450} ²	$R_{NZ,corr}$ (proper Mpc) ³	
		Free $\lambda_{0,Ly\alpha}$	Fixed $\lambda_{0,Ly\alpha}$
J0927+2001	-26.46	3.8±1.2	3.6±1.2
J0002+2550	-27.03	10.7±1.0	10.7±1.0
J0836+0054	-27.53	13.0±0.9	13.0±0.9
J0840+5624	-26.77	4.7±1.1	0.01 ^{+1.1} _{-0.1}
J0005-0006	-25.31	8.2±1.7	7.6±1.7
J1411+1217	-26.48	7.2±1.2	5.1±1.2
J1306+0356	-26.94	5.9±1.0	6.1±1.0
J1137+3549	-27.43	6.3±0.9	6.1±0.9
J1602+4228	-26.70	8.5±1.1	8.8±1.1
J0842+1218	-26.77	6.8±1.1	6.8±1.1
J0303-0019	-25.57	5.6±1.5	1.3 ^{+1.5} _{-1.3}
J2315-0023	-25.32	7.1±1.7	6.5±1.7
J1250+3130	-27.21	5.3±0.9	4.0±0.9
J1623+3112	-26.87	5.1±1.0	4.9±1.0
J1030+0524	-26.95	5.5±1.0	5.5±1.0
J1148+5251	-27.31	2.5±0.9	2.3±0.9

¹ QSO ID;

² absolute magnitude at the rest-frame wavelength $\lambda_{rest} = 1450 \text{ \AA}$;

^{3,4} NZ radius corrected for QSO intrinsic brightness using eq. 4.3, for M -free and M -fix respectively.

4.2 ABSORPTION FROM NEUTRAL HYDROGEN

The radiation emitted by the QSO at wavelengths shorter than the Ly α transition wavelength ($\lambda_{rest} < 1215.67 \text{ \AA}$) will be absorbed by neutral hydrogen in the intervening IGM (GP absorption). The theoretical GP optical depth for a generic Lyman transition is:

$$\tau_{GP} = \frac{\pi e^2}{m_e c} f \lambda H^{-1}(z) n_{\text{HI}}(z), \quad (4.10)$$

where f and λ are the oscillator strength and the central wavelength of the transition, $H(z)$ is the Hubble constant at redshift z , and n_{HI} is the number density of neutral hydrogen. The QSO radiation emitted at wavelengths bluer than the Ly α transition will be Ly α absorbed by IGM systems containing neutral hydrogen if they are at redshifts z_α such that:

$$(1 + z_\alpha) = \frac{(1 + z_{\text{sys}}) \lambda_{rest}}{\lambda_\alpha}, \quad (4.11)$$

where $\lambda_\alpha = 1215.67 \text{ \AA}$. In the same way, the radiation emitted at wavelengths shorter than the Ly β transition will be Ly β absorbed by neutral hydrogen at redshifts z_β such that:

$$(1 + z_\beta) = \frac{(1 + z_{\text{sys}}) \lambda_{rest}}{\lambda_\beta}, \quad (4.12)$$

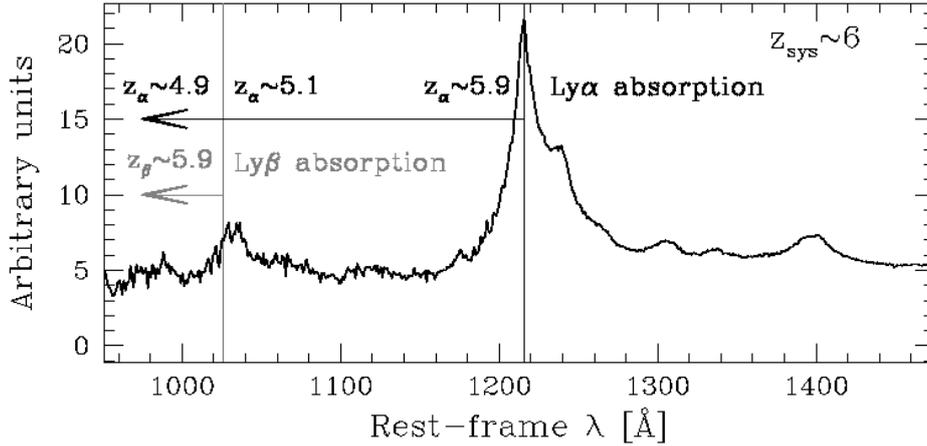
where $\lambda_\beta = 1025.72 \text{ \AA}$. For the same n_{HI} , the GP theoretical optical depth depends on the $f \lambda$ product (see eq. 4.10), resulting in a Ly β optical depth which is a factor of 6.2 lower with respect to the Ly α one. In fig. 4.5 we illustrate how the Ly α and Ly β GP absorptions are expected to affect the intrinsic spectrum of a QSO at $z_{\text{sys}} \sim 6$ in the rest frame wavelength range 980 – 1500 \AA . The QSO radiation emitted at wavelengths bluer than the Ly α transition wavelength is Ly α absorbed by neutral hydrogen at redshifts $z_\alpha < z_{\text{sys}}$. For the $z_{\text{sys}} \sim 6$ case the flux emitted just blue-wards of the Ly α transition will be absorbed by neutral hydrogen at $z_\alpha \sim 5.9$, while that emitted just red-wards of the Ly β one will be absorbed at $z_\alpha \sim 5.1$. Instead, the radiation emitted at wavelengths shorter than the Ly β transition, will be first Ly β absorbed by neutral hydrogen at $z_\beta < z_{\text{sys}}$ and subsequently Ly α absorbed by systems at $z_\alpha < z_\beta$. For the same case, the radiation emitted just below the λ_β will be first Ly β absorbed by H I at $z_\beta \sim 5.9$ and afterward Ly α absorbed by systems at $z_\alpha \sim 4.9$.

From the observational point of view, the effective GP optical depth is defined as:

$$\tau_{GP}^{eff} = -\ln(\langle f_\lambda^{obs} / f_\lambda^{em} \rangle), \quad (4.13)$$

where the average is over a certain wavelength or redshift range along the line of sight. This quantity depends on the ionization, density and UV background fields (Songaila & Cowie 2002;

Figure 4.5: $\text{Ly}\alpha$ and $\text{Ly}\beta$ GP absorptions for a QSO at $z_{\text{sys}} \sim 6$ ($\lambda_{\text{rest}} = [980 - 1500] \text{ \AA}$). Black arrow: wavelength range affected by the $\text{Ly}\alpha$ GP absorption. Grey arrow: wavelength range affected by the $\text{Ly}\beta$ GP absorption. We also indicate approximate redshifts for the $\text{Ly}\alpha$ (black) and $\text{Ly}\beta$ (grey) absorbers.



Fan et al. 2002). Fan et al. (2006b) found that while for $z_{\text{abs}} < 5.7$ the effective optical depth for $\text{Ly}\alpha$ photons can be parametrized as:

$$\tau_{GP}^{eff} = (0.85 \pm 0.06) \left(\frac{1 + z_{\text{abs}}}{5} \right)^{4.3 \pm 0.3}, \quad (4.14)$$

for $z_{\text{abs}} > 5.7$ there is an accelerated evolution that rapidly exceeds the extrapolation from lower redshifts, and an increase in the dispersion in τ_{GP}^{eff} . Assuming a photoionization and an IGM density distribution models, Fan et al. (2006b) interpreted the combination of the steep increase with redshift in the mean and in the scatter of τ_{GP}^{eff} as the reaching of the final stages of the reionization process. However Becker et al. (2007) showed that the steep rise in τ_{GP}^{eff} can be reproduced with alternative models of the density field, without having to invoke reionization. Additionally, the observed scatter in τ_{GP}^{eff} is consistent with it being driven solely by the density field (Lidz et al. 2006). Even though the driver of the τ_{GP}^{eff} evolution has not been clarified yet, we want to test if the steepening at $z_{\text{abs}} > 5.7$ is detectable while studying integral quantities, that are less dependent on the spectral resolution. In the following we outline a possible method to detect the evolution of τ_{GP}^{eff} through the analysis of the relation between the observed $\text{Ly}\alpha/\text{Ly}\beta$ flux ratios and the QSO systemic redshift. We test this method using our sub-sample of 16 optical spectra (see sec. 4.1).

4.2.1 BUILDING A MODEL

We want to build a spectral model in the wavelength range $\lambda_{rest} = [980 - 1500] \text{ \AA}$ considering the Ly α absorption due to the intervening neutral hydrogen. To do so we extrapolate the τ_{GP}^{eff} parametrization (see eq. 4.14), valid for H I absorbers at $z_{abs} < 5.7$, up to $z \sim 6.4$. This way we do not account for a possible excess at $z > 5.7$.

- **Step 1: the SDSS composite spectrum.**

In chap. 3 we have seen that for $\lambda_{rest} > 1215.67 \text{ \AA}$, the SDSS low redshift composite spectrum represents the median emitting characteristics of $z \sim 6$ QSOs very well. If we assume that this is valid also for shorter wavelengths, we can use the SDSS composite spectrum of lower redshift QSOs as the initial QSO flux in our model (see fig. 4.6, top panel). Even though individual QSO could deviate significantly from the median behavior, our model should be able to represent the global trend.

- **Step 2: Observed frame.**

For each QSO in our sample we compute the flux in the observed frame using the source systemic redshift. At $\lambda_{obs} = \lambda_{em} (1 + z_{sys})$ will correspond a flux $f_0(\lambda_{obs}) \propto f(\lambda_{em})/(1 + z_{sys})$. Since we are going to analyze flux ratios, we do not need to rescale the model fluxes to the observed ones (see fig. 4.6, central panel).

