



MASTER THESIS
30 ECTS

Where Is The Galaxy Counterpart of DLA2225+0527?

Johannes Møller Astrup

DARK COSMOLOGY CENTRE
NIELS BOHR INSTITUTE

Supervisor:
Johan Peter Uldall Fynbo

Submitted:
December 30, 2015

Contents

1	Introduction	1
2	Background	2
2.1	The Lyman- α line	2
2.2	Quasi-Stellar Objects	2
2.3	QSO absorbers	4
2.3.1	The Gunn-Peterson trough	4
2.3.2	Lyman- α forest absorbers	5
2.3.3	Lyman limit systems	5
2.3.4	Damped Lyman- α absorption systems	6
2.4	Galaxy counterparts	7
2.4.1	Lyman break galaxies	8
2.4.2	Metallicity-luminosity relations	9
2.4.3	Dust attenuation	10
2.5	DLA2225+0527	11
3	Observations and data reduction	12
3.1	Imaging	12
3.1.1	Basic data processing	12
3.1.2	Bias frames	13
3.1.3	Flat fields	14
3.1.4	Processing science frames	15
3.1.5	PSF-subtraction	16
3.2	Spectroscopy	18
3.2.1	Initial setup	19
3.2.2	SPSF-subtraction	19
3.2.3	Flux limits	20
4	Results	22
4.1	Imaging	22
4.2	Spectroscopy	25
5	Discussion	29
	References	31
	Appendices	35
A	Python code pipeline	35
B	Star formation rates with H α and Ly α	40

Abstract

I search for the galaxy counterpart of a solar metallicity damped Lyman- α absorber (DLA) found at $z = 2.13$ in the sightline towards the quasar QSO225+0527. Imaging data from Nordic Optical Telescope (NOT) and the Keck Telescope at Hawaii are combined with spectroscopic data from the X-shooter spectrograph at NOT. PSF-subtraction of the imaging data reveals a bright object composed of two disk-like structures so far undescribed in the literature. The object is not seen in the spectroscopic data, and thus redshift-determination is impossible with the existing data. Without an established redshift, it is not possible to confirm whether or not the object is the DLA galaxy counterpart. However, assuming that the object lies at the redshift of the absorber, magnitudes, physical size, and star formation rates are derived. It is concluded that these physical parameters, as well as the elongated morphology of the object, do not contradict with the hypothesis that the object is the searched for DLA emission counterpart.

1 Introduction

Damped Lyman- α absorbers (DLAs) are a class of absorbers seen in the spectra of bright, distant background sources typically high-redshift quasars or gamma-ray bursts. They have been discussed in the literature for several decades, but even today their nature is not fully understood (see Møller and Warren (1993) for an early presentation). It is widely believed that the absorption occurs in large slabs of neutral hydrogen representing stages of disk galaxy formation in the early Universe. This makes DLAs valuable as probes of forming galaxies and the corresponding reservoirs of gas in the high-redshift Universe.

DLAs account for the vast majority of neutral hydrogen at high redshifts, and are therefore also assumed to be the birthplace of stars and galaxies. However, previous attempts to connect high-redshift DLAs with the corresponding high-redshift galaxies which have been observed in the ultra-violet and optical regime (Lyman break galaxies), have not yet been successful. Together they form two populations of high- z galaxies; one population observed in absorption and one population observed in emission. The main obstacle in connecting the two groups is the detection of DLAs in emission, as DLA galaxy counterparts are very hard to observe. This problem is possibly connected to the fact that the more luminous DLAs are also the most chemically evolved, and therefore likely the ones to contain most obscuring dust.

This project is dedicated to the DLA DLA2225+0527 discovered by Krogager et al. (2015). The system contains high abundance of dust, and is also highly chemically evolved. The DLA has only been observed in its absorption spectrum. The purpose of my project is to search for emission from the galaxy-counterpart both through spectroscopy and imaging. A detection of the emitting part would shed light on the connection between metallicity, luminosity, and dust. It would also be the first direct detection of a solar-like DLA.

2 Background

In order to clarify the methods of this study, I will present the basic applied astrophysical concepts. The presentation is followed by a brief summary of the discovery history and science of damped Lyman- α systems and galaxies in the early Universe leading up to the recent papers on the quasar in question. The overview serves as a framework and motivation for my research.

2.1 The Lyman- α line

90 % of the baryonic matter in the Universe is in the form of hydrogen. Any astronomer working in the field of spectroscopy will devote a great deal of working hours to analyzing the spectrum of hydrogen. One of the strongest lines of neutral hydrogen is the Lyman α -line belonging to the Lyman series. An illustration of the Lyman spectrum is shown in figure 1. The lines of the Lyman series correspond to electron transitions between any excited state and the ground state. Transitions between the ground state and the first excited state correspond to the Lyman- α with the wavelength 1216 Å. As the rest of the Lyman spectrum, this line is in the ultraviolet regime. The Lyman limit absorption line at 912 Å corresponds to the energy threshold for a complete ionization of a hydrogen atom in the ground state.

Lyman- α photons are mostly produced in settings where a source of photons with energy above the Lyman limit is available. If neutral hydrogen atoms are also present, these can be ionized. When recombining, the atoms will decay through a cascade of transitions emitting various photons including Lyman- α in most of the cases. Star forming regions are typically strong emitters of Lyman- α .

With 'Lyman- α absorption' is generally meant a process in which radiation with the wavelength 1216 Å passes through a cloud with some content of neutral hydrogen. A hydrogen atom in the ground state can interact with a Lyman- α photon (Ly α photon) resulting in excitation of the atom from the ground state to the first excited state. At some later time, the excited atom will decay to the ground state emitting a new Ly α photon. This new photon will be emitted in a random direction. The net effect is that the flux of Ly α photons in the incident direction is reduced. The cross section area for neutral hydrogen to interact with Ly α is relatively large, meaning that neutral hydrogen is an efficient absorber of Ly α . Even a small HI-abundance will leave absorption features on incident Ly α radiation making the Ly α a powerful tracer of neutral hydrogen in the Universe.

2.2 Quasi-Stellar Objects¹

The two terms 'quasar' (Quasi-Stellar Radio Source) and 'QSO' (Quasi-Stellar Object) are today used without distinction. Historically, the two terms categorized the same type of

¹The following section is based on Mo et al. (2010, p. 618-624)

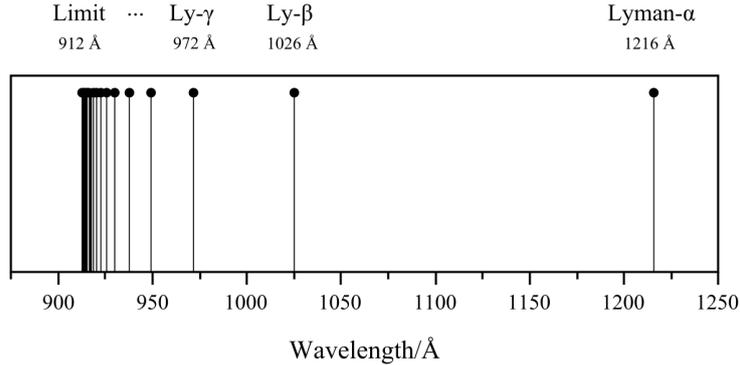


FIGURE 1: An illustration of the Lyman series - the first series of hydrogen emission lines. From <https://commons.wikimedia.org/wiki/File:LymanSeries.svg>

object into a radio-loud and a radio-quiet class with the latter being the far most numerous. I will mainly use the term 'QSO' for the remainder of this paper.

QSOs belong to the broader class 'Active Galaxy Nuclei' (AGN). AGN are defined to have one or more of the following properties:

- A compact nucleus which is much brighter than a similar region in a normal galaxy
- A non-thermal emission continuum with the full wavelength range from radio to X-ray
- Strong emission lines
- Variability on relatively short time scales

Being the most luminous types of AGN, QSOs normally completely outshine their host galaxies. They are differentiated from the similar, but less luminous, AGN-class 'Seyfert 1 galaxies' by the magnitude limit $M_B \lesssim -21.5 + 5 \log h$. QSOs are confined to very compact regions, which are normally optically unresolved at $\theta \sim 1''$. Having unusually blue colors, QSOs are mainly detected in optical using a color-selection technique. This has resulted in the detection of a majority of intrinsically bright and highly redshifted objects. Figure 2 shows a composite 2D spectrum of QSOs, which clearly shows strong emission lines and a continuum that extends blueward of 1000 Å.

The most widely accepted paradigm for explaining the large energy output generated within the compact regions of QSOs is based on the hypothesized existence of supermassive black holes (SMBHs) in the cores of galaxies. The energy is produced as a result of matter accreting onto the black hole. A simple spherical model (Mo et al., 2010, p. 624) will yield that a $\sim 10^8 M_\odot$ black hole can power a $L \sim 10^{46} \text{ erg s}^{-1}$ AGN. The corresponding mass accretion rate is $\dot{M} \sim 2 M_\odot \text{ yr}^{-1}$. Although successful as a theoretical explanation, the SMBH-paradigm is still just at hypothesis since such an accretion process has never been observed. However, strong evidence for the presence of a supermassive object in the center of the Milky Way has been provided by measuring the kinematics of individual stars in the inner region. An object with a density of $> 3 \times 10^{19} M_\odot \text{ pc}^{-3}$, centered on the radio source Sgr A*, has been proven to exist. No other astrophysical objects with such extreme physical conditions are known, and the theoretical supermassive black hole is the best candidate. Similar objects have been observed in nearby galaxies, but not with such impressive certainty as within our own galaxy. Unless the Milky Way is a very special place, many ellipticals should harbor a similar SMBH-like object that could possibly serve as engine for a QSO. It has been theorized that all galaxies go through one or several

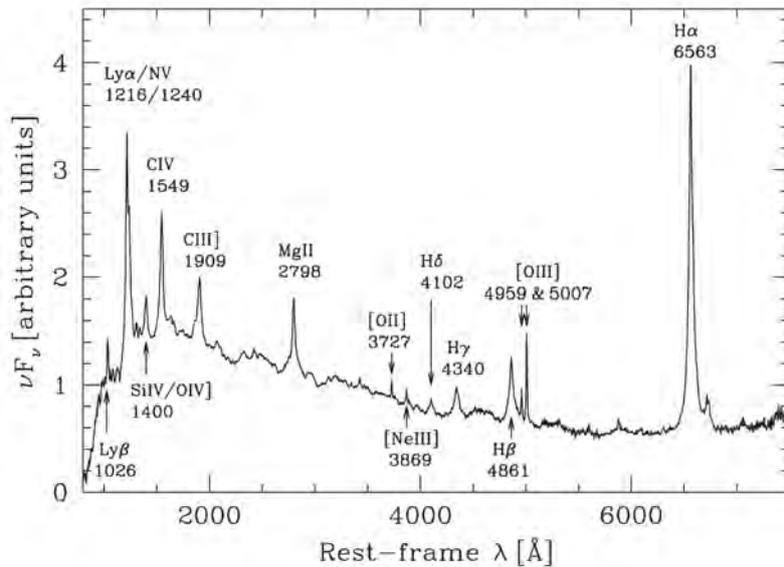


FIGURE 2: A composite 2D spectrum of a QSO. The continuum extends bluewards of the thermal regime as is typical of AGN. Strong emission lines are clearly present. Figure from Mo et al. (2010, p. 620)

AGN-phases. In this line of thought, AGN form a natural part of the evolution of normal galaxies.

2.3 QSO absorbers

The term 'QSO absorber' generally refers to any absorbing cloud of gas that lies along the sightline towards a QSO. Assuming that the gas is mostly hydrogen, the most prominent absorption line will be Ly α . Since the absorbing system will lie at a lower redshift than the background QSO, the Lyman absorption lines will be shifted bluewards of the Lyman lines in the QSO spectrum. The light from one distant source may likely travel through multiple gas clouds with different z -values, all of them leaving different imprints along the source spectrum. The spectrum becomes a probe of the gas density along the sightline, and can be used for mapping hydrogen clouds in the early Universe.²

QSO absorption systems are classified according to their hydrogen column densities. I will devote a few sections on presenting the main classes of QSO absorption systems along with a brief history of their prediction and discovery.

2.3.1 The Gunn-Peterson trough

In 1965, several independent studies predicted that in an expanding Universe, homogeneously filled with gas, the continuous redshifting of the Lyman- α line would produce an absorption trough bluewards of the Ly α emission line of any distant QSO (Gunn and Peterson, 1965; Scheuer, 1965; Shklovskii, 1965). As described previously, neutral hydrogen is a very efficient absorber of Lyman radiation. We know from classic cosmology and from observing the Cosmic Microwave Background (CMB), that the gas content of the Universe shifted phase from ionized to neutral roughly 380,000 years after the Big Bang in the cosmological epoch referred to as 'Recombination'. The astronomers in 1965 therefore expected the majority of the intergalactic hydrogen to be neutral. Light travelling from distant sources would traverse vast amounts of neutral hydrogen effectively absorbing

²Gamma-Ray Bursts are also used as illuminating sources for probing the gas content of the Universe. The physical mechanisms and spectrographic methods are essentially the same as with QSO absorbers. However, since this paper is devoted to a QSO absorber I will not discuss Gamma-Ray Bursts any further.

wavelengths smaller than the Ly α . This feature is known as the Gunn-Peterson trough. In high- z spectra the Gunn-Peterson trough should appear bluewards of the Lyman regime since the absorbing clouds lies at lower redshifts than the source. However, going towards smaller and smaller redshifts the Gunn-Peterson trough should work its way redwards towards the Ly α , and eventually 'eating' it up. Gunn and Peterson studied the spectrum of a QSO with redshift $z = 2.01$ catalogued as 3C 93C, and found such a spectral region with reduced flux. The authors used this finding to calculate an upper limit for the density of intergalactic neutral hydrogen. To their surprise, this limit turned out to be much lower than expected. Gunn and Peterson solved this problem by suggesting the - at that time rather radical idea - that the major part of the intergalactic hydrogen is ionized. Today, we know this to be true. The first perfect Gunn-Peterson trough was observed as late as 2001 by Becker et al. (2001) who observed the feature in a $z = 6.28$ QSO. The absence of Gunn-Peterson-troughs in spectra of QSOs for $z < 6$ is a strong indicator that the intergalactic medium of the Universe shifted from being neutral to ionized at $z \gtrsim 6$. This epoch is called 'Reionization'. An enormous source of energy is required to drive this phase shift. The source of this is not known today, but the most likely candidate is massive stars. Reionization is a far less understood cosmological epoch than its much older cousin Recombination.