- **Step 3: Mocking absorption.**

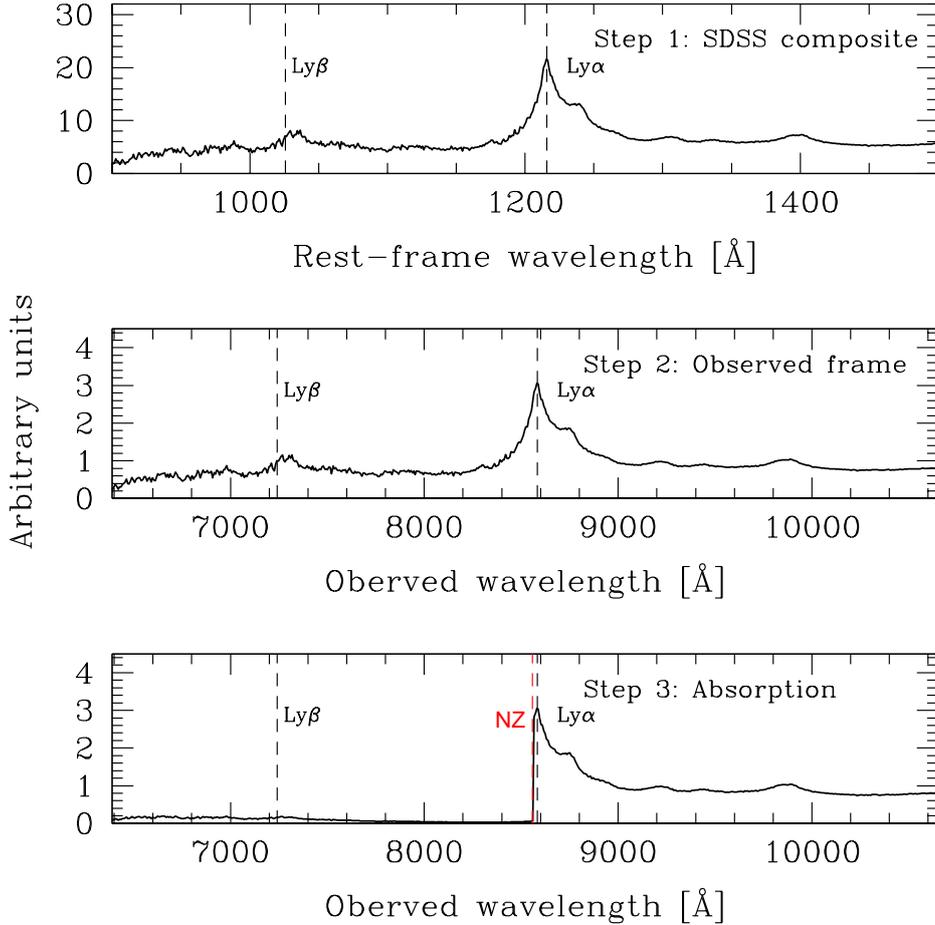
We finally mock the GP optical absorption. Extrapolating the parametrization of τ_{GP}^{eff} to redshifts higher than 5.7 (see eq. 4.14), the observed flux becomes:

$$f_{abs}(\lambda_{obs}) = f(\lambda_{obs}) e^{-\tau_{GP}^{eff}(z_\alpha)}, \quad (4.15)$$

where $z_\alpha = \frac{\lambda_{obs}}{\lambda_\alpha} - 1$. In sec. 4.1 we have seen that QSOs are able to ionize regions with typical sizes of few Mpc. These regions are characterized by an excess of radiation in the blue wing of the Ly α before the onset of the GP absorption. Hence we reproduce the H I absorption starting from the redshift z_{NZ} , which is the redshift corresponding to the NZ boundary, as estimated in sec. 4.1. We are considering the values obtained for the *M-fix* case. An example of the final model is shown in fig. 4.6, bottom panel.

Note that in this model we are not considering the additional Ly β GP absorption arising at $\lambda_{rest} < \lambda_\beta$.

Figure 4.6: Building the model. Top panel: SDSS composite spectrum. Central panel: flux in the observed frame, computed using the QSO systemic redshift. Bottom panel: flux in the observed frame after mocking $\text{Ly}\alpha$ GP absorption. The absorption starts from the redshift corresponding to the NZ boundary (red dashed line).



4.3 FLUX RATIOS

Following [Vanden Berk et al. \(2001\)](#) we define the $\text{Ly}\alpha$ range as $\text{Range}_{\text{Ly}\alpha} = [1160, 1290] \text{ \AA}$, and the $\text{Ly}\beta$ range as $\text{Range}_{\text{Ly}\beta} = [1012, 1055] \text{ \AA}$ (see fig. 4.7). The total emitted flux in $\text{Range}_{\text{Ly}\alpha}$ is thus given by the sum of the power-law continuum and of the $\text{Ly}\alpha$, N v and Si II lines. The total emitted flux in $\text{Range}_{\text{Ly}\beta}$ is equal to the sum of the continuum and of the $\text{Ly}\beta$ and O IV lines. In fig. 4.7 we illustrate how these regions are affected by the GP absorption. We can easily identify three spectral intervals:

1. $\text{Ly}\alpha_{dx} = [\lambda_{\text{NZ}}, 1290 \text{ \AA}]$, not affected by any GP absorption

2. $Ly\alpha_{sx} = [1160\text{\AA}, \lambda_{NZ}]$ and $Ly\beta_{dx} = [\lambda_\beta, 1055\text{\AA}]$, affected by $Ly\alpha$ absorption only
3. $Ly\beta_{sx} = [1012\text{\AA}, \lambda_\beta)$, first affected by $Ly\beta$ absorption and subsequently $Ly\alpha$ absorbed by neutral hydrogen sitting at $z_{abs,\alpha} < z_{abs,\beta}$

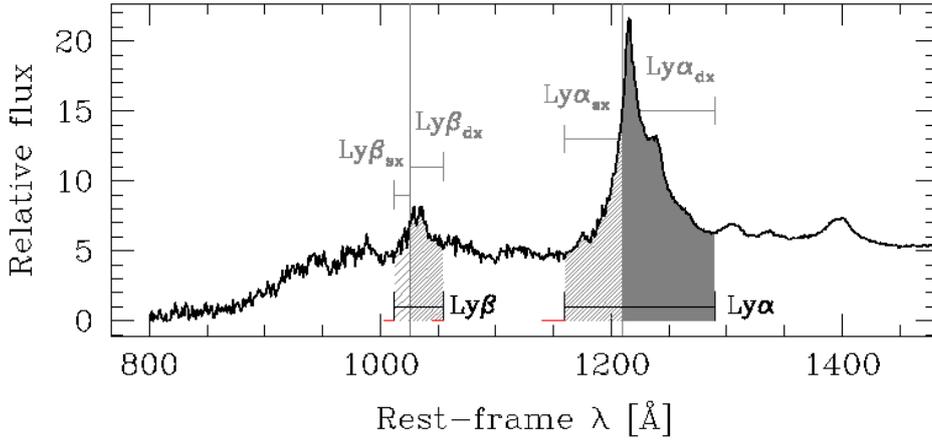
where λ_{NZ} is the wavelength corresponding to the boundary of the NZ region. Note that we are not taking into account the effects of the NZ region for the $Ly\beta$ emission since we are not simulating the $Ly\beta$ GP absorption and in the following analysis we will focus either on the $Ly\beta$ total flux or on the $Ly\beta_{dx}$ one. We numerically integrate the model and the QSO fluxes in the observed spectral regions corresponding to the $Ly\alpha_{dx}$, $Ly\alpha_{sx}$, $Ly\beta_{dx}$ and $Ly\beta_{sx}$ intervals. While measuring the integral fluxes for the models, we need to take into account that for $\lambda_{rest} < \lambda_\alpha$, the SDSS composite spectrum does not represent the median QSO emitting properties directly. The composite spectrum has in fact been obtained from spectra of QSO with $0.1 < z_{sys} < 4.8$. This means that for $\lambda_{rest} < \lambda_\alpha$ we would expect some residual GP absorption. It is not straightforward to quantify analytically the residual absorption for the considered spectral regions. Hence we choose suitable continuum windows free from line features, close to the three intervals that would potentially be affected by the absorption:

- $Ly\alpha_{sx}$: $Region_{cont1} = [1140 - 1160] \text{\AA}$;
- $Ly\beta_{dx}$: $Region_{cont2} = [1045 - 1060] \text{\AA}$;
- $Ly\beta_{sx}$: $Region_{cont3} = [1003 - 1012] \text{\AA}$;

(see horizontal red lines in fig. 4.7). We then fit the SDSS composite continuum for $\lambda_{rest} > \lambda_\alpha$ using a power-law with fixed slope $\alpha = -1.54$ (Vanden Berk et al. 2001). After extrapolating the continuum model to bluer wavelengths ($\lambda_{rest} < \lambda_\alpha$), we calculate for each continuum interval an absorption correction equal to the ratio between the extrapolated power-law and the SDSS composite flux $f_{corr} = f_{PL}/f_{SDSS}$. We then correct the $Ly\alpha_{sx}$, $Ly\beta_{dx}$ and $Ly\beta_{sx}$ fluxes computed for the models using the absorption correction obtained for the continuum. We finally obtain the total integral flux in $Region_{Ly\alpha}$ and $Region_{Ly\beta}$ as the sum of the integral fluxes of the corresponding sx and dx components.