2.3.2 Lyman- α forest absorbers³

Bahcall and Salpeter (1965) predicted that a set of discrete absorption lines could also be found in the spectra of high-redshift QSOs. These lines are not connected with intergalactic gas, but to the presence of a distribution of inhomogeneous clumpy gas filaments belonging to galaxies intersecting the line of sight towards the QSO. Since these neutral gas clouds are small, their absorption lines are thin. But owing to the great Ly α absorption cross section of neutral hydrogen, the lines are often deep. Since these small absorbing structures are predicted to be numerous, a large amount of distinct absorption lines should be found bluewards of the Ly α line of the QSO each of them corresponding to a different redshift. This phenomenon was later observed and the term 'Lyman- α forest' was introduced. The small absorbing filaments are strongest by numbers shortly after the reionization, and becomes fewer as the Universe expands further. A small- z Lyman- α forest will have fewer and shallower lines than one seen in a high- z spectrum. An example spectrum is plotted in figure 3. Besides the Lyman absorption, a minor part of the lines belong to various other ultraviolet transitions from C, O, Mg, Si, Fe, Al and other heavy elements. By definition, Lyman- α absorbers have neutral hydrogen densities of $N(\text{HI}) < 10^{17} \text{ cm}^{-2}$. However, the total hydrogen content is much greater with the rest of it being ionized. The neutral component is only a tiny fraction of the ionized, and it is often less than 0.1% of the total (Sparke and Gallagher, 2007, 395).

2.3.3 Lyman limit systems⁴

Lyman limit systems are defined to have $10^{17} < N(\text{HI}) < 2 \times 10^{20} \text{ cm}^{-2}$. At these densities, the clouds will effectively absorb light with energies high enough to ionize hydrogen. Equivalently, the clouds will absorb light with wavelengths that are equal to or greater than the Lyman limit (hence the name) of 912 Å. The flux of the QSO will drop nearly to zero at the wavelengths immediately bluewards of $912(z + 1)$. As with the Lyman forest absorbers, the main hydrogen content is in the ionized component.

Of all the different types of QSO absorption systems, the Lyman limit systems (LLSs) are both the least studied and least understood. High resolution spectroscopy have revealed

³The following section is based on Rauch (1998)

⁴The following section is partly based on Sparke and Gallagher (2007, p. 390)

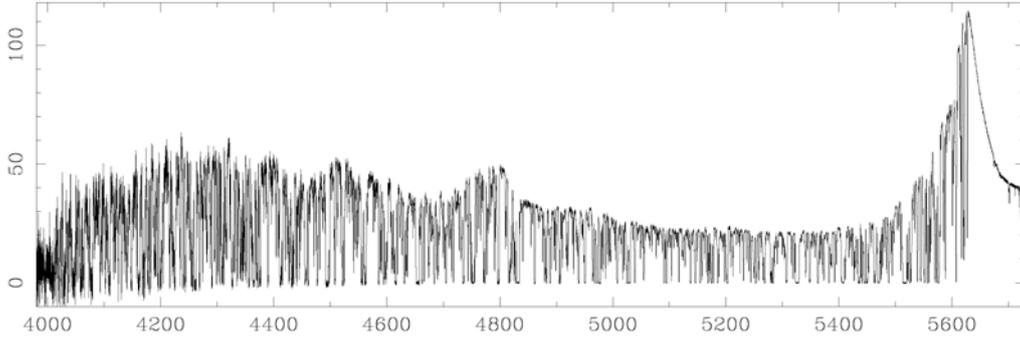


FIGURE 3: High resolution spectrum of *QSO1442 + 23* found at $z_{em} = 3.62$. The figure is clearly showing the Ly α forest. Figure from Rauch (1998) based on data from Womble et al. (1996, p. 137).

that some LLS are chemically evolved with $[\text{Fe}/\text{H}] = -0.5 \pm 0.1$, which resembles the metallicities of the DLAs (to be discussed below) (Prochaska, 1999). This could suggest that LLSs are the photoionized counterparts of the damped systems. However, other LLSs are found to be metal-poor and not easily associated with the DLAs (see Levshakov et al. (2003) for an extreme example). Simulations point towards LLSs being located within virial radii of galaxies with a wide range of masses, which suggests that the systems are somehow evolutionarily connected to them (Kohler and Gnedin, 2007).

2.3.4 Damped Lyman- α absorption systems

Like other classes of QSO absorbers, DLA systems are selected on the basis of high HI column densities. The criterion for the density is $N(\text{HI}) \geq 2 \times 10^{20} \text{ cm}^{-2}$, for which the Ly α is optically thick even in the absorption wings (Wolfe et al., 1986). The line will show the prominent damping wings that have given rise to the name 'Damped Lyman- α absorber'. In fact, the damping wings appear even for $N(\text{HI}) > 10^{19} \text{ cm}^{-2}$ (Petitjean, 1998), but the first surveys of damped system were performed at low resolution, and the high density limit was introduced to avoid blending errors. Another motivation for the choice of this limit is the assumption that these lines should be characteristic of galaxies at high redshifts. The limit is thus an association of DLAs with early galaxy structures. Of course, this classification is artificial and potentially might introduce misconceptions in our understanding of the nature of DLAs. It is worth stressing that the classification is just a way of introducing a nomenclature for the different observations in the spectra of QSOs. The actual physical systems are most likely less sharply distinguished.

The DLA systems differ from the other classes of QSO absorbers by more than just the somewhat arbitrary selection criterion; contrary to Lyman limit absorbers and Ly α forest absorbers, the gas content of the DLA systems is mainly neutral. This motivates the before mentioned hypothesis that DLA systems are the birthplace of stars and thus galaxies in the early Universe (Wolfe et al., 2005).

Figure 4 shows the column density distribution of HI clouds for $z \sim 2.8$. The distribution is well-fitted over two orders of magnitude with a power-law:

$$\frac{d^2n}{dNdz} \propto N^{-1.5} \quad (1)$$

Because of the shape of the distribution, most of the mass of the neutral gas will belong to the densest systems even though the diffuse clouds are by far the most numerous. In other words, the DLA systems represented in the figure by only four markers beyond $N(\text{HI}) = 10^{20}$ account for much more neutral gas than the remaining majority of systems, which all belong to the class of Ly α forest absorbers. This notion further strengthens the

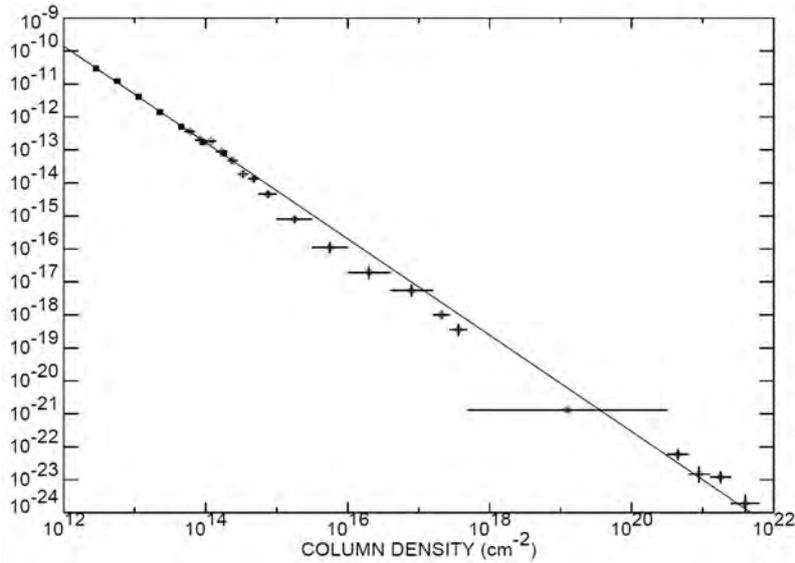


FIGURE 4: Plot of the column density distribution of HI clouds for $z \sim 2.8$. Figure from Petitjean (1998). The figure was printed without units on the y-axis in the original paper. The y-axis shows the differential column distribution $f(N, X)$ introduced in Petitjean et al. (1993), where it is defined such that $f(N, X)dNdX$ is the number of absorption systems with column density between N and $N + dN$ and radial coordinate between X and $X + dX$ with $X = 0.5[(1+z)^2 - 1]$ for $q_0 = 0$. It is assumed that it does not depend on X in the redshift range.

hypothesis that DLA systems is where we should look for early galaxy structures and star formation.

Besides hydrogen, DLAs are also observed to contain many different metals. These are in the form of ions, and are seen by emission from their various transitions. The strongest lines belong to low-ionization species such as MgII, SiII and OI. Higher ionization lines such as CIV and SiIV are also seen (Sparke and Gallagher, 2007, p. 393). The metallicities of DLAs have a large scatter at all redshifts, and for $z \sim 2-3$ the mean metallicity is $\sim 5-10\%$ of the solar value (Kulkarni and Fall, 2002). In accordance with the cosmological chemical evolution of the Universe, the mean metallicity is expected to increase with decreasing z . Kulkarni et al. (2005) showed that the abundance of zinc relative to hydrogen is slowly increasing with decreasing redshift confirming the expected relation.

The presence of metals is a strong indicator of active star formation and another link between DLAs and early galaxy structures. It is, of course, very interesting to ask whether or not the associated young stellar populations are actually visible themselves. A deeper understanding of the physical nature of the DLAs requires direct observation of the galaxy counterparts from emission and not just by their absorption of background sources.

2.4 Galaxy counterparts

The first paper to present a sample of DLAs was Wolfe et al. (1986). Already that same year, Foltz et al. (1986) suggested that it should be possible to directly detect Ly α emission from the absorbing structures. The emission should be seen as a sharp spike in the middle of the damped Ly α line trough. The first successful detection of Ly α emission from a DLA was done by Møller and Warren (1993). They used long slit spectroscopy and achieved a 4.2σ detection of Ly α emission from an object at an angular separation of $1''.32$ seen towards a QSO catalogued as PKS0528-250 at redshift $z = 2.77$. A frame of the 2D-spectrum showing the Ly α trough is shown in figure 5. The expected spectroscopic features suggested by Foltz et al. (1986) is clearly seen. The dark spot in the continuum gap is the Ly α emission, which the authors associated with the absorber on the basis that the

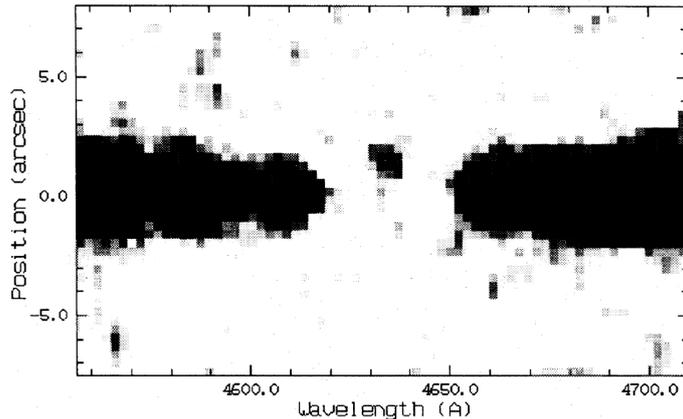


FIGURE 5: Two dimensional spectrum of the $z = 2.77$ -QSO, PKS0528-250. The dark band is the QSO-continuum. The gap is the damped $\text{Ly}\alpha$ line. The dark spot in the middle of the trough is $\text{Ly}\alpha$ emission associated with the absorber. Figure from Møller and Warren (1993).

velocity difference between absorber and emitter was only $50 \pm 100 \text{ km s}^{-1}$. The authors were aware that the emission could be the result of photoionization from the QSO, and not from star formation. However, based on the morphology of the features, the authors argued in favor of the latter explanation.

The detections of emission from DLA galaxy counterparts were sparse in the following years. Fynbo et al. (1999) detected unusually high emission from a DLA at $z = 1.93$. The anomaly was explained as a result of the DLA being photoionized by another QSO, which didn't show any damping features. The article also proposed the idea that the galaxy counterparts of DLAs could be connected to *Lyman break galaxies*.

2.4.1 Lyman break galaxies

Lyman Break Galaxies (LBGs) are a class of high-redshift ($3.0 \lesssim z \lesssim 3.5$) galaxies, which are simply defined by two spectral features: a roughly flat spectrum in the far-UV and a spectral break at the Lyman limit (912 \AA). They were first detected and described by Steidel et al. (1996) who were searching for high-redshift star forming galaxies. The authors assumed that these galaxies would show the before mentioned spectral features. The assumption was made for the following reasons: young massive stars should dominate the far-UV, hereby producing the flat spectrum and the Lyman break should be produced by intrinsic stellar energy distributions, the inherent opacity of galaxies to Lyman photons, and the effect of intervening absorbers. Using a color selection technique based on a custom suite of broadband filters, a population of the anticipated LBGs was discovered. Their spectral features were remarkably similar to present day star forming galaxies: flat continuum, weak $\text{Ly}\alpha$ emission, prominent high-ionization stellar lines of He II, C IV, Si IV, and N V, and strong interstellar absorption lines due to low-ionization stages of C, O, Si, and Al. It was hereby concluded that LBGs are the high- z counterparts of today's luminous star forming galaxies.

The suggestion by Fynbo et al. (1999) that DLAs should be associated with LBGs was partly motivated by models by Pei and Fall (1995). Based on an analysis of the cosmic decline of HI, Pei and Fall (1995) concluded that the star formation rates in DLAs match the star formation rates in LBGs. Fynbo et al. (1999) considered a scenario where every Lyman break galaxy is surrounded by a damped $\text{Ly}\alpha$ disk. In this model, DLAs are the gas reservoirs from which the stars in LBGs are formed. Møller et al. (2002) investigated this conjecture by comparing spectral properties of 3 DLAs with those of a larger catalogue of LBGs. The authors found that for half-light radius, radial profile, optical-to-near-infrared

color, morphology, Ly α emission equivalent width, and Ly α velocity structure there was no conflict with the prediction.

However plausible, the connection between DLAs and LBGs are far from firmly established. Even though both classes have been known for decades, very little observational overlap exists between the two; only very few DLAs have been observed from emission. Fynbo et al. (2008) offers a possible explanation for the missing link: where the detection of LBGs is limited by flux only, DLA detection is limited by the probability of a sightline towards a QSO passing through it, which scales with the cross section of the cloud. In the local Universe, this cross section scales with the luminosity to a given power. Assuming this relation holds for high z , it is found that DLAs are selected towards the faint end of the luminosity function. In this line of thought, LBGs and DLAs are respectively drawn from the bright and faint end of the luminosity function of the same population. To thoroughly unite DLAs and LBGs, it is necessary to observe more DLAs in emission. A method for selecting the most luminous DLAs would be valuable.

2.4.2 Metallicity-luminosity relations

In the local Universe, galaxies have well-known correlations between metallicity and star formation rate (SFR) in the sense that there is a tendency for high-mass galaxies to also have high SFRs and metallicities. It would be tempting to assume that a similar connection could be found for high- z galaxies. In order to test this hypothesis, Møller et al. (2004) looked at a NICMOS⁵ sample of 17 DLAs with known metallicities. Ly α emission have been observed for a subsample of 4 of these DLAs. Figure 6 shows a histogram of the metallicities of the sample. The population with observed Ly α emission is marked with blue markers whereas the population without is marked with red ones. There is a clear trend for the blue population to have higher metallicities. In the right plot, sub-DLAs (HI-column density smaller than the normal limit) have been removed. The general trend is unchanged. As previously described, Ly α emission is generally linked to star formation and hence the general trend in figure 6 is an indicator of a positive correlation between metallicity and star formation for high- z DLAs. The samples are too small to provide a proof of the relation, but the figure gives a good motivation for using high metallicity as a selection criterion when searching for luminous DLAs.

Fynbo et al. (2010) did just that. Using the, at that time newly commissioned, X-Shooter spectrograph on VLT (Very Large Telescope) they targeted high-metallicity DLAs. Noterdaeme et al. (2009) had developed a catalogue of 1426 DLAs in the redshift-range $2.15 < z < 5.2$, which were identified in the SDSS⁶ Data Release 7 with a fully automated procedure. Redshift, HI column density, and equivalent width of associated metal absorption lines were determined for each DLA. Using this catalogue, Fynbo et al. (2010) selected DLAs with rest-frame equivalent width of the SiII line larger than 1 Å under the well-justified assumption that this is a good indicator of high- Z ($> 0.1 Z_{\odot}$). The X-Shooter is a multi-wavelength three-arm echelle spectrograph covering 3000 to 25000 Å. The three arms provide a big field of view with a predicted less than 10% of galaxy counterparts uncovered by the slits (Fynbo et al., 2008). The authors only choose DLAs with redshifts close to 2.4. The spectral range covered by the X-Shooter at these redshifts extends from the Lyman limit to H α . The H β , [OII], [OIII] and H α wavelengths are in the near infrared. The method proved to be successful; from the first target, the QSO Q2222-0946, a DLA was observed in Ly α , [OIII] and H α emission. A metallicity of $[Zn/H] = -0.46 \pm 0.07$ was inferred corresponding to 2/5 of the solar value. From the emission properties, a lower limit on the star formation rate was set as high as $10 M_{\odot} \text{ yr}^{-1}$.

⁵The Near Infrared Camera and Multi-Object Spectrometer onboard the Hubble Space Telescope

⁶Sloan Sky Digital Survey II. A major multi-filter imaging and spectroscopic survey based on data collected with the Apache Point Observatory in New Mexico, United States.

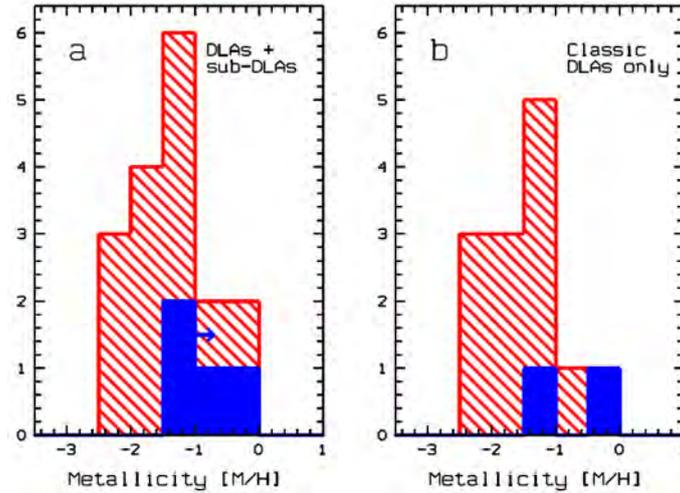


FIGURE 6: Histogram of metallicities for 17 DLAs. The blue markers corresponds to DLAs where $\text{Ly}\alpha$ -emission have been directly observed. These are generally found towards higher $[\text{M}/\text{H}]$ -values. In the right plot, sub-DLAs (HI-column density smaller than the normal limit) have been removed. The general trend is unchanged. Figure from Møller et al. (2004).

2.4.3 Dust attenuation

Fynbo et al. (2011) continued the method and observed a $z = 2.58$ DLA in the spectrum of Q0918+1636. The metallicity-luminosity conjecture was once again confirmed since the DLA had a very high metallicity. In fact, the metallicity was exceptionally high; possibly even higher than the solar value. This time the DLA was observed in [OII] and [OIII] emission, but no $\text{Ly}\alpha$ emission was detected. Based on a derived upper limit of the $\text{Ly}\alpha$ flux, an upper limit on the star formation rate⁷ was found, $\text{SFR} < 0.3 M_{\odot} \text{ yr}^{-1}$. At the same time, the [OII] and [OIII] fluxes yielded much higher star formation rates: $\text{SFR} \sim 20 M_{\odot} \text{ yr}^{-1}$. This contradiction led the authors to the conclusion that the Lyman emission was lost in dust obscuration. Given the high metallicity, the presence of dust was very likely. To test this hypothesis, the overall spectrum of the QSO was compared with a composite QSO spectrum. It was evident from the comparison that Q0918+1636 was much redder than a "normal" QSO. It had a rest-frame extinction of $A_V = 0.2$ mag, which was most likely due to dust attenuation from the observed DLA. The extinction had moved Q0918+1636 to the boundary color of QSOs included in the SDSS data collection. On these grounds, Fynbo et al. (2011) argued that many high-metallicity DLAs might possibly be systematically omitted from current catalogues. This could be another reason why so few DLAs have shown emission so far.

The possible dust bias of DLAs have been discussed in the literature for more than 30 years. Wolfe et al. (2005) presents a review. Several studies have been devoted to the subject, but their results do generally not agree. Fynbo et al. (2013) selected QSOs from a combination of optical and near-infrared imaging. This proved to be an efficient method for targeting reddened QSOs. The approach was further revised by Krogager et al. (2015) with a modification of the selection criteria that rejected contamination from cool stars and evolved galaxies, which had been present in the before mentioned survey. The survey identified almost 150 dust-reddened QSOs that was not present in the SDSS catalogue, and had "fallen out" as a result of the colour selection technique. Most of these QSO turned out to be reddened by dust in their host galaxy and not by intervening absorbers. However, the survey did identify one clear example of a QSO reddened by a DLA in the

⁷See appendix B for a review of the formalism for calculating star formation rates from $\text{H}\alpha$ and $\text{Ly}\alpha$ fluxes.

sightline. The QSO does not appear in samples based on the SDSS-catalogue and is an example of a dusty QSO omitted as a result of previous color selection techniques. The QSO is catalogued as Q2225+0527. The corresponding DLA is presented in Krogager et al. (2016).

2.5 DLA2225+0527

DLA2225+0527 lies at redshift $z = 2.13$ in the sightline of the QSO (with the same catalogue number) at $z = 2.32$. Krogager et al. (2016) fitted various dust-templates to the spectrum of the QSO. Models including a term of dust in the DLA gave equally good fit as models with dust in the QSO only. However, a very low very ratio of iron to zinc was measured, $[\text{Fe}/\text{Zn}] = -1.22$. This ratio is a measure of depletion on dust (De Cia et al., 2013)). The abundance of the two metals is assumed to be intrinsically equal, but iron depletes on dust whereas zinc rarely depletes. A small number for the ratio of the two is thus a tracer of dust, and in this case favors dust in the DLA over dust in the QSO. Also, Krogager et al. (2016) detected strong H_2 absorption lines. This also indicates dust in the DLA since molecular hydrogen form most effectively on the surface of dust grains (e.g., figure 8 of Noterdaeme et al. (2008)).

It has so far not been possible to observe the DLA in $\text{Ly}\alpha$ emission, which is not surprising since this is to be expected of a very dusty DLA as described earlier. More surprising is it that emission from [OIII] is also not present. Given the very high metallicity, the galaxy counterpart is expected to be very massive and luminous. An upper limit for the star formation rate has been calculated by Krogager et al. (2016) to $\text{SFR} < 6 M_{\odot} \text{ yr}^{-1}$. Again, this is a very low level considering the high metallicity. The authors notice that some emission might have been lost outside the slits and outside the imaging. Another explanation could be that the dust is not evenly distributed, and that the star forming regions are encapsulated by a region of even higher dust obscuration, which is not probed by the line-of-sight.

A detection of the galaxy counterpart of DLA2225+0527 is very desirable for various reasons: it would be another confirmation of the relation between high metallicity and luminosity, it would be the first solar-metallicity DLA to be directly detected, and it would confirm that some emitting DLAs are indeed lost by the SDSS color selection technique.

3 Observations and data reduction

The data used for this project is a compilation of data from different programs - both imaging and spectroscopy. The following is a presentation of the data, starting out with a description of the instrumentation and observational conditions. This is followed by a review of the performed data processing and some of the underlying theoretical aspects. I will treat imaging and spectroscopy separately in two sections.

DLA2225+0527 was initially observed and identified in the High A_v Quasar Survey described in Krogager et al. (2015). This survey selected red candidate QSOs based on photometric data from the SDSS and UKIRT Infrared Deep Sky Survey (UKIDSS). The photometry data were supplemented with mid-infrared photometry from the Wide-Field Infrared Survey Explorer (WISE). 159 candidates were selected and observed with low resolution spectroscopy using the Andalucia Faint Object Spectrograph and Camera (ALFOSC) at Nordic Optical Telescope (NOT) on La Palma during observations runs in 2012, 2013 and 2014. The DLA was observed on the night of August 27, 2014. The following nights it was observed at high resolution. The metallicity was estimated by observing the ZnII λ 2026 transition with the high resolution grism 17 covering 6500 Å.

3.1 Imaging

Imaging was done using the ALFOSC in imaging mode. Only one filter was used: the r' -filter centered on 618 nm. The resolution of the CCD is $0''.19/\text{pixel}$. The observations were made using a five-point dither with sub-integrations of 330 seconds. A total of 42 integrations were made. Krogager et al. (2016) chose the 21 frames with the best seeing. Out of these, I discarded 5 frames that had a different size-format which would complicate the combining of the frames. The final 16 frames give a total integration time of 5280 seconds.

The following subsections are a step-by-step presentation of the data reduction that I have performed - first the basic imaging processing followed by the actual data analysis.

3.1.1 Basic data processing

The theoretical background is partly based on chapter 4 of Howell (2006). I refer to this book for a more thorough introduction on the subject. I perform the reduction using various tasks in IRAF⁸. These will be presented and explained along with relevant parameters and results. For displaying and examining the images I use the programs GAIA⁹ and SAOImage DS9¹⁰

⁸Image Reduction and Analysis Facility - A collection of software written at the National Optical Astronomy Observatory.

⁹Graphical Astronomy and Image Analysis Tool. A derivative of the Skycat catalogue developed for the VLT project at ESO

¹⁰Astronomical imaging and data visualization application developed at the Smithsonian Astrophysical Observatory.

The purpose of the basic data processing is to remove various instrumental sources of error in order to recover the original true signal. Each step will introduce some noise and uncertainty, which of course should be minimized as much as possible. The main technical corrections are bias subtraction and flat fielding. Bias is an electronically added positive offset value that is added to each pixel of the CCD in order to avoid negative pixel values. Negative pixel values are the result of unexposed pixels. The value for zero received photons will result in different values with a small distribution about zero in the final image after read-out and A/D-conversion (Howell, 2006, p. 52). The bias is generally not spatially constant over the CCD, and bias frames should be extracted and used for correction. A bias frame is a zero seconds exposure with closed shutter. Bias is an additive effect, which means that bias frames should be subtracted from the science frames.

Flat fielding is a correction for response gradients across the CCD. The pixels of the CCD are generally not equally responsive. This effect depends both on their position on the CCD and the spectral properties of the incoming light. Flat fields are exposures of an evenly illuminated surface. An evenly illuminated surface can be achieved either by observing the sky at twilight or the inside of the closed telescope illuminated by a lamp. The correction is a multiplicative error, and the science frames should be divided by the flat field. The final flats should be normalised before division.

Written in an equational form, the reduction process looks like this:

$$\text{Final science frame} = \frac{\text{Raw frame} - \text{Bias frame}}{\text{Flat field frame}} \quad (2)$$

Besides flat fielding and bias subtraction a number of other operations are sometimes performed e.g. correction for dark current. That won't be necessary for this data set. For both flats and bias frames many exposures are combined to get better signal-to-noise ratios. Combining exposures is also done when processing science frames. For each type of frames it is necessary to decide which combination technique should be used and how many frames. For each combination I use the IRAF command **IMCOMBINE** with different chosen parameters.

Figure 7 shows an unprocessed science image. I notice that there is ~ 50 pixels of overscan along all edges except the bottom. Overscan regions are additional pseudo-pixels that are added to the frame during read-out. They do not correspond to any physical CCD-pixels. Before performing operations on any of the images, I crop these regions away using **IMCOPY** and specifying a new smaller area of [60 : 2008, 1 : 2030].

3.1.2 Bias frames

The bias-frames show no clear gradients and no cosmic ray hits are seen. With the IRAF command **IMSTAT**, the bias-frames are found to have mean values of 10006 - 10015 ADU¹¹ and standard deviations of 12.03-12.7 ADU. According to Howell (2006, p. 73):

$$\text{RON} = \frac{\text{gain} \cdot \sigma_{B1-B2}}{\sqrt{2}} \quad (3)$$

where $B1 - B2$ is the difference-frame produced by subtracting two different BIAS-frames. σ refers to the standard deviation of this frame. The gain is found in the FITS¹²-header of the images: $\text{gain} = 0.33 \text{ e}^{-1}/\text{ADU}$. However, this factor only serves the purpose of conversion between $\text{e}^{-1}/\text{pixel}$ and ADU/pixel . Since I want to keep the units in ADU, I

¹¹Analog-to-Digital unit

¹²Flexible Image Transport System. A digital file format for scientific images which can contain a large amount of metadata stored in the 'header'.



FIGURE 7: An unprocessed science frame. The overscan is visible towards the edges. The brighter glow towards the center is not an actual astrophysical signal, but it is the effect of pixel-to-pixel response variations that will be corrected for in the flat fielding.

will omit the gain in the above formula. I produce the difference-frame with the IRAF-command **IMARITH**, and find the standard deviation with **IMSTAT**: $\sigma_{B1-B2} = 17.2$ ADU. Putting the values in the formula above, I get $\text{RON} = 12.1$ ADU. This method should be more precise than just calculating the RON from the standard deviations mentioned above since any gradient in the frame is removed by producing the difference picture.

To produce the master-bias-frame, I use the IRAF-function **IMCOMBINE** with the combining algorithm set to **average**. This setting produces a weighted average for all pixels. I also use a clipping algorithm for rejecting extreme pixels values e.g. hot pixels or cosmic ray hits. I choose the algorithm **ccdclip**, which is using the noise-parameters from the FITS-header for rejecting bad pixels. I find that these choices for the **IMCOMBINE**-function produces the lowest standard deviation in the master-bias frame, which is $\sigma_{\text{masterbias}} = 3.651$ ADU - much lower than the standard deviation in any of the uncombined bias-frames.

3.1.3 Flat fields

All flats are automated sky-flats. Before combining them into one frame, I correct them for bias by subtracting the above produced master-bias-frame from each of the 3 frames. Again, I use **IMARITH** for subtraction. I combine the flats using **IMCOMBINE**, but this time I set the scale parameter to **MODE**. This setting multiplies each frame with a factor close to one, which makes the mode (most common value) of each frame the same. I use the combining algorithm **MEDIAN**. This algorithm will choose the median value for each pixel when combining. Having produced the combined flat-frame, I normalise it by dividing the whole frame with the mode (which I find with **IMSTAT**). The division is also performed with **IMARITH**. Afterwards, I check the normalisation with **IMSTAT**. The mean is 1.024 ADU with a standard deviation of 0.04666 ADU. The quality of the flat field is checked by flat fielding a flat field - that is dividing one of the flat fields by the normalized master flat field. The result is an almost completely blank frame confirming the quality of the flat field frame.