In fig. 4.8 we show the measured $Ly\alpha/Ly\beta$ flux ratios as a function of the QSO systemic redshifts. Black points indicate our measurements while red points indicate the ratios obtained for the models. Even though the model traces very well the trend of the actual measurements, black points tend to be systematically higher than the red ones. This is due to the fact that in our model we are not accounting for any additional $Ly\beta$ absorption, resulting in an overestimate of the total $Ly\beta$ flux (lower $Ly\alpha/Ly\beta$ ratios). As we have pointed out already, strong deviations between the two samples are possible and expected (as in the case of J1411+1217). The model has in fact

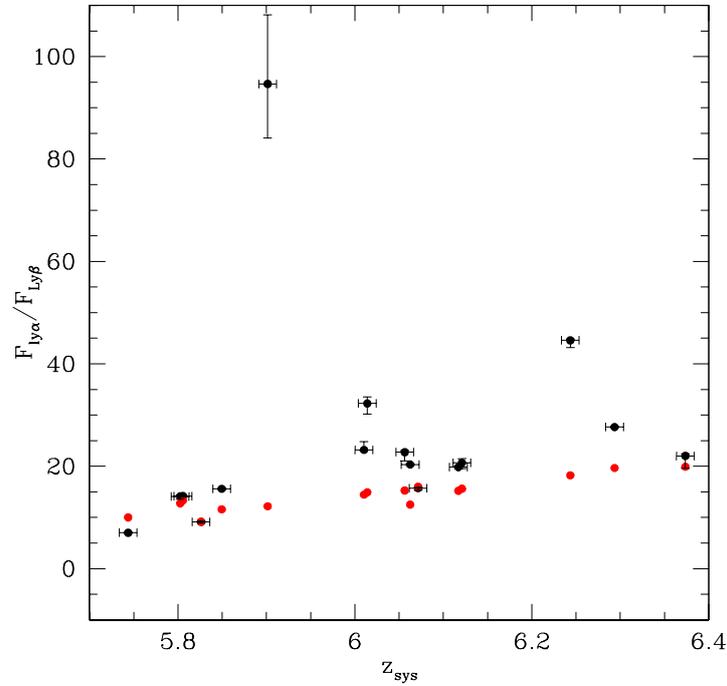
Figure 4.7: SDSS composite spectrum: GP absorption in $\text{Range}_{\text{Ly}\alpha}$ and $\text{Range}_{\text{Ly}\beta}$. We indicate with decreasing grey tones increasing GP absorption: $\text{Ly}\alpha_{dx} = [\lambda_{NZ}, 1290\text{\AA}]$, not affected by any GP absorption (dark grey); (2) $\text{Ly}\alpha_{sx} = [1160\text{\AA}, \lambda_{NZ})$ and $\text{Ly}\beta_{dx} = [\lambda_{\beta}, 1055\text{\AA}]$, affected by $\text{Ly}\alpha$ absorption only (light grey); (3) $\text{Ly}\beta_{sx} = [1012\text{\AA}, \lambda_{\beta})$, first affected by $\text{Ly}\beta$ absorption and subsequently $\text{Ly}\alpha$ absorbed by neutral hydrogen sitting at $z_{abs,\alpha} < z_{abs,\beta}$ (very light grey). The red horizontal lines indicate the continuum regions we select to correct the SDSS composite spectrum for pre-existing GP absorption.



been obtained from the simple assumption that the median emission properties of the low redshift sample represents the typical properties of the high redshift QSOs well. Hence individual sources can still be significantly different. From the analysis of the total $\text{Ly}\alpha/\text{Ly}\beta$ flux ratios, we are not able to infer much about the steepening of the effective optical depth for $z > 5.7$. If there is any detectable effect it is diluted by the total flux measurements.

If we want to detect a steepening in the evolution of τ_{GP}^{eff} we need to focus on spectral regions that are solely $\text{Ly}\alpha$ GP absorbed: i.e., the $\text{Ly}\alpha_{sx}$ and $\text{Ly}\beta_{dx}$ windows. In fig. 4.9 we show the measured $\text{Ly}\alpha_{sx}/\text{Ly}\beta_{dx}$ flux ratios as a function of the QSO systemic redshifts (black points: actual measurements; red points: models). Note that for this particular flux ratio, the correction for the pre-existing absorption in the SDSS composite spectrum is only of 7%. Also in this case, the model traces very well the actual measurements. From $z_{sys} \sim 6.1$ onwards though, the ratios measured for the model are systematically higher the ones obtained for our sources. From this systemic redshift on, the flux corresponding to the $\text{Ly}\alpha_{sx}$ spectral region is more absorbed with respect to the simple extrapolation of the optical depth parametrization valid at lower redshifts. This results in an overestimate of the $\text{Ly}\alpha_{sx}/\text{Ly}\beta_{dx}$ ratio for the models. For $z_{sys} \gtrsim 6.1$ the $\text{Ly}\alpha_{sx}$ corresponds to $\text{Ly}\alpha$ absorbers with $z_{\alpha} \gtrsim 5.7$. We thus conclude that it is possible to detect the steepening of τ_{GP}^{eff} at $z \gtrsim 5.7$ through the analysis of the relation between the QSO systemic redshift and the ratio between line fluxes, once they are measured over suitable spectral regions.

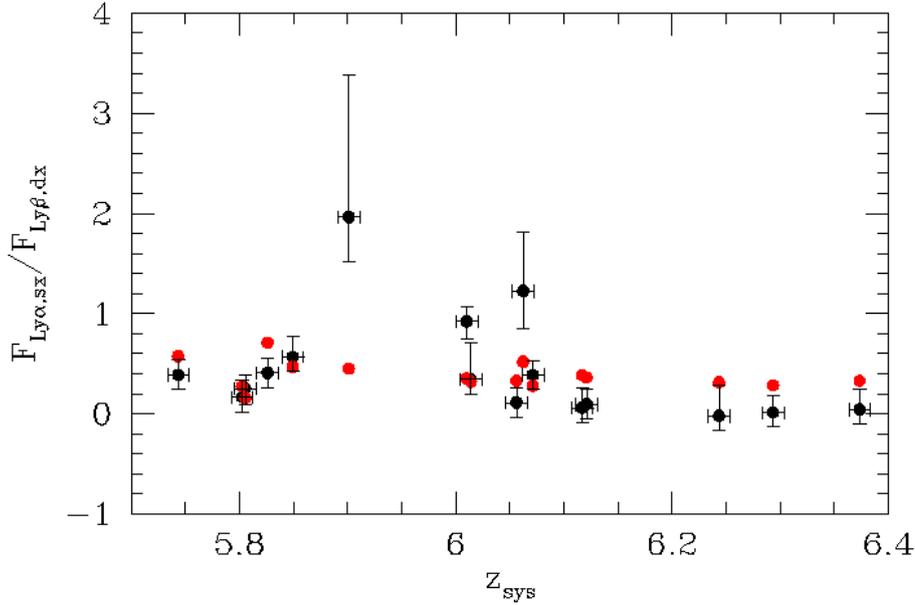
Figure 4.8: $\text{Ly}\alpha/\text{Ly}\beta$ flux ratios as a function of the QSO systemic redshifts. Black points: our measurements. Red points: ratios obtained for the models.



4.4 SUMMARY AND CONCLUSIONS

We have measured the NZ radii for 16 of the 27 $z \sim 6$ QSOs described in chap. 3. For all the sources in the sub-sample we have a reliable estimate of the QSO systemic redshift, which has been measured either from the O I+Si II (13 sources) or from the Mg II emission lines (3 objects). In the literature, the NZ radius has been arbitrarily defined as the distance at which the ratio between the observed and the emitted flux drops to 0.10 for the first time after smoothing the spectrum to a resolution of 20 \AA . We have tested the dependence of the R_{NZ} definition on the smoothing length adopted. We have found that at high redshifts ($z \gtrsim 6$) the definition is robust and does not change significantly for different choices of the smoothing kernel. At lower redshifts instead the definition of a common global smoothing kernel is not straightforward. At $z < 6$ there is an increase in the fluctuations of the IGM transmission: if the smoothing length is not properly chosen, it is easy to wash out the signature of the H II region by convolving the excess of flux in the NZ region with a transmission spike originated in the surrounding IGM. The length of the smoothing kernel at which the convolution with the IGM spikes starts, depends on the IGM structure surrounding the QSO. For comparison with previous results, we have adopted the literature definition of R_{NZ} . After correcting the measurements for the intrinsic QSO luminosities,

Figure 4.9: $Ly\alpha_{sx}/Ly\beta_{dx}$ flux ratios as a function of the QSO systemic redshifts. Black points: our measurements. Red points: ratios obtained for the models. From $z_{\text{sys}} \sim 6.1$ onwards, the ratios measured for the model are systematically higher the ones obtained for our sources, indicating additional absorption in the observations compared to the model (See text for details).



we have studied the evolution of $R_{NZ,corr}$ with the QSO systemic redshift while considering two different models for the QSO emitted flux. We have found that while performing a consistent analysis, choices of different models for the emitted flux cause variations on the $R_{NZ,corr}$ estimates that are on average lower than the typical errors. Comparing our measurements with previous results though, the degree of correlation between $R_{NZ,corr}$ and z_{sys} is significantly reduced. We thus conclude that it is not straightforward to give relative constraints on the IGM ionization state through the analysis of the relation between the $R_{NZ,corr}$ and the redshift, once the spectra are analyzed and fitted consistently.