The final combined flat field is shown in figure 8. There are clear gradients seen in the frame. The field also reveals two dead rows and some round structures that are the result of dust specks on the window or filter. All of these effects will be removed when flat fielding the science images.

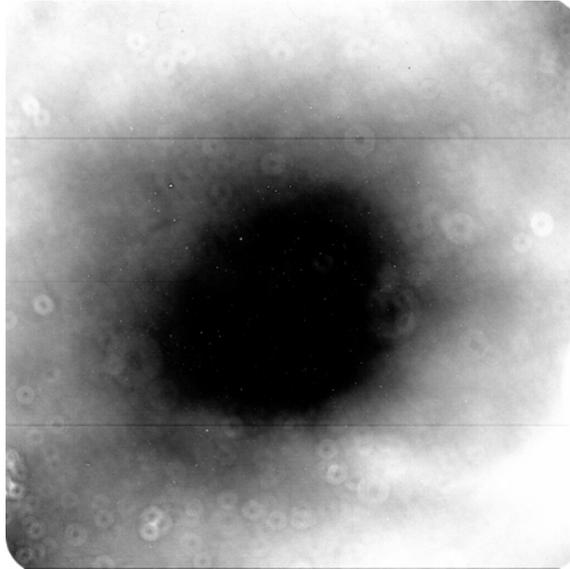


FIGURE 8: The final combined master flat image. There are clear gradients in the image. Also present are two dead rows and some dust specks on the window or filter, which are seen as round structures.

3.1.4 Processing science frames

I perform the basic operations on the science frames: cropping the overscan, correcting for bias, and flat fielding. The resulting images have standard deviations spanning from 2447 to 3363 ADU.

The science frames should be combined to achieve better signal-to-noise ratios. All of them have the same exposure time $t_{exp} = 330$ s. Before combining the frames, I have to align them to adjust for small offsets in the x,y-directions due to imperfect tracking. This can be done with **IMALIGN**. The program needs a list of coordinates on bright stars in the field and estimates of the shifts. These two lists can be produced with **DAOFIND** and **IMCENTROID**. These tasks are found in the aperture photometry package **APPHOT**. Before this can run correctly, several parameters must be set in **DATAPARS**. These are seeing, airmass, sky background, RON, gain, and a maximum and minimum value for good pixel data. Not all these parameters are strictly important for running the operations and analysis that I perform in this research, but since they form part of a routine IRAF-run, I have included them here for the sake of completeness:

- Seeing is blurring caused by turbulence in the atmosphere. It causes point sources to appear as "fuzzy blobs". Seeing can be measured as the full-width-at-half-maximum of the point spread function (PSF)¹³, and is actually a measure of the true resolution that the atmospheric conditions allows for. I choose to measure the FWHM of the PSF of Q2225+0527 in the middle of the frame, and I find it to be 3.95 pixels. Converted to arc seconds, this corresponds to 0".75.
- In order to correct for extinction in the atmosphere (primarily Rayleigh-scattering), IRAF needs to know the optical path length of the atmosphere above the telescope. This number is causally referred to as the "airmass". The airmass is given in the FITS-header = 1.31689. It is given relative to the path length at sea level, which per definition is 1.
- The standard deviation parameter is important for IRAF in order to determine which imaging features should be considered as true objects, or which are just fluctuations

¹³The point spread function describes an imaging system's response to point sources.

of the background. The standard deviation of the sky background is found by running **IMSTAT** on a small region (60×60 pixel) containing no objects or cosmic ray hits. I get $\sigma = 83.9$ ADU.

- IRAF uses the read-out noise and gain level for calculating errors. I have previously found both the read-out noise, $\text{RON} = 12.1$ ADU and $\text{gain} = 0.33 e^{-1}/\text{ADU}$.
- It is necessary to set a maximum and minimum level for the values that IRAF should take under consideration when performing the various operations. This prevents IRAF from including e. g. saturated stars in the calculation. For the maximum good pixel value I choose 295000 ADU, which is safely below the saturation level: 300000 ADU. I set the minimum value to 2043 ADU, which is more than 5 sigmas below the mean sky background: 2463 ADU.

Having set all the values, I can now run **DAOFIND** on the first science frame. I choose 70 sigmas for the detection threshold in **FINDPARS**. This very high value is chosen because the purpose of this mapping is only to find strong sources, which I can use for calculating offsets between the frames (not mapping every object in the frame). The **DAOFIND** task produces a list of coordinates of all the objects found in the image. With **TVMARK** I overplot the science images with all the found sources in DS9. Afterwards, I manually remove the coordinates of a few wrong markers from the list. **IMCENTROID** can be used for calculating the x,y-offsets between the frames using the first science frame as a reference. Setting the parameters **BOXSIZE** and **BIGBOX** correctly is very important; they approximate values for fine centering and coarse centering. I manually check for approximate maximum offsets in the frames and afterwards choose **BOXSIZE** = 25 and **BIGBOX** = 60. **IMCENTROID** produces a list of offsets. Finally, **IMALIGN** can align all frames when given the lists of coordinates and offsets and the same parameters for **BOXSIZE** and **BIGBOX**. The program also trims the frames to the matching areas, which means that ~ 30 pixels are cut off along each edge. I combine all the aligned science frames with **IMCOMBINE** using the algorithm **MEDIAN** and the rejection mechanism **AVSIGCLIP**, which rejects extreme pixel values based on average pixel values in the frames.

The final combined science image is shown in figure 9. Using **IMSTAT**, I find that the pixel values range from 377.1 to 614562 ADU with a standard deviation of $\sigma = 2965$ ADU. The FWHM is 3.43 pixels or $0''.65$, which is $0''.10$ lower than in the single frames. No cosmic ray hits are seen confirming that the various rejection mechanisms have been effective. Comparing with figure 7, it is obvious that the combined science frame shows much more detail than the single science frame. This can serve as a qualitative confirmation of a higher signal-to-noise ratio for the combined and processed science image.

3.1.5 PSF-subtraction

To reveal any faint structures close to the QSO, it is necessary to 'remove' the bright QSO from the image. This can be done by modelling the point spread function of the image. I re-run **DAOFIND** on the final science image with new chosen parameters in **DATAPARS**. The FWHM is now measured to 3.43 pixels, the background mean is 2425 ADU, and sigma is 24.1 ADU. I notice that these values have been improved from the single science frames. I run the task with a threshold of 10, and by choosing this much lower value I ensure that all significant features of the images get marked. I check the resulting coordinates with **TVMARK**. The task is now much more precise and efficient in finding features in the image - a result of the better statistics of the final image.

The task **PHOT** can be used to calculate instrumental magnitudes for all objects in the list. The magnitudes for each object is calculated as

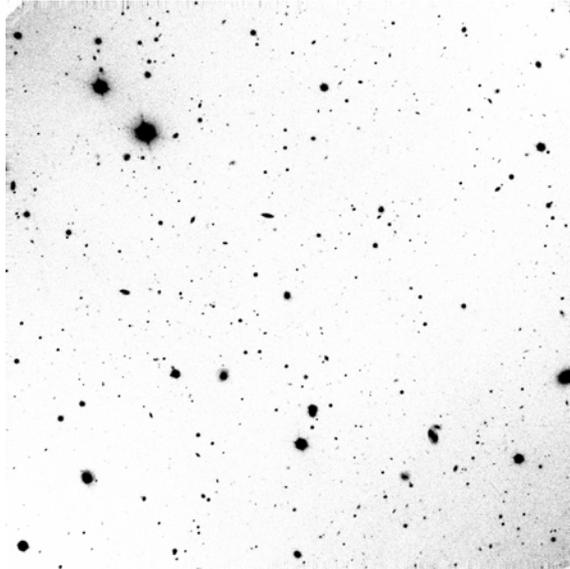


FIGURE 9: The final combined science image. The frame shows more detail than a single science exposure. This can be seen by comparison with figure 7. The colormap has been inverted in this figure. No apparent cosmic ray hits are seen in the image. Some pixels have been "stretched" along the edges. This is an unwanted side-effect of the alignment and combining of the frames. However, the edges contain no objects of interest, and the stretching does not constitute a problem for the further processing.

$$m = m_{zero} - 2.5 \log F + 2.5 \log t \quad (4)$$

where t is the integration time, m_{zero} is a chosen constant which is 25 by default. F is the flux of the object as calculated by

$$F = \Sigma - A \times m_{sky} \quad (5)$$

where Σ is the sum of all pixels in an aperture with area A centered on the object. m_{sky} is an estimate of the sky value per pixel. The aperture radius should be specified. I choose 10 pixels based on a qualitative analysis of the average object diameter in the field. The estimate of the pixel sky value is calculated by IRAF by fitting a function on the sky within an annulus centered on the star. The inner and outer radius of this annulus should be set. In a rather uncrowded field as the relevant, these values can safely be set as large as 30 and 60 pixels for the inner and outer respectively. After having plotted a few histograms of sky-only regions, I choose a Gauss-function for the sky-fitting.

The next step is to produce the PSF-model, which can be done with the task **PSF**. This task produces a two-component model based on a selection of suitable stars in the field. The first component is an analytic function that approximates the light distribution in a specified central region of the stars. The radius of this region is set within **DAOPARS**. A good choice for this parameter is the seeing, which I earlier measured as the FWHM. In **DAOPARS** is also chosen a function for the analytic fit. I choose a Moffat¹⁴ function since it is more precise in the wings of the PSF than the Gaussian.

The second component of the model is a lookup table with all the residuals of the individual fits. The sampling rate of this table is twice that of the original image. The

¹⁴The Moffat function is a continuous probability distribution based upon the Lorentzian distribution. In terms of a vector r it has the functional form $f(r) = a[1 + (r/\alpha)^2]^{-\beta}$ (from https://www.gnu.org/software/gnuastro/manual/html_node/PSF.html)

brightness of any arbitrary pixel is calculated by integrating the analytic function over the area of the pixel. An additive correction is then determined by interpolating within the lookup table. The parameter **PSFRAD** defines the region that IRAF subtracts from each star. This should be set equal to the radius for which the biggest star of interest fades into noise. I qualitatively determine this value to be 25 pixels for the science image.

Having all parameters set, **PSF** can run. The task is run either automatically or interactively. Choosing the latter possibility allows the user to manually select stars to build the PSF from by clicking on them in **DS9**. When a star is clicked, a 3D-model is shown along with the calculated magnitude, coordinates, and other data. Hereafter, the star can either be accepted for the model or rejected. I proceed in this manner, and build a list of 10 stars with magnitudes in the range of the QSO that all have well-defined round shapes. Having set the max limit for good data values below the saturation level, all saturated objects are automatically pre-rejected by the program. After the PSF is built, the task **ALLSTAR** can be run with the PSF-model. The task subtracts all stars in the field, so the quality of the model can be checked. With the parameters chosen above and the selected list of PSF-stars, the subtraction is robust and efficient and the immediate surroundings of the QSO can now be searched for faint objects that might constitute as DLA-candidates. Images of the quasar before and after PSF-subtraction is presented under 'Results'. This concludes the basic imaging data reduction.

3.2 Spectroscopy

Following the initial selection and identification, the DLA was observed with the X-Shooter spectrograph at the VLT on October 30, 2014 and November 23, 24, and 25, 2014. The X-Shooter is presented in Vernet et al. (2011). The spectrograph splits the light into three so-called arms: ultra-violet (3000 - 5500 Å), visible (5500 - 10000 Å) and near-infrared (10000 - 25000 Å). In this way, the X-Shooter covers the wavelength range from 3000 Å to 2.5 μm simultaneously.

The X-Shooter is equipped with an atmospheric dispersion corrector (ADC) consisting of two counter-rotating double prisms. This device prevents the light from being dispersed in the vertical direction as a result of terrestrial atmospheric refraction. The dispersion would cause some wavelengths to fall outside the slit. The ADC was malfunctioning on the nights of observation, and in order to prevent slit loss the observations had to be performed at parallactic slit angle, i.e. the angle at which the slit is aligned with the direction from the target to the zenith. For this reason, only two different angles were obtained: one from the night in October, $\text{PA} = -4^\circ$ east of north (hereafter referred to as A) and three almost identical angles from the November nights, $\text{PA} \sim 39^\circ$ east of north (hereafter referred to as BCD). See figure 15 for a plot of the angles.

All observations were carried out at same slit-width hereby making the spatial direction of a 2D-spectrum correspond to the actual physical distance from the QSO. The wavelength resolution of the X-Shooter for this survey was 0.06 nm/pixel and the spatial $0''.15/\text{pixel}$. The basic data reduction was done by Krogager et. al using the official X-Shooter pipeline **ESOREX**. Besides the flux data, each **FITS**-image also contain an error image and a so-called quality image where erroneous or bad pixels are marked.

The main purpose of the spectroscopic analysis is to look for DLA emission at wavelengths corresponding to nebular emissions lines at the DLA redshift. The redshift is determined in Krogager et al. (2015). The lines of interest are [OIIIa], $\lambda_{\text{OIIIa}} = 5007 \text{ \AA}$ and $\text{H}\alpha$, $\lambda_{\text{H}\alpha} = 6562.8 \text{ \AA}$. These are strong emission lines, and their redshifted wavelengths can be calculated with the general formula:

$$\lambda_{\text{obs}} = \lambda_{\text{emit}}(z + 1) \quad (6)$$

At $z_{\text{DLA}} = 2.132487$ the lines are found to have the following observed redshifts: $\lambda_{\text{OIIIa}} = 15683.86 \text{ \AA}$ and $\lambda_{\text{H}\alpha} = 20557.89 \text{ \AA}$. These wavelengths are well-covered by the wavelength range of the X-Shooter.

The following subsections are a review of the various data reduction and analysis that I have performed on the spectroscopic data. The reduction is done with a program that I have written for the project. The code is inspired by different routines written by Krogager, J.-K. and Fynbo, J. P. U. The programming language is Python, and the full pipeline can be found in appendix A.

3.2.1 Initial setup

I use the package **PYFITS** for loading the FITS-data. For each image three arrays are produced: the spectroscopic data (**image**), the error image (**sig**), and the FITS-header (**header**). The spectra can now be subject to regular array manipulations. The combined B + C + D spectra have had cropped away the 4 first and last pixels in spatial direction, because the edges of the spectra are very noisy and erroneous (see figure 10). Before proceeding further, I crop the A-spectra in the same way. For simplifying later calculations, the error data from standard deviation is converted into variance by squaring the **sig**-array.

Header data of interest can be extracted from the header array by calling it with the relevant keywords. I extract the radial velocity of the QSO, and calculate the Doppler shift of the quasar lines using the usual formula

$$\Delta\lambda = \left(1 + \frac{v_{\text{rad}}}{c}\right) \quad (7)$$

where v_{rad} is the radial velocity of the QSO. I also extract the value of the smallest wavelength and the wavelength resolution. Combining these values with the Dopplershift, I produce an array **wl_obs** that contains the value of the observed wavelength corresponding to each column in the **image**-array

$$\lambda_{\text{obs}} = \left(1 + \frac{v_{\text{rad}}}{c}\right) \times \text{res}(\text{"/pixel}) + \lambda_{\text{min}} \quad (8)$$

where λ_{min} is the smallest wavelength of the spectrum and **res** is the resolution.

3.2.2 SPSF-subtraction

As with the imaging data, it is necessary to subtract the QSO emission from the spectrum to look for nearby DLA emission. To do this, I model the QSO spectral point-spread-function and subtract this from the image. For each line of interest I have by eye found two nearby small sections of wavelength-width 15 – 30 pixels ($\sim 15 \text{ \AA}$) left and right of the expected location of the emission line. These sections will be used for calculating the basic properties of the SPSF-model. They are chosen where the QSO line is continuous and no skylines are seen. Also, they must be fairly far away from any line of interest in order not to include an emission line in the SPSF-model. See figure 10 for an example. I manually declare the beginnings and ends of these sections with their pixel coordinates. I also declare the approximate line wavelength and a cut-out region of 402 pixels somewhat centered on the line.

The code proceeds by looping over all columns in the spectra cut-out. For each column, I fit a normal Gauss-function in the spatial direction around the QSO line. Before fitting, I calculate a standard deviation. If this is larger than a selected threshold, I skip the column to avoid eventual divergence. For each fitted Gauss, I store the amplitude, center,

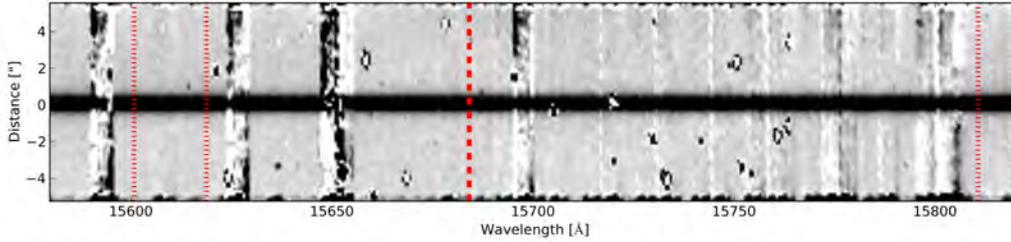


FIGURE 10: 2D-spectrum in the A-direction showing the cut-out around the OIII α line at the DLA absorption redshift. The thick line marks the approximate expected location of the emission line. The thinner lines mark the two sections chosen for building the model SPSF. Also visible in the figure are the noisy edges in the spatial direction found in the A-spectra. These will be cropped off before further processing.

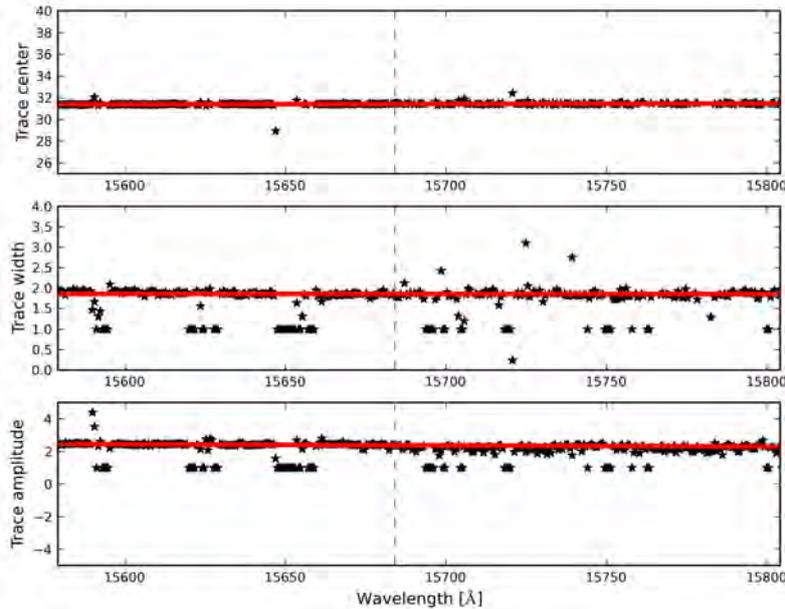


FIGURE 11: The gauss-parameters for the 402 columns cut out around the OIII α -line in the spectrum of angle A. All three parameters have been fitted with a linear function along the wavelength-direction.

and width parameter in three new arrays. All three parameters can be fitted with a linear function as is seen in figure 11.

The SPSF is then built by Gauss-fitting each column across the spatial direction in the two chosen SPSF-regions. Each fitted function is normalised by dividing all pixels with the amplitude parameter from the fit. Afterwards, all the fits are summed and then divided by the total number of columns contributing to the model. The SPSF is subtracted from the image by looping along the columns in a region of ± 100 pixels around the line center. For each column the loop subtracts $\text{SPSF} \times (b + a \times \text{wl})$ where **SPSF** is the column array built above, a and b are the two parameters from the linear fit of the amplitudes produced above, and wl is a wavelength-index array. Plots with the QSO line before and after SPSF-subtraction, for each line and angle, is shown in the results section.

3.2.3 Flux limits

Having subtracted the QSO continuum from the images, the next step is to produce 1D-spectra and compare fluxes with noise. For each image, I place 3 apertures along the spatial direction. Each aperture measures $3''.3$ in the spatial direction. The fluxes of each column are summed together and an 1D-spectra is then plotted for each aperture. I overplot each

spectrum with a with a 3σ detection limit by plotting the square root of the **sig**-array multiplied with 3 for each wavelength-index. The plots are shown and evaluated in the following section. In (Krogager et al., 2016) a velocity spread is calculated for the DLA by fitting the broadening of the different absorption lines. The result is $\Delta V = 331 \pm 30 \text{ km s}^{-1}$. From this result I set the aperture limits in the wavelengths direction as

$$\lambda_{min}^{max} = \lambda_{obs} \pm \frac{165 \text{ km s}^{-1} \lambda_{obs}}{c(1 + z_{dla})} \quad (9)$$

where λ_{obs} is the expected observed wavelength of the line in question. For each aperture all values of the **sig**-array are summed. The square root of this is multiplied with 3 and outputted as a 3σ total flux limit for the given aperture. This concludes the primary part of the spectroscopic data analysis.

4 Results

In following section the results of my study are presented. The primary results are further analyzed and discussed. As with the preceding chapter, imaging and spectroscopy are treated in two separate subsections. However, some overlaps naturally occur.

4.1 Imaging

Figure 12 shows the primary results of the imaging data processing. The first panel is the final processed science frame zoomed on the QSO. Up is north and left is east in this plot and all the following images if nothing else is stated. The second panel is the same zoom after having subtracted the PSF of the QSO. The central part of the QSO is still not accessible after the PSF-subtraction due to strong residuals. No clear DLA emission candidates are seen in the image. However, $\sim 3''$ South-West of the quasar is seen a faint patch that partly overlaps the QSO, and was invisible before PSF-subtraction. To further investigate this structure, I have smoothed the PSF-subtracted image with a gauss smoothing of $4 \times 4 \text{ pixel}^2$ and $\sigma = 1.8$. I have used the IRAF-task **IMSMOOTH** to perform the smoothing. The result is shown in the third panel of figure 12. A red arrow is inserted to mark the location. The patches are clearly seen in the smoothed image, and I find it unlikely that they should be the result of pure noise.

The "fuzzy blob" is also described by Krogager et al. (2016). Here it is compared with a Ly α nebula from the QSO host galaxy, which was observed in radio emission as presented in Barthel et al. (1988). In this article, the quasar is catalogued as 2222+051. The radio contour plot from Barthel is shown in the left panel of figure 13. The shape of the radio emission is somewhat reminiscent of the structures seen in the gauss smoothed figure. To test this further, I've scaled the radio contour plot with the smoothed figure and overplotted the two. The result is seen in the right panel of figure 13. The color scale

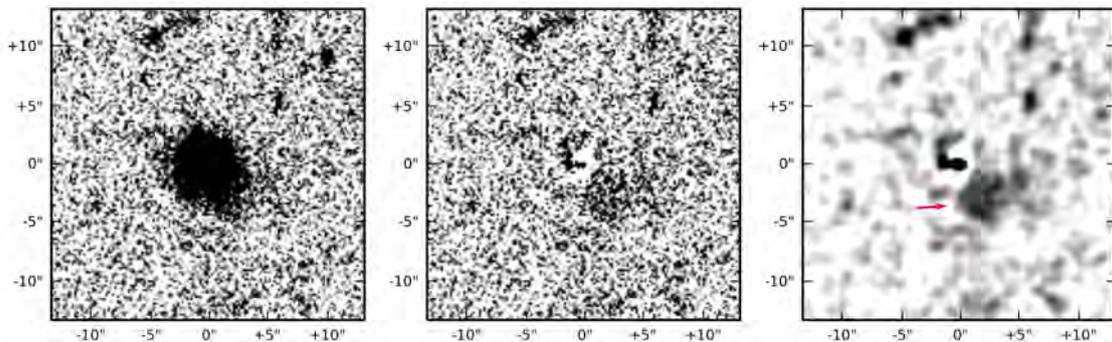


FIGURE 12: The primary results of the imaging data processing. The first panel is the final science frame zoomed on the QSO. The second panel is the image after PSF-subtraction of the QSO. The third panel shows the image after being processed with a $4 \times 4 \text{ pixel}^2$, $\sigma = 1.8$ gauss-smoothing. The red arrow marks the location of the patch described in the text. Up is north and left is east for these images and for all other if nothing else is stated.

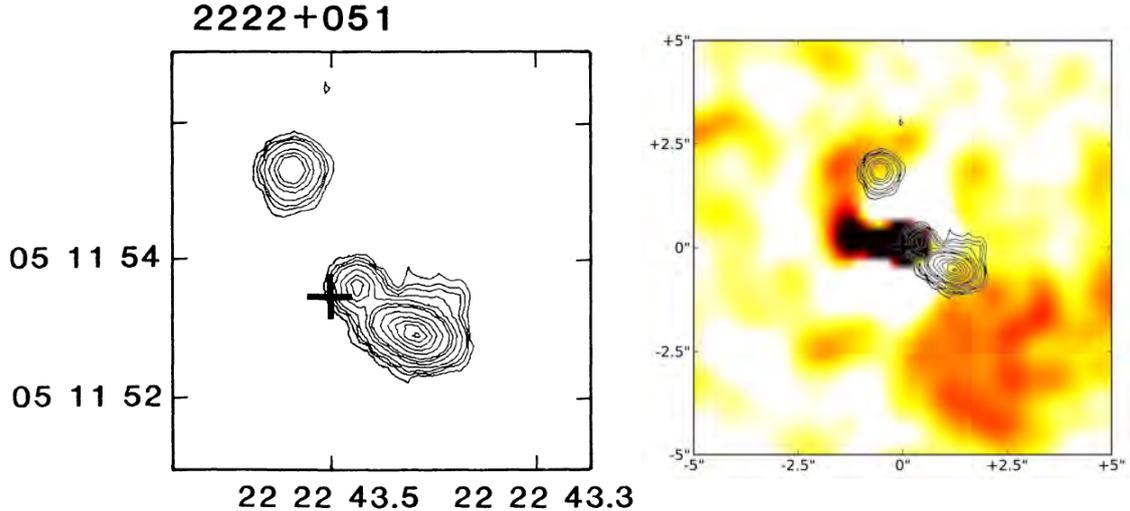


FIGURE 13: *Left panel:* The radio contour plot of the Ly α nebula as presented in Barthel et al. (1988). In this article the QSO is catalogued as 2222+051. *Right panel:* The gauss smoothed image with the South-West emission patch overplotted with the radio contour plot. The two structures do not show the same physical size.

of the QSO figure have been changed from the preceding figures, to make the overplot more clearly visible. It is evident from the figure that the emission patches and the radio contours have very different physical sizes. For this reason, I highly doubt that they should emerge from the same physical object.

To study this new object further, I've gotten hold of high resolution imaging data from the Deimos instrument at the Keck II observatory. The data set was collected by Stockton et al. (2008) who were studying two nearby massive $z = 2.5$ galaxies containing old stellar populations. QSO2225+0527 is by coincident in the field. The adaptive optics and $0''.105/\text{pixel}$ resolution of the Deimos allows for studying the emission patches in greater detail. Figure 14 shows an I-band image of the QSO. The Keck-images are rotated $\sim 90^\circ$ clockwise compared to the NOT-data. In the right panel, the QSO have been subtracted. The subtraction has been based on a spherical model. The QSO is saturated in the Keck-images, and for this reason a true PSF-model cannot be produced. The subtraction has been done by J. P. U. Fynbo. Even without smoothing the images, the faint patches are clearly visible. In this deeper image, the shape of the structure is somewhat reminiscent of a 'T'. A very plausible guess would be that the feature is two antiparallel-aligned disk galaxies seen edge-on.

To further study these objects, I calculate their apparent magnitudes. I use **IRAF** to calculate instrumental magnitude for the entire object choosing an aperture of radius of 20 pixels. I convert from instrumental to apparent magnitudes by conversion with a bright galaxy in the field that is also found in the SDSS-catalogue (RA: 336.317, DEC: 5.456). I use the following standard formula: $m_1 - M_1 = m_2 - M_2$ where m and M correspond to the instrumental and apparent magnitudes respectively. The apparent magnitude of the disks combined in the I -band is: $m_I = 23.56$ mag. By assuming that the disks are found at the absorption redshift, a corresponding absolute magnitude can be calculated:

$$M_{abs} = M - \mu + 2.5 \log(1 + z) \quad (10)$$

where the last term corrects for the fact that we're considering a flux density for a given filter and not a true bolometric magnitude. μ is the distance modulus, which can be calculated as

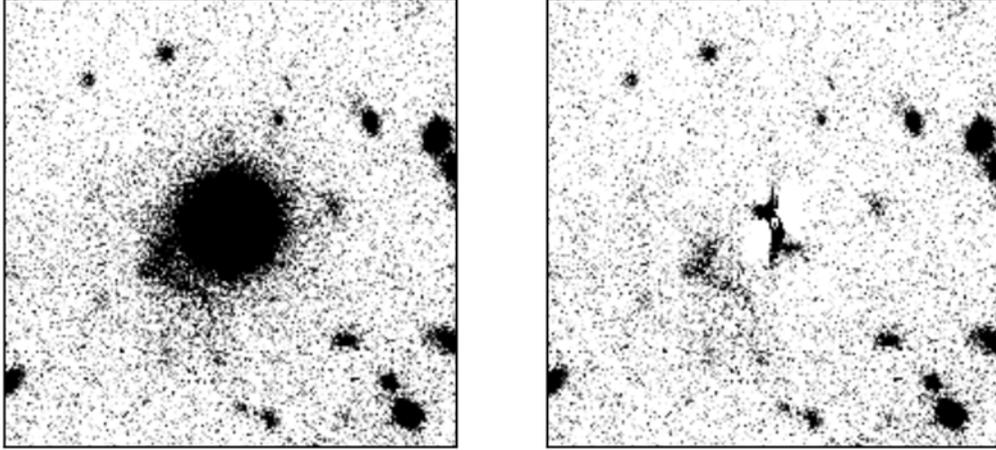


FIGURE 14: The Deimos I-band images, centered on the QSO. In the right panel, the QSO have been subtracted with a spherical model. The images are rotated $\sim 90^\circ$ clockwise compared to the previous images; Up is east and right is north.

$$\mu = 5 \log d_L - 5 \quad (11)$$

d_L is the luminosity distance, which can be found from the redshift by assuming that the disks are found at the absorption redshift, $z_{dla} = 2.13$. For a flat universe, the luminosity distance as a function of redshift is

$$d_L(z) = (1 + z)d_C(z) \quad (12)$$

where d_C is the comoving distance. According to Ryden (2003, p. 106), this can approximated as

$$d_C = \frac{c}{H_0} z \left[1 - \frac{1 + q_0}{2} z \right] \quad (13)$$

I assume a standard Λ CDM-cosmology with $H_0 = 67.8 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $\Omega_\Lambda = 0.693$ and $\Omega_M = 0.307$. This gives the deceleration parameter $q_0 = (1/2)\Omega_M - \Omega_\Lambda = -0.54$. Inserting all values, I find that the absolute magnitude of the disks in the I-band at the absorption redshift is $M_{abs} = -21.38$.