We have analyzed the evolution of the $Ly\alpha/Ly\beta$ flux ratios as a function of the QSO systemic redshifts in order to detect possible effects due to a steep increase in the GP optical depth for $Ly\alpha$ absorbers at $z > 5.7$. We have built a model for our 16 sources, starting from the SDSS composite spectrum of lower redshift quasars. We have simulated the GP absorption extrapolating up to $z_{\text{sys}} \sim 6.4$ the optical depth parametrization obtained for lower-redshift absorbers. We have then measured the $Ly\alpha/Ly\beta$ and $Ly\alpha_{sx}/Ly\beta_{dx}$ flux ratios for the sources and for the models. From the analysis of the total $Ly\alpha/Ly\beta$ flux ratios though, we cannot infer much about the steepening of the effective optical depth. If there is any detectable effect it is diluted by the total flux measurements. We are instead able to detect it through the analysis of the relation between the $Ly\alpha_{sx}/Ly\beta_{dx}$

flux ratio and the QSO systemic redshift. From $z \sim 6.1$ onwards, the ratios measured for the models are systematically lower with respect to the ones measured for our sources. I.e., from that systemic redshift on, the Ly α line is more absorbed with respect to what we would expect from the simple extrapolation of the absorption parametrization valid at lower redshifts. It is actually possible to detect a steepening in the evolution of the effective optical depth through the analysis of the relation between the QSO systemic redshift and the ratio between line fluxes, once they are measured over suitable spectral regions. Since we are dealing with integral flux measurements, this analysis does not require high resolution spectra. A larger sample and a more detailed modeling would be needed to quantitatively test the utility of this method in constraining the evolution of τ_{GP}^{eff} .

When and how the universe reionized remains one of the fundamental question of modern cosmology. In order to fully understand the best way to use QSO spectra to robustly constrain the reionization epoch, we need to extend a consistent analysis as shown here to dataset as large as possible.

SUMMARY AND FUTURE PERSPECTIVES

THE PAST

In the past 10 years, more than 60 QSOs at redshift $5.7 < z < 6.5$ have been discovered (Fan et al. 2000, 2001, 2003, 2004, 2006; Jiang et al. 2008, 2009; Willott et al. 2007, 2010) thanks to large multivalength surveys such as SDSS and CFHQS. Only recently, the $z = 6.5$ redshift barrier has been broken and the first $z \sim 7$ QSO has been discovered in the near-infrared (NIR) UKIDSS large area survey (ULAS, Mortlock et al. 2011). High-redshift QSOs are among the most luminous sources known to date and are direct probes of the universe less than 1 Gyr after the Big Bang. They are fundamental in studying the early growth of supermassive BHs, galaxy formation and interstellar medium chemical evolution (e.g. Kauffmann & Haenelt 2000; Wyithe & Loeb 2003; Hopkins et al. 2005). Strong emission features excited by the central engine and its immediate surroundings can be used to infer properties of the powering black hole (BH) and of the circumnuclear gas. At the same time, studying the absorption systems along the line of sight of bright quasars, one can put constraints on the cosmological evolution of the intergalactic medium (IGM) ionization state and of its chemical enrichment. QSOs emission covers the entire electromagnetic spectrum. At $z \sim 6$, key diagnostic UV emission features such as the Ly α , N v, C iv and Mg ii lines are redshifted to the optical and NIR bands. Therefore, spectra covering a broad wavelength range, [7000, 24000] Å, are needed. Moreover, a consistent analysis has to be performed for statistically significant samples spanning a wide range of luminosities. Until now, only the most luminous QSOs have been studied, often one object at the time. In this thesis we have presented the most extended consistent analysis to date of optical and NIR spectra of $z \sim 6$ QSOs.

In chapter 2 we have focused on NIR spectra covering the Mg ii and Fe ii emission features, which are powerful probes of the BH mass (M_{BH}) and of the chemical enrichment of the broad

emission line region (BLR). We have presented new spectroscopic VLT-ISAAC observations of three $z \sim 6$ SDSS QSOs. The new observations extend the existing SDSS sample towards the faint end of the QSO luminosity function. We have collected 22 spectra from the literature (19 different sources) of high-redshift ($z > 4$) QSOs covering the range of interest. The final sample is composed of 22 sources: our three new sources at $z \sim 6$, 10 spectra from the literature of QSOs with $4 < z < 5.7$ and 9 spectra of QSOs with $z \sim 6$. Using a maximum likelihood fitting routine, optimized for our spectral decomposition, we have estimated M_{BH} , the Eddington ratio (defined as $L_{\text{bol}}/L_{\text{Edd}}$) and the Fe II/Mg II line ratio, a proxy for the chemical composition, to characterize both the central object and the broad line region gas.

We have found that the QSOs in our sample host BHs with masses of $\sim 10^9 M_{\odot}$. Our results are in good agreement with previous estimates, indicating that different fitting procedures produce systematic variations smaller than the errors due to the intrinsic scatter of the M_{BH} estimator. High-redshift QSOs are accreting close to the Eddington luminosity. Moreover, the distribution of Eddington ratios for the $z \sim 6$ QSOs is significantly different (average Eddington ratio ~ 8 times higher) from that of a comparison sample at $0.35 < z < 2.25$ (SDSS, DR7). Since the difference in the two Eddington ratio distributions persists, even if one considers two luminosity matched subsamples, we conclude that this result is not biased by the luminosity selection of the $z \sim 6$ QSOs: the high- z sources are accreting faster than the ones at $0.35 < z < 2.25$. Finally, we have shown that the Fe II/Mg II ratios depend strongly on the fitting method employed. From our self-consistent, homogeneous analysis the Fe II/Mg II scatter found in previous studies is significantly reduced, implying that the BLRs in the highest- z QSOs studied to date all show comparable level of chemical enrichment. The measured Fe II/Mg II line ratios show no evolution with cosmic time in the redshift range $4 < z < 6.5$. Using the Fe II/Mg II line ratio as a secondary proxy of the Fe/Mg abundance ratio, we have been able to conclude that the QSOs in our sample have undergone a major episode of Fe enrichment in the few 100s Myr preceding the cosmic age at which they are observed.

In the third chapter we have focused on the optical emission properties of $z \sim 6$ QSOs. We have assembled a sample of 26 good quality QSOs spectra covering the redshift range $z = 5.70 - 6.42$ and the wavelength range $900 \text{ \AA} < \lambda_{\text{rest}} < 1500 \text{ \AA}$, which is characterized by the Ly α complex (Ly α , N V, Si II), and the O I+Si II, C II and Si IV+O IV] broad lines. The Ly α complex is severely affected by the neutral hydrogen absorption for wavelengths shorter than the Ly α transition wavelength ($\lambda_{\text{rest}} < 1215.67 \text{ \AA}$). We have measured the power-law continuum normalization and the central wavelength, width and amplitude for each of the emission lines detected in the QSOs. For the Ly α complex, the line properties have been measured using two different ways in order to test how the derived emission properties depend on the fitting model adopted. In the first case, (*M-free*), we have considered all the Ly α parameters as free parameters, while the second case (*M-fix*), we have fixed the Ly α central wavelengths using a prior on the QSO systemic

redshift. We showed that the redshifts inferred from O I+Si II complex are consistent with those from Mg II. This allows us to use z_{OI} as a homogeneous proxy for the QSO systemic redshift, also for the objects where NIR spectra covering Mg II are not available.

Concerning the modeling, our results show that, in case of strong absorption of the Ly α line, the Ly α line flux can be significantly underestimated if one does not fix its central wavelength. In general, the properties of the Ly α complex lines are better constrained if the prior on the QSO systemic redshift is used to fix the central wavelengths. Using our estimates of the systemic redshifts we have built a median composite spectrum for the $z \sim 6$ QSOs and we have compared it with the SDSS composite spectrum obtained from QSOs with $0.1 < z < 4.8$. We have found that the two composite templates are astonishingly similar red-wards of the Ly α transition wavelength. The only exception is the Ly α + N v complex, which is slightly brighter in the $z \sim 6$ composite. We have interpreted this difference as resulting from a selection luminosity bias: luminous sources with strong Ly α lines have higher probabilities to be selected as drop-out objects. We have then analyzed the emission line shifts through the comparison of the redshifts inferred from the optical lines with the ones obtained from Mg II and C IV lines available from the literature. While the C IV line is on average blue-shifted with respect to all the measured optical lines, for the Mg II line the measured absolute offsets are correlated with the line ionization potentials. An exception to this trend is the Si IV+O IV] line: the redshift inferred from this line is in exceptional good agreement with the Mg II estimate even though it is a high ionization complex (average ionization potential of 61.28 eV).