I apply the angular distance relation to estimate the diameter of the disks. For a small angle:

$$d_A = \frac{l}{\delta\phi} \quad (14)$$

where l is the actual physical extent of the object, $\delta\phi$ is the angular extent in radians, and d_A is the cosmological angular-diameter distance. According to Ryden (2003, p. 112), the latter can be given by

$$d_A = \frac{d_L}{(1 + z)^2} \quad (15)$$

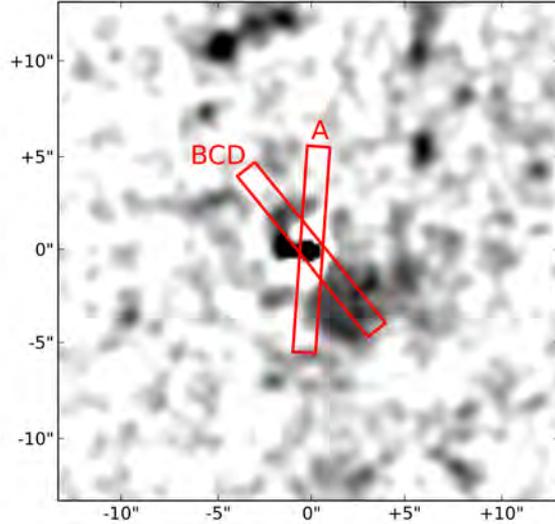


FIGURE 15: The gauss smoothed image overplotted with the location of two X-Shooter slit-angles: A: PA = -4° east of north and BCD: PA $\sim 39^\circ$ east of north. It is seen that the disks are covered by the $-3''.00$ -aperture of the BCD-angle.

The diameter of the disk measures approximately 33 pixels in figure 14. The resolution is $0''.105/\text{pixel}$. Converting this to radians and using equation 15 and 14 yields $l = 25.7$ kpc. This value is approximately the same size as the Milky Way disk, $D_{MW} \sim 30$ kpc. If this is truly the galaxy counterpart, then the physical scale suggest a highly evolved structure resembling our own galaxy. This is in good agreement with what should be expected from the Solar metallicity of the DLA.

4.2 Spectroscopy

In figure 16 to 19 we see the plots produced with the Python code. The figures show a cut-out of the spectra centered on the OIII λ and H α lines in both the A- and BCD-directions. In the two upper panels of all plots, the 2D-spectrum is shown before and after the SPSF-subtraction of the QSO continuum. The three lower panels shows 1D-spectra extracted for three $3''.3$ -apertures along the spatial direction. The 1D-spectra has been overplotted with a 3σ -curve based on the corresponding error-images. The SPSF-subtraction has been less successful around the H α line due to irregularities in the QSO continuum. Unfortunately, the H α line is also coinciding with a skyline.

As seen from figure 15, the disks are partly covered by the $-3''.00$ -aperture of the BCD-angle. However, the flux curves are everywhere below the error curve for both spectral lines except for a few spots coinciding with skylines, cosmic rays, or otherwise erroneous pixels. I've also qualitatively studied the remaining parts of the BCD-spectra for any emission features, but I have not found anything of significance. Even though they do not cover the disks, I have included the figures for the A-angle for the sake of possible detection of emission features not visible in the imaging data. However, as with the BCD-plots these are not showing any emission features.

I have extracted flux limits for the $-3''.00$ -aperture of the BCD-angle: $F_{OIII} < 9.96 \times 10^{-17} \text{erg s}^{-1} \text{cm}^{-2}$ and $F_{H\alpha} < 12.21 \times 10^{-17} \text{erg s}^{-1} \text{cm}^{-2}$. From the value for H α , I can calculate an upper limit on the star formation rate of the disks based on the Kennicutt (1983)-relation:

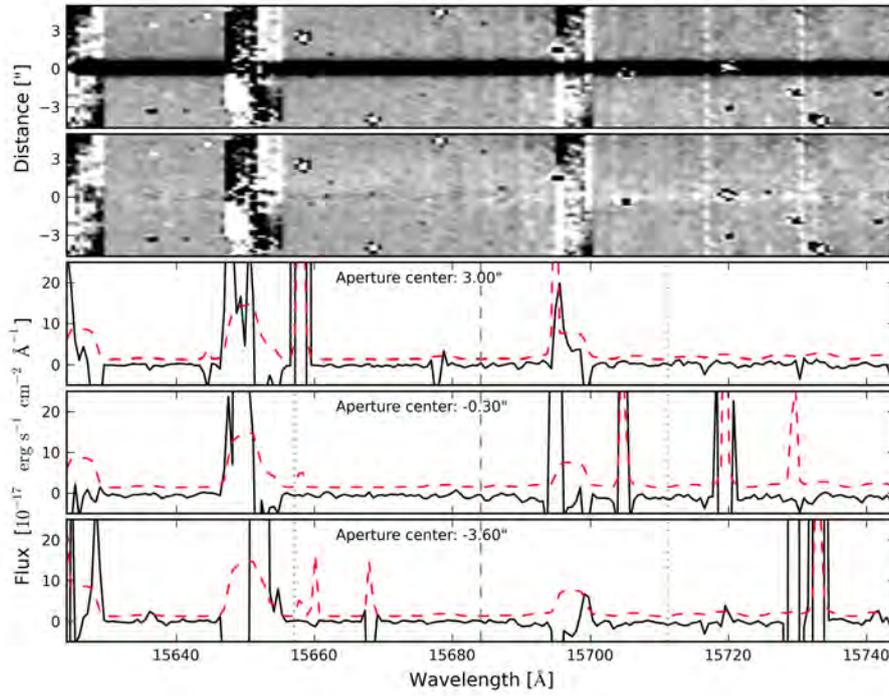


FIGURE 16: The NIR-spectral data in the A-direction, centered on the OIIIa line at the absorption redshift. The two upper panels show the 2D-spectra before and after subtraction of the QSO continuum. The three lower panels show 1D-spectra with fluxes extracted in three apertures with spatial widths of $3''.3$. The two thinner lines mark the spectral widths of the apertures. The spectra have been overplotted with a 3σ detection curve based on the corresponding error images.

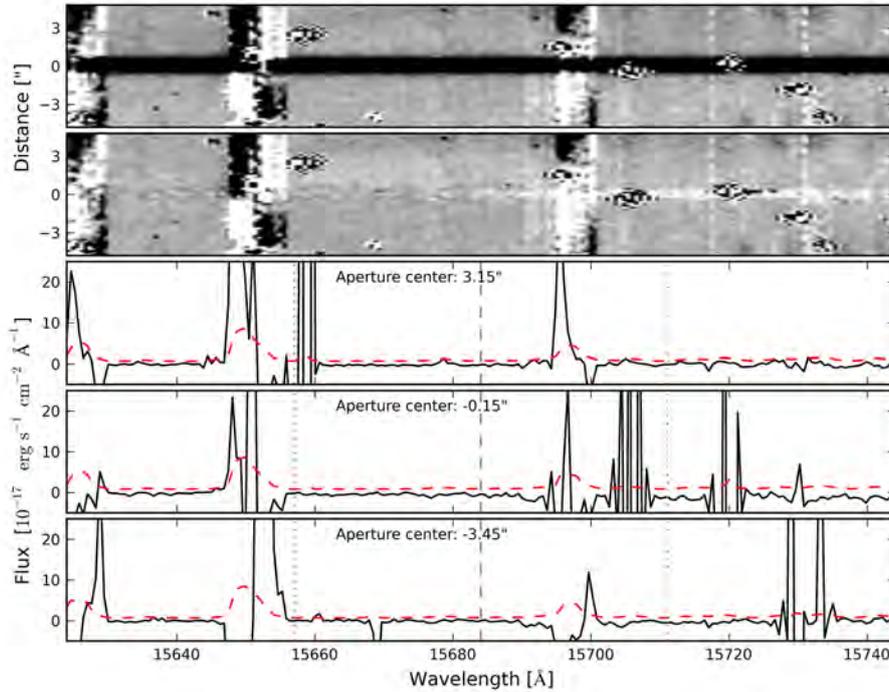


FIGURE 17: Same as figure 16 but in the BCD-direction.

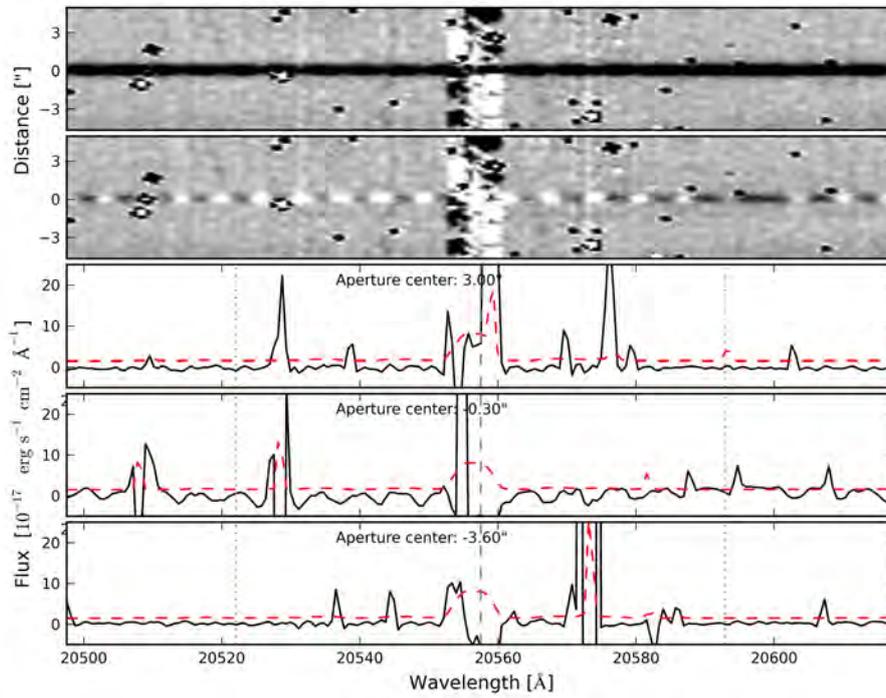


FIGURE 18: Same as figure 16 but centered on $H\alpha$ in the A-direction.

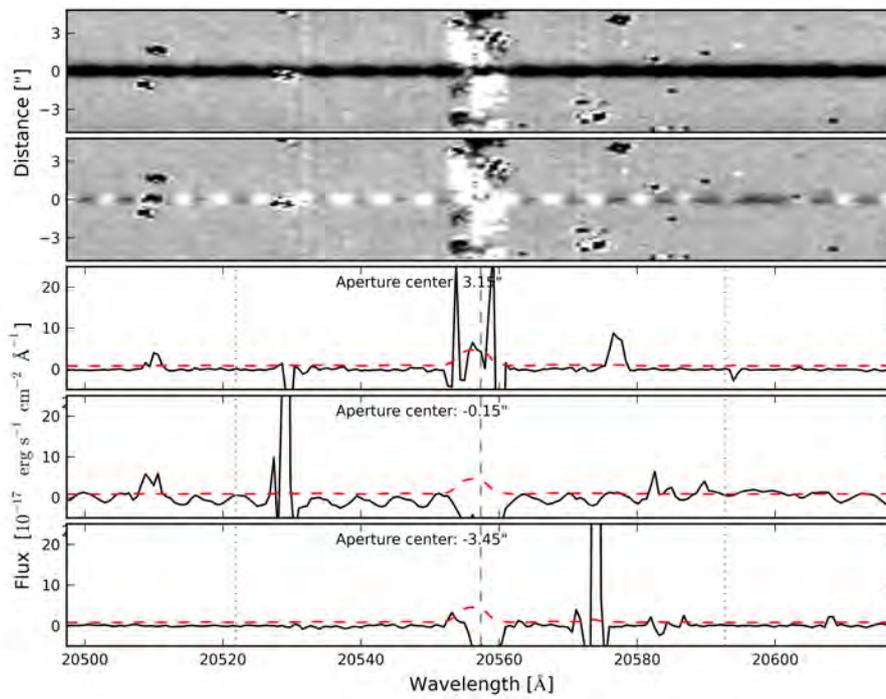


FIGURE 19: same as figure 18 but in the BCD-direction.

$$\text{SFR} = \frac{L(\text{H}\alpha)}{1.12 \times 10^{41} \text{ erg s}^{-1}} \text{M}_{\odot} \text{ yr}^{-1} \quad (16)$$

See appendix B for a derivation and review of the formula. I convert the flux limits to luminosity limits with the inverse-square relation:

$$L(\text{H}\alpha) = 4\pi d_L^2 F_{\text{H}\alpha} \quad (17)$$

D_L is the luminosity distance found with equation 12. Inserting all values and calculating, I get an upper limit for the star formation rate of the disks, $\text{SFR} < 38.55 \text{ M}_{\odot} \text{ yr}^{-1}$. The limit may be even higher since this calculation has not considered either galactic extinction, nor corrected for dust in the DLA. Since the slit is only covering approximately half of the area of the disks, the calculation is only an estimation of order of magnitude. It is also worth stressing that the extracted flux limits for the $\text{H}\alpha$ line is probably influenced by the sky line centered on the expected wavelength.

5 Discussion

Based on a combination of imaging and spectroscopy from various programs, I have presented the discovery of a new object that might constitute as a candidate for the galaxy counterpart of DLA2225+0527. The lack of spectroscopic data for the disks make it impossible to determine a redshift. An emission redshift matching the absorption redshift would make the object a very plausible galaxy counterpart candidate. However, in the lack of redshift I will discuss the physical properties derived under the assumption that the disks lie at the absorption redshift. The derived magnitude of -21.38 mag is almost one order of magnitude brighter than the Milky Way value of ~ -20.5 whereas the derived physical length $l = 25.7$ kpc is of the same order. The values thus suggest that the structures are of the physical size of the Milky Way but brighter. Part of the explanation for this is possibly that the structures appear to be *two* disk galaxies of Milky Way size. The magnitude is a measure of the flux of the two combined. Part of the explanation can also be found by considering the rather high upper limits on the star formation rates. $\text{SFR} < 38.55 \text{ M}_{\odot} \text{ yr}^{-1}$ is far above the value of the Milky Way, $\text{SFR}_{mw} \sim 1.45 \text{ M}_{\odot} \text{ yr}^{-1}$ (Robitaille and Whitney, 2010). This suggests that the disks consists of young bright stars contributing to the magnitude. All in all, I conclude that the derived physical values are not inconsistent with the hypothesis that the presented disks are the galaxy counterpart of the DLA.