After characterizing the line shifts, we have analyzed the QSO emission properties (continuum luminosity; EW, FWHM and luminosity of each line), looking for possible correlations among these quantities. We have not found any significant correlation between the line EW and continuum luminosity. While for the low ionization lines the absence of correlation is in agreement with the *Baldwin effect*, we were expecting an anti-correlation for the high ionization lines (Si IV+O IV] and N v), which was not found. We have therefore investigated the relation between the line and continuum luminosities, which is simply a re-formulation of the EW-continuum luminosity one. A significant correlation has been found for all the emission lines with the exception of N v. Our conclusion is that the lack of correlation for the N v line might correspond to a weak detection of the expected anti-correlation between its EW and the continuum luminosity. Nevertheless, we should point out that the N v line can be severely affected by the delicate spectral decomposition of the Ly α complex. Additionally, we have found a strong correlation between the line FWHM and its luminosity. However no significant correlation has been found between the line FWHM and the continuum luminosity. This difference is most probably arising from the fact that the range of $L(1350 \text{ \AA})$ spanned in our analysis is very narrow (half of the range spanned in line luminosity). As we have seen in the third chapter, the distribution of Eddington ratios for $z \sim 6$ QSOs is relatively narrow, therefore the correlation between the line luminosity and FWHM is

supposedly indicating that QSOs with brighter lines are the ones hosting more massive BHs. This has been further confirmed by testing the relation between the line luminosity and M_{BH} , obtained in chap. 2 from the Mg II line. We have found a significant correlation between the two quantities in all cases except for the N V (independent on the model adopted for the Ly α line) and for the Si II (only for the *M-free* model). Our results support the hypothesis that our sample is composed by QSOs with similar Eddington ratios but different BH masses, and that the line luminosity is increasing with increasing M_{BH} . However, there might be significant biases arising from the low number of objects and the small range in M_{BH} spanned by the current sample.

Finally, we have found a possible interpretation of our results through the *disk-wind* model for the BLR. In this model the BLR is a combination of a disk and a wind component, the relative importance of which depends on the actual QSO spectral energy distribution. High ionization lines can be originated in winds, while intermediate/low ionization lines are always generated in the accretion disk or at the base of the wind. From our analysis, N V is consistent with being mainly originating from the BLR wind component, while all the other emission lines seem to be mainly originating from the disk, in agreement with results for lower redshift QSOs.

In the fourth chapter we have used a sub-sample of the optical QSOs analyzed in third chapter to study the ionization state of the IGM at high redshift. We have first focused on the so called QSO *near zones* (NZs): H II regions ionized by the QSO UV photons. The evolution of the NZ radii as a function of the QSO systemic redshifts has been used in the literature to give relative constraints on the IGM ionization state. We have selected a sub-sample of 16 sources out of the 26 presented in the previous chapter, removing the weak line QSOs (WLQs), the broad absorption line QSOs (BALs) and those sources for which we do not have a reliable measurement of the systemic redshift. In the literature, the NZ radius has been observationally defined as the distance at which the ratio between the observed and the emitted flux drops to 0.10 after smoothing the spectrum to a resolution of 20 Å.

We have tested the dependence of the R_{NZ} definition on the adopted smoothing length. We have found that at high redshifts ($z \gtrsim 6$) the definition is robust and does not change significantly for different choices of the smoothing kernel, while, at lower redshifts, the definition of a common global smoothing kernel is not straightforward. At $z < 6$ there is in fact an observed increase in the fluctuations of the IGM transmission. As a consequence, if the smoothing length is not properly chosen, it is easy to wash out the signature of the H II region by convolving the excess of flux in the NZ region with a transmission spike originated in the surrounding IGM. To compare with previous results, we have the literature definition of R_{NZ} . After normalizing the measurements for the intrinsic QSO luminosities, we have studied the evolution of $R_{\text{NZ},\text{corr}}$ with the QSO systemic redshift while considering the two different models for the QSO emitted flux outlined in the third chapter (*M-free*: free Ly α central wavelengths; *M-fix*: fixed Ly α central wavelengths). We have found that while performing a consistent analysis, choices of different models for the emitted

flux cause variations on the $R_{NZ,corr}$ estimates that are on average lower than the typical errors. From our results, the degree of correlation between $R_{NZ,corr}$ and z_{sys} is significantly reduced with respect to previous studies. Hence, once the spectra are analyzed and fitted consistently, it is not straightforward to give relative constraints on the IGM ionization state through the analysis of the relation between the $R_{NZ,corr}$ and the redshift.

Secondly, we have outlined a possible method to put constraints on the Gunn Peterson (GP) effective optical depth through the analysis of the evolution of the $Ly\alpha$ and $Ly\beta$ flux ratios. From the observational point of view, the effective GP optical depth (τ_{GP}^{eff}) is defined as minus the natural logarithm of the ratio between the observed and emitted flux, averaged over a certain wavelength range. Fan et al. (2006b) found that, for $Ly\alpha$ absorbers at $z_{abs} > 5.7$, there is an accelerated evolution that rapidly exceeds the extrapolation from lower redshifts.

We have analyzed the evolution of the $Ly\alpha/Ly\beta$ flux ratios as a function of the QSO systemic redshifts in order to detect possible effects due to a steep increase in the GP optical depth for $Ly\alpha$ absorbers at $z > 5.7$. We have built a model for our 16 sources, starting from the SDSS composite spectrum of lower redshift quasars, where we have simulated the GP absorption extrapolating the optical depth parametrization obtained for lower-redshift absorbers up to $z_{sys} \sim 6.4$. We have measured the $Ly\alpha$ and $Ly\beta$ integral fluxes both over the entire line wavelength range and over suitable spectral windows that are only affected by the $Ly\alpha$ GP absorption ($Ly\alpha_{sx}$, $Ly\beta_{dx}$). We have then measured the $Ly\alpha/Ly\beta$ and $Ly\alpha_{sx}/Ly\beta_{dx}$ flux ratios for the sources and for the models. While from the analysis of the total $Ly\alpha/Ly\beta$ flux ratios, we have not been able to detect the steepening of the effective optical depth, we have been able to measure a change in τ_{GP}^{eff} through the analysis of the $Ly\alpha_{sx}/Ly\beta_{dx}$ flux ratios. Our results show that from $z \sim 6.1$ onwards, the ratios measured for the models are systematically lower compared to the ones measured for our sources. This means that, from that systemic redshift on, the $Ly\alpha$ line is more absorbed with respect to what we would expect from the simple extrapolation of the absorption parametrization valid at lower redshifts. We have concluded that it is possible to detect a steepening in the evolution of the effective optical depth through the analysis of the relation between the QSO systemic redshift and the ratio between line fluxes, once they are measured over suitable spectral regions. Thanks to this method it might be possible to constrain the evolution of the GP optical depth by analyzing large samples of low resolution spectra.

THE (NEAR) FUTURE

The principal biases in our analysis are represented by the size of the sample and by the narrow range of luminosities spanned by our sources. Even though we are starting to sample lower QSO luminosities (e.g. see chap. 2), we are still not probing the bulk of the $z \sim 6$ QSO population.

We need larger samples, spanning a wide range of luminosities if we want to prove the validity of our results for the typical $z \sim 6$ QSO, well below the currently accessible *peak of the iceberg*. Moreover, we need these samples to be observed over a wide wavelength range, covering the highest possible number of key diagnostic features, in order to enable a clean and consistent spectral decomposition and a complete characterization of each object.

THE PAN-STARRS 3 π SURVEY

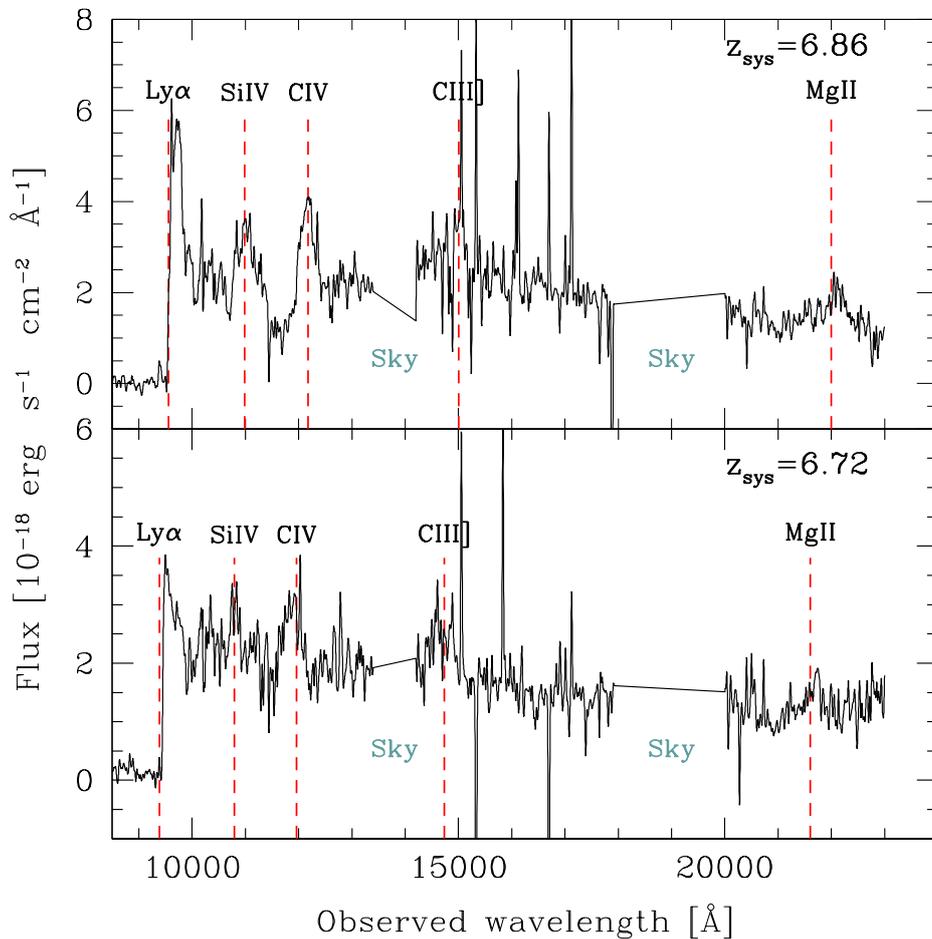
Although surveys such as SDSS have been highly successful in finding quasars at $z \sim 6$, they are limited to $z \sim 6.5$ due to the lack sensitivity at wavelengths higher than 9000 Å. Currently only few quasars at $z > 6.5$ are known (B. Venemans, private communication, [Mortlock et al. 2011](#)). The Panoramic Survey Telescope & Rapid Response System (Pan-STARRS, PS1) has started surveying the sky in the g , r , i , z and y bands. In particular, the PS1 3π survey is mapping three quarters of the sky by taking 4 exposures per year of the same area in each filter. The 3π survey represents a fundamental step forward in high- z quasar searches for the following reasons:

- **Large area:** with its 30000 deg² of sky coverage, it is significantly larger than both the SDSS DR8 (14000 deg²) and ULAS (7500 deg²) surveys. This is critical in order to build statistically significant samples of these rare objects.
- **Depth:** after three years, the survey will be significantly deeper than SDSS both in the i (~ 0.5 mag) and in the z band (~ 1 mag), that are critical to detect $z > 5.7$ QSOs ([Morganson, De Rosa et al. 2011](#)).
- **Colors:** thanks to the y band filter, it will open a new redshift window, allowing the search for sources with $6.5 < z < 7.5$.

Quasars at $5.5 < z < 7.5$ are selected as i ($z \sim 6$) or z ($z \sim 7$) band drop-out objects. I.e., they are expected to show virtually no emission in the i ($z \sim 6$) and z ($z \sim 7$) bands and to have comparatively bright detections in the y and J filters. The main contaminant to the drop-out selected candidates are asteroids and brown dwarfs. To remove false detections and contaminants, additional photometric observations are required in the i , z , y and J filters. Finally, spectroscopic follow-up is needed for ultimate confirmation. Our group has been heavily involved both in the candidate selection process and in the subsequent follow-up. We have recently published the discovery of the first PS1 QSO at $z \sim 5.73$ ([Morganson, De Rosa et al. 2011](#), see chap. 3). In the near future we expect to find ~ 100 similar i drop-out objects, that will allow us to perform a statistically significant characterization of the $z \sim 6$ QSO population. This will address two of the main shortcomings of the current analysis (sample size and luminosities probed).

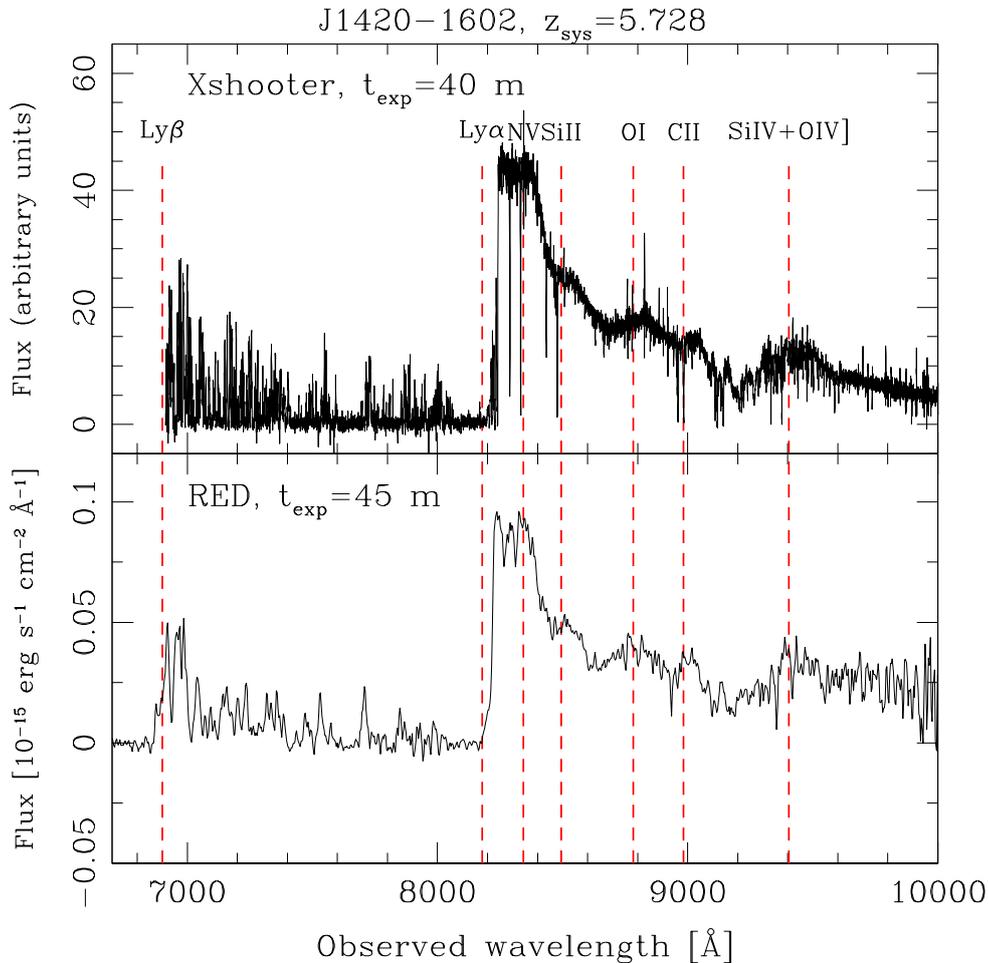
AN OBSERVATIONAL GAME CHANGER: XSHOOTER

Figure 5.1: Preliminary reduction of the Xshooter spectra (z, J, H and K bands) of two $z \sim 7$ QSOs, selected in the VIKING survey (VISTA, P.I. Venemans) observed during P87 (19th-21st of August, 2011). The observed flux is in units of $10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$ and it is smoothed over 20 \AA . The QSOs were observed for a total exposure time of only $\sim 1 \text{ h } 40 \text{ m}$ each, under non optimal weather conditions. We indicate the main emission features. A preliminary estimate of the redshifts has been obtained from the central wavelengths of the Si IV+O IV] and Mg II lines. These are two of the three highest redshift QSOs known at the time of writing this thesis.



One key goal in the study of the high redshift QSO population is to assemble observations of as many diagnostic lines as possible. In the past only few of such lines could be proved with available spectrographs. However, this situation is now changing with the recently commissioned Xshooter spectrograph at the VLT: Xshooter is a multiwavelength single object spectrograph. It consists of three arms (ultra-violet/blue, UVB; visual, VIS; near-infrared, NIR) covering in one single shot the spectral range between the U and the K band at an intermediate resolution

Figure 5.2: The first Pan-STARRS QSO J1420-1602: comparison between the visual spectra (i and z bands) obtained with Xshooter ($t_{exp} = 40$ m, top panel, preliminary reduction, observations from Aug 2011) and with the Red Channel Spectrograph mounted on MMT ($t_{exp} = 45$ m, bottom panel). The Xshooter spectrum has been binned over 1.6 \AA . Main emission features are indicated.



($R=4000-14000$, depending on the wavelength and on the slit width). Its wide wavelength coverage combined with the high sensitivities makes Xshooter the perfect spectrograph to observe high redshift QSOs. We have been awarded 3 nights of observations (visitor mode, P.I. De Rosa) during last semester (P87) and 24 h (service mode, priority A, P.I. De Rosa) for next semester (P88) to obtain coherent observations of a statistically significant sample of $z \sim 6$ QSOs covering a large range of intrinsic luminosities. Among the sources targeted during the past run (19th-21st of August, 2011), we have observed two new $z \sim 7$ QSOs selected by our collaborators in the VIKING survey (VISTA, P.I. Venemans) and our new PS1 $z \sim 6$ QSO. In fig. 5.1 we show the preliminary reduction of the Xshooter spectra obtained for the VIKING sources, in the z, J, H, and K bands, smoothed at 20 \AA . These two $z \sim 7$ QSOs were observed for a total exposure time

of only ~ 1 h 40 m each, under non optimal weather conditions. However, they still provide the first statistical analysis of the emission properties of $z \sim 7$ QSOs. In the figure we also indicate the main emission lines. A preliminary estimate of the redshifts has been obtained from the central wavelengths of the Si iv+O iv] and Mg II lines, following the techniques developed in this thesis. In fig. 5.2 we show the comparison between the visual spectra (i and z bands) of the first Pan-STARRS QSO J1420-1602, obtained with Xshooter ($t_{exp} = 40$ m, top panel) and with the Red Channel Spectrograph mounted on MMT ($t_{exp} = 45$ m, bottom panel). These spectra prove the astonishing power of Xshooter for our science. These new Xshooter observations of $z \sim 7$ QSOs pave the way to explore the next redshift frontier: now that (at least) the $z \sim 6$ QSOs are well understood, will we finally be able to detect a change in QSO properties by pushing the time back another few hundred Myr?

ACKNOWLEDGMENTS

So, it is all true: I am finally writing this last chapter. It sounds unreal, but I have made it through all of this. Being honest, I have not done it by myself. There are many people I need to thank for their support during these three long years (I hope I am not forgetting anyone).

First of all, my wonderful Bosses (with capital B): Fabian Walter (Big Boss) and Roberto Decarli (Little Boss, also known as my beloved *bossolo*). It has been a long journey, with many problems, but we have always managed to solve them together. I have learned a lot, from both of you, different key lessons that are going to be part of my (not only scientific) life. Fabian, I was not supposed to be your student, and in the end, I hope you do not regret the choice of supervising me. Thanks for being there pushing me whenever I needed it, for the patience in our daily meetings and for the constant support you gave me. I really admire your astronomical feeling: thanks for teaching me how important it is to zoom out and to look at problems from another perspective, to *feel* them before sinking into the potential well of the details. Bossolo, (oh bossolo!) you know that if I call you this way it is only to stress how important you have been during my PhD, right? You have been a mentor and a friend. I will never thank you enough for everything you have done for me (especially for being my Baiocchi dealer). Without you, this Thesis would have literally been impossible.

How can I not mention the remaining Team Quasar mates? Eric Morganson and Bram Venemans (Eduardo, you are taking my place: take good care of the old ones.. and of my baby-mac). Eric, I will never forget our days in La Silla, looking for the first PS1 QSOs while throwing up our souls. And finally the happiness of finding the first object, now THAT has been cool! A special thanks goes to Bram, who helped me a lot while I was feeling lost at the VLT, observing with Xshooter for the first time. Thanks for all your detailed emails, they have been really precious!

A big thanks to my (ex)MPIA collaborators: Anna Pasquali and Jaron Kurk. Zia Anna, what can I say? You have been there when I was alone. Thanks for sitting with me for hours, reducing by hand the first Xshooter spectra: I have had a great time, and I have really missed you here at the institute. You are a great teacher. Thanks to Jaron, for offering me the possibility of a PhD Plan-B, and for helping me out with the first analysis of the NIR data.

I am sincerely in debt with my collaborators at the Steward Observatory (Arizona). Thanks to Xiaohui Fan for his support during these three years: without his contribution this Thesis would have not been possible. Thanks for hosting me in Tucson: I have always felt very welcome! Thanks to Ian McGreer for the very useful scientific conversations, for helping me out in many occasions.. and for being a patient teacher while answering my naive questions.

Thanks to Hans-Walter Rix and Jochen Heidt for being the correctors of this Thesis, to Eva Grebel and Matthias Bartelmann for being part of my Thesis committee (I hope I will be still able to talk after the 1st of December ;-).

To properly conclude this first part, I would really like to thank the Pan-STARRS project: its delayed start allowed me to learn much more than I was initially supposed to, and it helped me finding my own way in science and life.

And now, let's get to my (real) life, let's get to the people that were with me during these three years, that helped me out during the tough periods and laughed with me during the happy moments, the people that love me even if I am a drama queen, even if I am Gisella-Alexandra even if I am me. You are really too many, but *in for a penny, in for a pound!* Let's get it started!

My wonderful crazy lost IMPRS generation (plus adopted ones)! We have had a lot of fun together! Impossible to mention all the great moments, the laughs, the tears (oh yeah, I remember many of them), the parties, the seminars, the after-seminars.. the after-after-seminars ;-). A special thanks goes to Ana (I'll never forget our first night drive to Koenigstuhl) and the adopted Emanuel, Alex H. and his (and our) beloved Maria, Alex K. and his car (yes Alex), Carolina and Uli, Dading and his camera, Julian (green furry for everyone!), Milica and her happy ending, Oliver, a wonderful dancer and the most patient office-mate, my very dark very German friends Roman and Anika. A special special thanks to the people who had to stand me the most: my wise Indian friend Bhargav, my tough (but sweet-hearted) Danish friend Kasper, and Mauricio, the most funny-ironic (Chilean) person I know. It has (almost) been a soap-opera, but I am glad I have shared it with you guys. And don't forget: the day I will be able to beat you all at table soccer is soon to come!

Thanks to the all the closest MPIA students and Post-Docs: in three years I have had the possibility to meet so many new interesting people that is difficult to mention all of them! A special thanks goes to Somayeh and her smiley face, to my wonder-Tessel and her morning coffees (I will always be there for you) and to the pure-Brit-Rory.

How to forget the Italian crew!!! Ale (how is your husband doing?), *βiβi* (sorry girl, you have definitely been adopted, take care of my bossolo), Cla, Den, Fede, Fiola, Giovi, Giua' (gna 'sti, 'sti bbone?), Giulia, Sara (with whom will you go relaxing in the princess venue after I leave?), Tere, Vivi and the new-entry Agnese! It has been great to have the possibility to meet you all and

to share sooooo much with you. My dear flatmates, my dear dirty Italian people, I will always remember our Sunday morning waking-up, our dinners, our laughs, Giovi's sweetness, Cla's pigeons, Vivi's Waka-Waka and Tere's energy. I will miss you all. Den & Fiola, this time I have betrayed you, I am not following your path! Now, this is a tragedy: the kid is leaving the nest! Thanks for being there for me, and helping me out both in Trieste and here in Heidelberg. I know I'll miss all of this Fiola, I know I will regret leaving, I know parents are always right ;-) Fede, it has been awesome to find you. There are not many words to say it, just: thanks for being there with me. You know how much I am going to miss you, but please, don't be sad! I am sure we will be in touch for our entire life!

Thanks to my Sushi-girls: Aina, Judith and Yara, and our happy chitchats about.. the beloved Martì!! OK, Martì, not always about you ;-) Thank you all for helping me out in one of the worst moments of my life, thanks for helping Gisella-Alexandra to stop crying, and for letting me have so much fun! We need to have our sushi night at least once again, maybe in the States since both me and Judith will be there?! And thanks to the beloved Martì for our wonderful ironing night, for the constant *GimMeFive* and the great swing performances!

Thanks to all my friends from the University: Anna, Beppe, Cima, Fiore, Flake, Eugi, Giambo, Gian, Kerlo, Las, Sara, Seba, Urpi.. et al. It is amazing how easy can be to stay in touch in the Internet era :-D Thanks all for the endless mailing lists about nothing, and for being part of my life even when we are physically spread all over Europe. Next frontier: the World!

Thanks to my family, the entire large-one. Mamma, Papa', Gio-gio, spero riusciate a venire prima o poi! Grazie per avermi supportato e sopportato, so quanto io possa essere difficile. And thanks to my 20 years old cat Gigia, for having been with me all this time.

Finally, because I hope he is going to be my happy-ending, thanks to Mario, my love. I have loved you as a friend, I have hated you as my best-friend, and now I love you as my life-mate. You know that I am thanking Pan-STARRS and uncle Max (Plank) everyday for giving me the possibility of finding you. I believe we will manage, I am sure we will.

It is difficult to say what is impossible, for the dream of yesterday is the hope of today and the reality of tomorrow.

(Robert H. Goddard)

BIBLIOGRAPHY

- Antonucci, R., 1993, *ARA&A*, 31, 473
- Baldwin, J. A., 1977, *ApJ*, 214, 679
- Baldwin, J. A., Ferland, G. J., Korista, K. T., Hamann, F., LaCluyzé, A., 2004, *ApJ*, 615, 610
- Barkana, R., Loeb, A., 2003, *Nat*, 421, 341
- Barth, A. J., Martini, P., Nelson, C. H., Ho, L. C., 2003, *ApJ*, 594L, 95
- Baskin, A., Laor, A., 2004, *MNRAS*, 350L, 31
- Becker, R. H., et al., 2001, *AJ*, 122, 2850
- Becker, G. D., Rauch, M. & Sargent, W. L. W., 2007, *ApJ*, 662, 72
- Bertoldi, F., et al., 2003b, *A&A*, 409, L47
- Bolton, J., & Haenelt, M., 2007a, *MNRAS*, 374, 493
- Bolton, J., & Haenelt, M., 2007b, *MNRAS*, 381, L35
- Bolton, J., et al., 2010, *MNRAS*,
- Boroson, T. A., 2005, *AJ*, 130, 381
- Boroson, T. A., & Green, R. F., 1992, *ApJS*, 80, 109
- Cardelli, J. A., Clayton, G. C., & Mathis, J. S., 1989, *ApJ*, 345, 245
- Carilli, C., et al., 2010, *ApJ*, 714, 834
- Cen, R., & Haimann, Z., 2000, *ApJ*, 542, L75
- Decarli, R., Dotti, M., Treves, A., 2011, *MNRAS*, 413, 39
- Decarli, R., Falomo, R., Treves, A., Labita, M., Kotilainen, J. K., Scarpa, R., 2010, *MNRAS*, 402, 2453
- Decarli, R., Labita, M., Treves, A., Falomo, R., 2008, *MNRAS*, 387, 1237
- De Rosa, G. et al., 2011, *ApJ*,