The morphology of the disks does not resemble the typical DLA emission counterpart, which are most often found to be confined to single compact cores. In Møller et al. (2002) three confirmed DLA galaxy counterparts are presented and compared to Lyman break galaxies. Out of these three, two of them show irregular structures that are extended or elongated in ways, which can be compared to the morphology of the object presented in this paper. I therefore find that the disks cannot be discarded as a galaxy counterpart candidate based on the morphology only.

Krogager et al. (2016) present a different candidate for the galaxy counterpart found in imaging from Keck. After having subtracted the QSO with a model PSF based on the QSO itself, one compact structure remained in both H and K' images. It lies at an impact parameter of roughly 4 kpc (see figure 20). The fact that the same structure was found in both bands supports the hypothesis that the structure is a real physical object and not just residual noise from the PSF-subtraction. On this ground, Krogager et al. (2016) regard the object a 'tentative' detection of the galaxy counterpart. As with the above presented disks, the object lacks a spectroscopically determined redshift to confirm it as the emission counterpart.

The question that constitutes the title of this project remains unanswered. More precise data is needed in order to firmly detect the DLA emission counterpart. Part of the problem is the extraordinary brightness of the QSO. This makes background subtraction in the spectroscopic data complicated. An X-Shooter-survey with the slits centered on

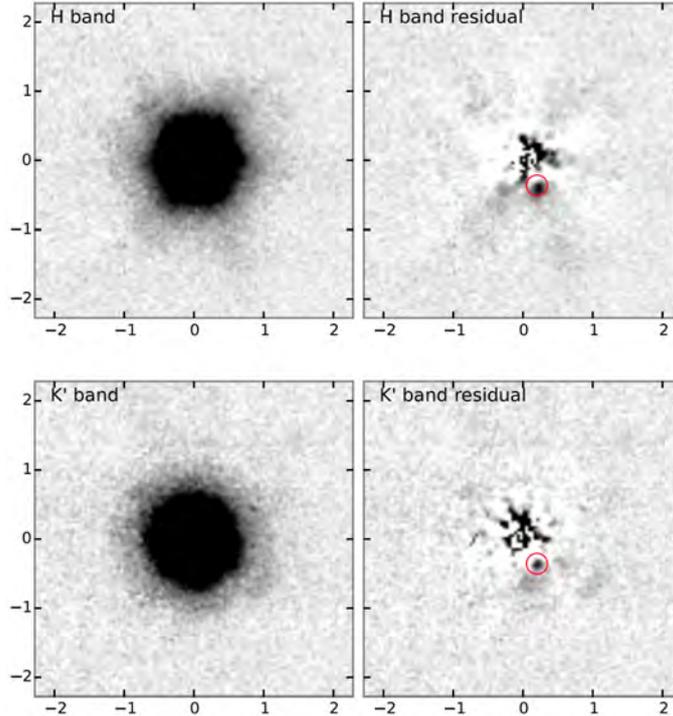


FIGURE 20: The Keck images of the QSO in H and K' bands, from Krogager et al. (2016). The left and right panels show the image before and after PSF-subtraction. The red circle marks the tentative detection of the emission counterpart of the DLA.

the disks should result in some direct spectroscopic emission when considering the fairly bright apparent magnitude of the object. A redshift could possibly be determined from such data making it easy to judge whether or not the object is a serious candidate. Also, the disk spectrum could be compared with the metal lines of the QSO absorption spectrum and a more precise $H\alpha$ flux limit could be established hereby narrowing down the limit on the star formation rate. Should future surveys conclude that the disks are not a plausible DLA counterpart, I argue that the object is still interesting in its own right since it is an undescribed bright object in an interesting field.

The Krogager et al. (2016) tentative emission candidate could be further analyzed by producing high resolution imaging data which are more suitable for performing PSF-subtraction. Spectroscopic data on the object is still needed to verify it as a galaxy candidate. Spectroscopic emission data is mostly likely difficult to obtain considering the small impact parameter and the bright QSO.

It is of course possible that a future survey will reveal the true DLA counterpart to be a third so far unobserved object.

References

- Bahcall, J. N. and Salpeter, E. E. (1965). On the Interaction of Radiation from Distant Sources with the Intervening Medium. *The Astrophysical Journal*, 142:1677–1680.
- Barthel, P. D., Miley, G. K., Schilizzi, R. T., and Lonsdale, C. J. (1988). Observations of the large scale radio structure in high redshift quasars. *Astronomy and Astrophysics Supplement Series*, 73:515–547.
- Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, J., Bahcall, N. A., Brinkmann, J., Connolly, A. J., Csabai, I., Czarapata, P. C., Doi, M., Heckman, T. M., Hennessy, G. S., Ivezić, Z., Knapp, G. R., Lamb, D. Q., McKay, T. A., Munn, J. A., Nash, T., Nichol, R., Pier, J. R., Richards, G. T., Schneider, D. P., Stoughton, C., Szalay, A. S., Thakar, A. R., and York, D. G. (2001). Evidence for Reionization at $z \sim 6$: Detection of a Gunn-Peterson Trough in a $z=6.28$ Quasar. *The Astronomical Journal*, 122(6):2850–2857. arXiv: astro-ph/0108097.
- De Cia, A., Ledoux, C., Savaglio, S., Schady, P., and Vreeswijk, P. M. (2013). Dust-to-metal ratios in damped Lyman-alpha absorbers. Fresh clues to the origins of dust and optical extinction towards gamma-ray bursts. *Astronomy and Astrophysics*, 560:A88.
- Foltz, C. B., Chaffee, Jr., F. H., and Weymann, R. J. (1986). Limits on the surface brightness of the 'Lyman Alpha Disk' absorbers toward the QSOs PHL 957 and MC 1331 + 170. *The Astronomical Journal*, 92:247–249.
- Fynbo, J. P. U., Krogager, J.-K., Venemans, B., Noterdaeme, P., Vestergaard, M., Møller, P., Ledoux, C., and Geier, S. (2013). Optical/Near-infrared Selection of Red Quasistellar Objects: Evidence for Steep Extinction Curves toward Galactic Centers? *The Astrophysical Journal Supplement Series*, 204:6.
- Fynbo, J. P. U., Laursen, P., Ledoux, C., Møller, P., Durgapal, A. K., Goldoni, P., Gullberg, B., Kaper, L., Maund, J., Noterdaeme, P., Östlin, G., Strandet, M. L., Toft, S., Vreeswijk, P. M., and Zafar, T. (2010). Galaxy counterparts of metal-rich damped Ly-alpha absorbers - I. The case of the $z = 2.35$ DLA towards Q2222-0946. *Monthly Notices of the Royal Astronomical Society*, 408:2128–2136.
- Fynbo, J. P. U., Ledoux, C., Noterdaeme, P., Christensen, L., Møller, P., Durgapal, A. K., Goldoni, P., Kaper, L., Krogager, J.-K., Laursen, P., Maund, J. R., Milvang-Jensen, B., Okoshi, K., Rasmussen, P. K., Thorsen, T. J., Toft, S., and Zafar, T. (2011). Galaxy counterparts of metal-rich damped Ly-alpha absorbers - II. A solar-metallicity and dusty DLA at $z = 2.58$. *Monthly Notices of the Royal Astronomical Society*, 413:2481–2488.
- Fynbo, J. P. U., Prochaska, J. X., Sommer-Larsen, J., Dessauges-Zavadsky, M., and Møller, P. (2008). Reconciling the Metallicity Distributions of Gamma-Ray Burst, Damped Ly-alpha, and Lyman Break Galaxies at $z \sim 3$. *The Astrophysical Journal*, 683:321–328.
- Fynbo, J. U., Møller, P., and Warren, S. J. (1999). Extended LY alpha emission from a damped LY alpha absorber at $z=1.93$, and the relation between damped LY alpha absorbers and Lyman-break galaxies. *Monthly Notices of the Royal Astronomical Society*, 305:849–858.
- Fynbo, J. U., Thomsen, B., and Møller, P. (2000). LY alpha emission from a lyman limit absorber at $z=3.036$. *Astronomy and Astrophysics*, 353:457–464.
- Gunn, J. E. and Peterson, B. A. (1965). On the Density of Neutral Hydrogen in Inter-galactic Space. *The Astrophysical Journal*, 142:1633–1641.

- Howell, S. B. (2006). *Handbook of CCD Astronomy*. Cambridge University Press.
- Kennicutt, Jr., R. C. (1983). The rate of star formation in normal disk galaxies. *The Astrophysical Journal*, 272:54–67.
- Kohler, K. and Gnedin, N. Y. (2007). Lyman Limit Systems in Cosmological Simulations. *The Astrophysical Journal*, 655:685–690.
- Krogager, J.-K., Fynbo, J. P. U., Noterdaeme, P., Zafar, T., Møller, P., Ledoux, C., Krühler, T., and Stockton, A. (2016). A quasar reddened by a sub-parsec sized, metal-rich and dusty cloud in a damped Lyman-alpha absorber at $z=2.13$. *Monthly Notices of the Royal Astronomical Society*, 455(3):2698–2711. arXiv: 1510.04695.
- Krogager, J.-K., Geier, S., Fynbo, J. P. U., Venemans, B. P., Ledoux, C., Møller, P., Noterdaeme, P., Vestergaard, M., Kangas, T., Pursimo, T., Saturni, F. G., and Smirnova, O. (2015). The High AV Quasar Survey: Reddened Quasi-Stellar Objects Selected from Optical/Near-Infrared Photometry—II. *The Astrophysical Journal Supplement Series*, 217:5.
- Kulkarni, V. P. and Fall, S. M. (2002). Metallicity Evolution of Damped Lyman Alpha Galaxies. *The Astrophysical Journal*, 580:732–741.
- Kulkarni, V. P., Fall, S. M., Lauroesch, J. T., York, D. G., Welty, D. E., Khare, P., and Truran, J. W. (2005). Hubble Space Telescope Observations of Element Abundances in Low-Redshift Damped Lyman Alpha Galaxies and Implications for the Global Metallicity-Redshift Relation. *The Astrophysical Journal*, 618:68–90.
- Levshakov, S. A., Agafonova, I. I., Centurion, M., and Molaro, P. (2003). Extremely metal-poor Lyman limit system at $z = 2.917$ toward the quasar HE 0940-1050. *Astronomy and Astrophysics*, 397(3):851–857. arXiv: astro-ph/0210619.
- Møller, P., Fynbo, J. P. U., and Fall, S. M. (2004). Detection of Lyman-alpha emission from a DLA galaxy: Possible implications for a luminosity-metallicity relation at $z = 2-3$. *Astronomy and Astrophysics*, 422:L33–L37.
- Møller, P. and Warren, S. J. (1993). Emission from a damped Ly-alpha absorber at $Z = 2.81$. *Astronomy and Astrophysics*, 270:43–52.
- Møller, P., Warren, S. J., Fall, S. M., Fynbo, J. U., and Jakobsen, P. (2002). Are high redshift damped lyman alpha galaxies lyman break galaxies? *The Astrophysical Journal*, 574:51–58.
- Mo, H., van den Bosch, F., and White, S. (2010). *Galaxy Formation and Evolution*. Cambridge University Press.
- Noterdaeme, P., Ledoux, C., Petitjean, P., and Srianand, R. (2008). Molecular hydrogen in high-redshift damped Lyman-alpha systems: the VLT/UVES database. *Astronomy and Astrophysics*, 481:327–336.
- Noterdaeme, P., Petitjean, P., Ledoux, C., and Srianand, R. (2009). Evolution of the cosmological mass density of neutral gas from Sloan Digital Sky Survey II - Data Release 7. *Astronomy and Astrophysics*, 505:1087–1098.
- Pawlik, A. H., Milosavljević, M., and Bromm, V. (2011). The First Galaxies: Assembly of Disks and Prospects for Direct Detection. *The Astrophysical Journal*, 731:54.

- Pei, Y. C. and Fall, S. M. (1995). Cosmic Chemical Evolution. *The Astrophysical Journal*, 454:69.
- Petitjean, P. (1998). QSO Absorption Line Systems. *arXiv:astro-ph/9810418*. arXiv: astro-ph/9810418.
- Petitjean, P., Webb, J. K., Rauch, M., Carswell, R. F., and Lanzetta, K. (1993). Evidence for structure in the H I column density distribution of QSO absorbers. *Monthly Notices of the Royal Astronomical Society*, 262:499–505.
- Prochaska, J. X. (1999). The Physical Nature of the Lyman-Limit Systems. *The Astrophysical Journal Letters*, 511:L71–L74.
- Rauch, M. (1998). The Lyman Alpha Forest in the Spectra of QSOs. *Annual Review of Astronomy and Astrophysics*, 36:267–316.
- Robitaille, T. P. and Whitney, B. A. (2010). The present-day star formation rate of the Milky-Way determined from Spitzer detected young stellar objects. *The Astrophysical Journal*, 710(1):L11–L15. arXiv: 1001.3672.
- Ryden, B. (2003). *Introduction to Cosmology*. Addison Wesley.
- Schaerer, D. (2003). The transition from Population III to normal galaxies: Lyalpha and He II emission and the ionising properties of high redshift starburst galaxies. *Astronomy and Astrophysics*, 397:527–538.
- Scheuer, P. A. G. (1965). A Sensitive Test for the Presence of Atomic Hydrogen in Inter-galactic Space. *Nature*, 207:963.
- Shklovskii, I. S. (1965). Physical Conditions in the Gaseous Envelope of 3c-273. *Soviet Astronomy*, 8:638.
- Sparke, L. S. and Gallagher, J. S. (2007). *Galaxies in the Universe, 2nd Edition*. Cambridge University Press.
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., and Adelberger, K. L. (1996). Spectroscopic Confirmation of a Population of Normal Star-forming Galaxies at Redshifts $Z > 3$. *The Astrophysical Journal Letters*, 462:L17.
- Stockton, A., McGrath, E., Canalizo, G., Iye, M., and Maihara, T. (2008). Morphologies of Two Massive Old Galaxies at $z \sim 2.5$. *The Astrophysical Journal*, 672:146–152.
- Vernet, J., Dekker, H., D’Odorico, S., Kaper, L., Kjaergaard, P., Hammer, F., Randich, S., Zerbi, F., Groot, P. J., Hjorth, J., Guinouard, I., Navarro, R., Adolfse, T., Albers, P. W., Amans, J.-P., Andersen, J. J., Andersen, M. I., Binetruy, P., Bristow, P., Castillo, R., Chemla, F., Christensen, L., Conconi, P., Conzelmann, R., Dam, J., de Caprio, V., de Ugarte Postigo, A., Delabre, B., di Marcantonio, P., Downing, M., Elswijk, E., Finger, G., Fischer, G., Flores, H., François, P., Goldoni, P., Guglielmi, L., Haigron, R., Hanenburg, H., Hendriks, I., Horrobin, M., Horville, D., Jessen, N. C., Kerber, F., Kern, L., Kiekebusch, M., Kleszcz, P., Klougart, J., Kragt, J., Larsen, H. H., Lizon, J.-L., Lucuix, C., Mainieri, V., Manuputy, R., Martayan, C., Mason, E., Mazzoleni, R., Michaelsen, N., Modigliani, A., Moehler, S., Møller, P., Norup Sørensen, A., Nørregaard, P., Péroux, C., Patat, F., Pena, E., Pragt, J., Reinero, C., Rigal, F., Riva, M., Roelfsema, R., Royer, F., Sacco, G., Santin, P., Schoenmaker, T., Spano, P., Sweers, E., Ter Horst, R., Tintori, M., Tromp, N., van Dael, P., van der Vliet, H., Venema, L., Vidali, M., Vinther, J., Vola, P., Winters, R., Wistisen, D., Wulterkens, G., and Zacchei, A. (2011).