- Diamond-Stanic, A. M., et al., 2009, *ApJ*, 699, 782
- Dietrich, M., Hamann, F., Appenzeller, I., Vestergaard, M., 2003, *ApJ*, 596, 817
- Dietrich, M., et al., 2002, *ApJ*, 581, 912
- Fan, X., et al., 2000, *AJ*, 120, 1167
- Fan, X., et al., 2001, *AJ*, 122, 2833
- Fan, X., et al., 2002, *AJ*, 123, 1247
- Fan, X., et al., 2003, *AJ*, 125, 1649
- Fan, X., et al., 2004, *AJ*, 128, 515
- Fan, X., et al., 2006, *AJ*, 131, 1203
- Fan, X., et al., 2006, *AJ*, 132, 117
- Freudling, W., Corbin, M. R., Korista, K. T., 2003, *ApJ*, 587L, 67
- Espey, B. R., Andreadis, S., 1999, in *ASP Conf. Ser. 162, Quasars and Cosmology*, ed. G.J. Ferland & J. A. Baldwin (San Francisco: ASP), 351
- Gaskell, C. M., 1982, *ApJ*, 263, 79
- Goto, T., 2006, *MNRAS*, 371, 769
- Grandi, S. A., 1982, *ApJ*, 255, 25
- Green, P. J., et al., *ApJ*, 450, 51
- Gunn, J. E., & Peterson, B. A., 1965, *ApJ*, 142 1633
- Haimann, Z., 2002, *ApJ*, 576, L1
- Hamann, F., Ferland, G., 1993, *ApJ*, 418, 11
- Hamann, F., Korista, K. T., Ferland, G. J., Warner, C., Baldwin, J., 2002, *ApJ*, 564, 592
- Heger, A., Woosley, S. E., 2002, *ApJ*, 567, 532
- Hogg, D., 1999, arXiv: astro-ph/9905116
- Hopkins, P. F., Hernquist, L., Cox, T. J., Di Matteo, T., Martini, P., Robertson, B., Springel, V., 2005, *ApJ*, 630, 705
- Hryniewicz, K., Czerny, B., Nikolajuk, M., & Kuraszkiewicz, J., 2010, *MNRAS*, 404, 2028
- Iwamuro, F., Kimura, M., Eto, S., Maihara, T., Motohara, K., Yoshii, Y., Doi, M., 2004, *ApJ*, 614, 69
- Iwamuro, F., Motohara, K., Maihara, T., Kimura, M., Yoshii, Y., Doi, M., 2002, *ApJ*, 565, 63
- Jiang, L., et al., 2007, *AJ*, 134, 1150
- Jiang, L., et al., 2008, *AJ*, 135, 1057
- Jiang, L., et al., 2009, *AJ*, 138, 305
- Just, D. W., et al., 2007, *ApJ*, 665, 1004
- Kaspi, S., Smith, P. S., Netzer, H., Maoz, D., Jannuzi, B. T., Giveon, U., 2000, *ApJ*, 533, 631
- Kauffmann, G., Haehnelt, M., 2000, *MNRAS*, 311, 576

-
- Kelly, B. C., Vestergaard, M., Fan, X., Hopkins, P., Hernquist, L., Siemiginowska, A., 2010, *ApJ*, 719, 1315
- Komatsu, E., et al., 2010 *ApJS*, in press (arXiv: 1003.6097v2)
- Kramer, R. H., Haimann, Z., 2009, *MNRAS*, 400, 1493
- Kurk J. D., et al., 2007, *ApJ*, 669, 32
- Kurk, J. D., Walter, F., Fan, X., Jiang, L., Jester, S., Rix, H. -W., Riechers, D. A., 2009, *ApJ*, 702, 833
- Labita, M., Decarli, R., Treves, A., & Falomo, A., 2009, *MNRAS*, 399, L2099
- Laor, A., *New Astron. Rev.*, 44, 503
- Lawrence, A., et al., 2007, *MNRAS*, 379, 1599
- Leighly, K. M., et al., 2004, *ApJ*, 611, 125
- Leighly, K. M., et al., 2007, *ApJ*, 663, 103
- Leighly, K. M., Halpern, J. P., Jenkins, E. B., & Casebeer, D., 2007, *ApJS*, 173, 1
- Lidz, A., Oh, S. P., & Furlanetto, S. R., 2006, *ApJ*, 639, L47
- Madau, P., Hardt, F., & Rees, M. J., 1999, *ApJ*, 514, 648
- Madau, P., & Rees, M. J., 2000, *ApJ*, 542, L69
- Maiolino, R., Juarez, Y., Mujica, R., Nagar, N. M., Oliva, E., 2003, *ApJ*, 596L, 155
- Maiolino, R., Mannucci, F., Baffa, C., Gennari, S., Oliva, E., 2001, *A&A*, 372L, 5
- Maselli, A., Ferrara, A., & Gallerani, S., *MNRAS*, 395, 1925
- Maselli, A., Gallerani, S., Ferrara, A., & Choudhury, T., *MNRAS*, 376, L34
- Matteucci, F., Recchi, S., 2001, *ApJ*, 558, 351
- McGill, K. L., Woo, J. -H., Treu, T., Malkan, M. A., 2008, *ApJ*, 673, 703
- McGreer, I., Mesinger, A., Fan, X., 2011, *MNRAS*, 415, 3237
- McLure, R. J., Dunlop, J. S., 2004, *MNRAS*, 352, 1390
- McLure, R. J., Jarvis, M. J., 2002, *MNRAS*, 337, 109
- Mesinger, A., Haimann, Z., 2004, *ApJ*, 661, L69
- Mesinger, A., Haimann, Z., 2007, *ApJ*, 660, 923
- Morganson, E., et al., 2011, arXiv: 1109.6241M
- Mortlock, D. J., et al., 2011, *Nat*, 474, 616
- Murray, N., Chiang, J., Grossman, S. A., & Voigt, G. M., 1995, *ApJ*, 451, 498
- Nagao, T., Maiolino, R., Marconi, A., 2006, *A&A*, 447, 863
- Peebles, P.J.E., 1993, *Principles of Physical Cosmology*, Princeton University Press
- Peterson, B. M., 2004, in *IAU Symp*, 222, 15P
- Peterson, B. M., 2010, in *IAU Symp.*, Co-Evolution of Central Black Holes and Galaxies, ed. B. Peterson, R. Somerville, & T. Storchi-Bergmann (Cambridge: Cambridge Univ. Press), 151

- Pentericci, L., Fan, X., Rix, H.-W., Strauss, M. A., Narayanan, V. K., Richards, G. T., Schneider, D. P., Krolik, J., 2002, *AJ*, 123, 2151
- Pickles, A. J., 1998, *PASP*, 110, 863
- Rees, M. J., 1984, *ARA&A*, 22, 471R
- Richards, G. T., et al., 2002, *AJ*, 124, 1
- Richards, G. T., et al., 2011, *AJ*, 141, 167
- Riechers, D., et al., 2009, *ApJ*, 703, 1338
- Ryan-Weber, E. V., Pettini, M., Madau, P., & Zych, B. J., 2009, *MNRAS*, 395, 1476
- Schlegel, D. J., Finkbeiner, D. P. & Davis, M. 1998, *ApJ*, 500, 525
- Shang, Z., et al., *ApJ*, 586, 52
- Shang, Z., Wills, B. J., Wills, D. & Brotherton, M. S., 2007, *AJ*, 134, 294
- Shapiro, P. R., & Giroux, M. L., 1987, *ApJ*, 321, L107
- Shemmer, O., et al., 2010, *ApJ*, L722, 152
- Shen, Y., et al., 2011, *ApJS*, 194, 45
- Shen, Y., Greene, J. E., Strauss, M. A., Richards, G. T., Schneider, D. P., 2008, *ApJ*, 680, 169
- Sigut, T. A. A., Pradhan, A. K., 2003, *ApJS*, 145, 15
- Songaila, A. & Cowie, L. L., 2002, *ApJ*, 123, 2183
- Spergel, D. N., et al., 2007, *ApJS*, 170, 377
- Steffen, A. T., et al., 2006, *AJ*, 131, 2826
- Trac, H., & Gnedin, N. Y., 2009, in press (arXiv: 0906.4348)
- Trakhtenbrot, B., Netzer, H., Lira, P., & Shemmer, O. 2011, *ApJ*, 730, 7
- Urry, M., & Padovani, P., 1995, *PASP*, 107, 803
- Vasudevan, R. V., & Fabian, A. C., 2007, *MNRAS*, 381, 1235
- Vanden Berk, D. E., et al., 2001, *AJ*, 122, 549
- Venkatesan, A., Schneider, R., Ferrara, A., 2004, *MNRAS*, 349L, 43
- Vestergaard, M., Wilkes, B. J., 2001, *ApJS*, 134, 1
- Volonteri, M., 2010, *A&ARv*, 18, 279
- Volonteri, M., Rees, M. J., 2005, *ApJ*, 633, 624
- Walter, F., et al., 2004, *ApJ*, 615, L17
- Wang, R., et al., 2010, *ApJ*, 714, 699
- Wilkes, B. J., 1984, *MNRAS*, 207, 73
- Willott, C. J., et al., 2007, *AJ*, 134, 2435
- Willott, C. J., et al., 2010, *AJ*, 140, 546
- Willott, C. J., Delorme P., Reyl  C., Albert L., Bergeron, J., Crampton, D., Delfosse, X., Forveille, T., 2010, *AJ*, 139, 906

- Willott, C. J., McLure, R. J., Jarvis, M. J., 2003, *ApJ*, 587L, 15
- Wills, B. J., et al., 1999, *ApJ*, 515, L53
- Wills, B. J., Netzer, H., Wills, D., 1985, *ApJ*, 288, 94
- Wyithe, J. S. B., Bolton, J., & Haenelt, M., 2008, *MNRAS*, 383, 691
- Wyithe, J. S. B., Loeb, A., 2003, *ApJ*, 595, 614
- Wyithe, J. S. B., & Loeb, A., 2004, *Nature*, 427, 815
- Wyithe, J. S. B., Loeb, A., & Carilli, C., 2005, *ApJ*, 528, 575
- Xu, Y., et al., 2008, *MNRAS*, 389, 1703
- York, D. G., et al., 2000, *AJ*, 120, 1579
- Yu, Q., & Lu, Y., 2005, *ApJ*, 620, 31