- X-shooter, the new wide band intermediate resolution spectrograph at the ESO Very Large Telescope. *Astronomy and Astrophysics*, 536:A105.
- Wolfe, A. M., Gawiser, E., and Prochaska, J. X. (2005). Damped Ly Alpha Systems. *Annual Review of Astronomy and Astrophysics*, 43:861–918.
- Wolfe, A. M., Turnshek, D. A., Smith, H. E., and Cohen, R. D. (1986). Damped Lyman-alpha absorption by disk galaxies with large redshifts. I - The Lick survey. *The Astrophysical Journal Supplement Series*, 61:249–304.
- Womble, D. S., Sargent, W. L. W., and Lyons, R. S. (1996). *Cold Gas at High Redshift*.

Appendices

A Python code pipeline

```
import pyfits as pf
import numpy as np
from scipy.optimize import curve_fit
import matplotlib.pyplot as pl
import matplotlib.gridspec as gs
import copy

pl.close('all')

# define Gaussian fit function
def fit_gauss(x, a0, a1, a2):
    return a0*np.exp(-(x - a1)**2/(2*a2**2))

# define linear function
def fit_lin(x, a0, a1):
    return a0+a1*x

##### MAIN #####

#line = 'OIII'      # setting emission line
line = 'Halpha'

angle = 'A'        # setting angle
#angle = 'BCD'

if angle == 'A':
    fitsdata = 'flux2d_NIR_A.fits'
    factor = 10
else:
    fitsdata = 'BCD_combined_NIR.fits'
    factor = 1

image = pf.getdata(fitsdata)*1.e17 # read fits image
header = pf.getheader(fitsdata) # read fits header
sig = pf.getdata(fitsdata,1)*1.e17 # read fits std. dev.
variance = sig**2 # std. dev. into variance

if angle == 'A':
    image = image[4:69,:]
    sig = sig[4:69,:]
else:
    image = image[:,:]
    sig = sig[:,:]

z = 2.132487 # redshift of absorption lines
c = 299792. # speed of light in vacuum, km/s
vrad = header['ESO_QC_VRAD_HELICOR'] # radial velocity, from header
d_s = (1 + vrad/c) # dopplershift of absorption lines
wl_low = header['CRVAL1']*factor # lowest wavelength in Angstrom, from header
wl_bin = header['CD1_1']*factor # delta-wavelength pr. bin in Angstrom, from header
arcsec_bin = header['CD2_2'] # delta-distance pr. bin in arcseconds, from header
xaxis = header['NAXIS1'] # length of x-axis
yaxis = ['NAXIS2'] # length of y-axis

wl_obs = d_s*np.arange(image.shape[1])*wl_bin+wl_low # observed wavelengths in Angstrom
dist_obs = dist_bin*np.arange(image.shape[0]) # distance in arcsecs

velocity_span = 330/2 # velocity spread from Krogager article, +/- center in km/s

if line == 'OIII':
    xline = 9574 # approximate OIIIa line (in pixels)
```

```

x1psf1 = 9435 # start psf 1 (in pixels)
x2psf1 = 9465 # end psf 1 (in pixels)
x1psf2 = 9785 # start psf 2 (in pixels)
x2psf2 = 9800 # end psf 2 (in pixels)
colstart = 9400 # begin of data crop (in pixels)
ncols = 402 # length of data crop (in pixels)

else:
xline = 17697 # approximate Halpha line (in pixels)
x1psf1 = 17625 # start psf 1 (in pixels)
x2psf1 = 17640 # end psf 1 (in pixels)
x1psf2 = 17765 # start psf 2 (in pixels)
x2psf2 = 17780 # end psf 2 (in pixels)
colstart = 17580 # begin of data crop (in pixels)
ncols = 402 # length of data crop (in pixels)

wl_min = wl_obs[xline] - velocity_span*wl_obs[xline]/c*(1+z) # defining wavelength limits of flux apertures
wl_max = wl_obs[xline] + velocity_span*wl_obs[xline]/c*(1+z) # from velocity spread

amps = np.zeros(ncols) # zero-value arrays of length ncol
centers = np.zeros(ncols)
widths = np.zeros(ncols)

# beginning loop to fit gauss

for colnr in range(colstart, colstart+ncols-1):
    pl.figure() #starting new plot
    pix = np.arange(yaxis) #array with numbers 1 to height of column
    pl.plot(pix, image[pix,colnr], 'r+') #plotting fluxes of column nr colnr
    ii = np.where((pix > 20) & (pix < 45)) #approximate index of quasar line top and bottom
    n = len(pix[ii])
    mean = sum(pix[ii]*image[pix[ii],colnr])/n #calculating mean
    sigma = sum(image[pix[ii],colnr]*(pix[ii]-mean)**2)/n #calculating standard deviation
    print(sigma)
    print(colnr)

    if 0 < sigma < 200: #testing if standard deviation is too large for the function to converge
        [fitpar, covar] = curve_fit(fit_gauss, pix[ii], image[pix[ii],colnr], p0=[max(image[pix[ii],colnr]),mean,sigma])
        a = fit_gauss(pix[ii],*fitpar) #gauss fitting across the quasar line in the column
        amps[colnr-colstart] = fitpar[0] #saving amplitude parameter
        centers[colnr-colstart] = fitpar[1] #saving center parameter
        widths[colnr-colstart] = abs(fitpar[2]) #saving width parameter
        pl.plot(pix[ii],a,'k') #overplotting gaussian function
        pl.plot(pix[ii],image[pix[ii],colnr]-a,'b+') #plotting residuals

    else:
        amps[colnr-colstart] = 1 #setting empty values
        centers[colnr-colstart] = 1
        widths[colnr-colstart] = 1
        print('Failed_to_converge')

# calculating medians
amps_med = round(np.median(amps))
widths_med = round(np.median(widths))
centers_med = round(np.median(centers))

pl.figure() # starting new figure

wl = np.arange(ncols)+colstart # wavelength index array

# plotting gaussian center for each column
pl.subplot(311)
pl.plot(wl_obs[wl],centers,'k*')
pl.ylim(25,40)
pl.xlim(wl_obs[xline]-175, wl_obs[xline]+200)
pl.ylabel('Trace_center', fontsize="10")
pl.axvline(wl_obs[xline], ls='--', color='gray')
pl.yticks(fontsize="8")
pl.xticks(fontsize="8")

```

```

# finding index of xvalues not too close to xline and centers not far off and amplitudes not too big or low
ii1 = list(np.where((wl < xline-15) & (centers > 30) & (centers < 40) & (amps > 0) & (amps < 4)))
ii2 = list(np.where((wl > xline+15) & (centers > 30) & (centers < 40) & (amps > 0) & (amps < 4)))

ii = list(ii1[0]) + list(ii2[0])

# linear fit through centers in the region chosen above
[fitpar_center, covar] = curve_fit(fit_lin, wl[ii], centers[ii])
a = fit_lin(wl,*fitpar_center)
pl.plot(wl_obs[wl],a,'r', linewidth=3) #overplotting with linear fit

# plotting gaussian width for each column
pl.subplot(312)
pl.plot(wl_obs[wl],widths,'k*')
pl.ylim(0,4)
pl.xlim(wl_obs[xline-175],wl_obs[xline+200])
pl.ylabel('Trace_width', fontsize="10")
pl.axvline(wl_obs[xline], ls='--', color='gray')
pl.yticks(fontsize="8")
pl.xticks(fontsize="8")

# linear fit through width
[fitpar_width, covar] = curve_fit(fit_lin, wl[ii], widths[ii])
b = fit_lin(wl,*fitpar_width)
pl.plot(wl_obs[wl],b,'r',linewidth=3) #overplotting with linear fit

# plotting gaussian amplitude for each column
pl.subplot(313)
pl.plot(wl_obs[wl],amps,'k*')
pl.ylim(-5,5)
pl.xlim(wl_obs[xline-175],wl_obs[xline+200])
pl.ylabel('Trace_amplitude', fontsize="10")
pl.xlabel('Wavelength_[$\\AA$]', fontsize="10")
pl.axvline(wl_obs[xline], ls='--', color='gray')
pl.yticks(fontsize="8")
pl.xticks(fontsize="8")

# linear fit through amplitude
[fitpar_amp, covar] = curve_fit(fit_lin, wl[ii], amps[ii])
c = fit_lin(wl,*fitpar_amp)
pl.plot(wl_obs[wl],c, 'r',linewidth=3) #overplotting with linear fit

pl.savefig('plot0.png', format='png', dpi=600, bbox_inches='tight')

# summing all flux of chosen psf-regions
ii1 = list(np.where((wl > x1psf1) & (wl < x2psf1) ) #index of wavelengths of psf1
ii = ii1[0]
nsum = 0
countsum = 0

# summing total flux approximately around continuum, for each column in region of psf1
for n in range(0,len(ii)-1):
    countsum = countsum + sum(image[30:45,ii[n]])
    nsum=nsum+1.

print(countsum/nsum) # average column sum in region

# building PSF 1
PSF = np.zeros(yaxis)
nsum = 0
for ncol in range(x1psf1,x2psf1): #fitting gauss-function along column for each colum in region psf1
    pix = np.arange(yaxis)
    ii = np.where((pix > 20) & (pix < 45))
    n = len(pix[ii])
    mean = sum(pix[ii]*image[pix[ii],ncol])/n
    sigma = sum(image[pix[ii],ncol]*(pix[ii]-mean)**2)/n
    [fitpar, covar] = curve_fit(fit_gauss, pix[ii], image[pix[ii],ncol], p0=[max(image[pix[ii],ncol]),mean,sigma])
    a = fit_gauss(pix[ii],*fitpar) # gauss fitting across the quasar line in the column
    if fitpar[0] > 1: # test if amplitude is greater than 1

```

```

    PSF = PSF+image[:,ncol]/fitpar[0] # if so, pixel values are divided by column-amplitude and added to total psf
    nsum=nsum+1

# building PSF 2

for ncol in range(x1psf2,x2psf2):
    pix = np.arange(yaxis)
    ii = np.where((pix > 20) & (pix < 45))
    n = len(pix[ii])
    mean = sum(pix[ii]*image[pix[ii],ncol])/n
    sigma = sum(image[pix[ii],ncol]*(pix[ii]-mean)**2)/n
    [fitpar, covar] = curve_fit(fit_gauss, pix[ii], image[pix[ii],ncol], p0=[max(image[pix[ii],ncol]),mean,sigma])
    a = fit_gauss(pix[ii],*fitpar)
    if fitpar[0] > 1:
        PSF = PSF+image[:,ncol]/fitpar[0]
        nsum=nsum+1

PSF = PSF/nsum # dividng total psf with number of coloumns to get average column-psf

pl.figure() # previewing the model PSF
pl.plot(PSF)
pl.title('Average_PSF')

image_res = copy.copy(image) # copying image data

# for each column substract psf times amplitude function to get residual image
for colnr in range(xline-100,xline+100): # looping around xline-100 to xline+100
    image_res[:,colnr] = image[:,colnr]-1.00*PSF*(fitpar_amp[0]+fitpar_amp[1]*wl[colnr-colstart])

pf.writeto('test0IIIa.fits', image_res) # write the resulting fits

# cropping images
image_crop = image[:,xline-100:xline+100]
image_res_crop = image_res[:,xline-100:xline+100]
variance_crop = variance[:,xline-100:xline+100]

# beginning to determine flux limits
n_aper = 3 # number af apertures
size_aper = np.round(yaxis/n_aper)
size_aper_arcsec = size_aper*header ['CD2_2']

pl.figure()
gs1 = gs.GridSpec(n_aper + 2,1)
gs1.update(wspace=0.025, hspace=0.05) # set the spacing between axes.

ax1 = pl.subplot(gs1[0]) # plotting image with quasar
pl.imshow(image_crop, vmin=-0.2, vmax=0.2*amps_med,
          cmap=pl.cm.gray_r, origin='lower', aspect='auto',
          extent=[wl_obs[colstart],wl_obs[colstart+ncols],dist_obs[0]-dist_obs[centers_med],dist_obs[yaxis-1]-dist_obs[centers_med])
pl.xticks([])
pl.yticks([-3,0,3])

ax1 = pl.subplot(gs1[1]) # plotting image with quasar substracted
pl.imshow(image_res_crop, vmin=-0.2, vmax=0.2*amps_med,
          cmap=pl.cm.gray_r, origin='lower', aspect='auto',
          extent=[wl_obs[colstart],wl_obs[colstart+ncols],dist_obs[0]-dist_obs[centers_med],dist_obs[yaxis-1]-dist_obs[centers_med])
pl.xticks([])
pl.yticks([-3,0,3])
pl.ylabel('Distance_["]', fontsize="10")

i = 0
for i in range(n_aper): # looping around number of apertures

    aperture_image = image_res_crop[(n_aper-i-1)*size_aper:size_aper*(n_aper-i)-1,:] # cropping image
    aperture_variance = variance_crop[(n_aper-i-1)*size_aper:size_aper*(n_aper-i)-1,:] # cropping error image

    mask = np.ones_like(aperture_image)

    sum_image = np.sum(aperture_image*mask, axis=0) # summing columns
    sum_variance = np.sum(aperture_variance*mask, axis=0) # summings columns of error image

```

```

sum_sigma = 3*np.sqrt(sum_variance)                    # converting to sigma 3

ax1 = pl.subplot(gs1[i + 2])                          # plotting
flux = pl.plot(wl_obs[xline-100:xline+100],sum_image,'k')
stddev = pl.plot(wl_obs[xline-100:xline+100],sum_sigma,'r--')
pl.axvline(wl_obs[xline], ls='--', color='gray')
pl.axvline(wl_min, ls=':', color='gray')              # drawing aperture limits
pl.axvline(wl_max, ls=':', color='gray')
pl.xlim(wl_obs[xline-100],wl_obs[xline+100])
center_value = -(i*(size_aper)+(size_aper/2))*header['CD2_2']+centers_med*header['CD2_2']
pl.text(wl_obs[xline-35],20,"Aperture_center:_.2f\" %center_value, fontsize="8")
pl.yticks([0, 10, 20],fontsize="8")
pl.xticks(fontsize="8")
pl.ylim(-5,25)

if i == 1:
    pl.ylabel('Flux_{\AA} [10^{-17}] \, \, {\rm erg} \, \, s^{-1} \, \, cm^{-2} \, \, \AA^{-1}]$', fontsize="10")
else:
    pl.ylabel('')

# extracting flux from apertures
aperture_wavelength = (wl_obs[xline-100:xline+100]>wl_min)*(wl_obs[xline-100:xline+100]<wl_max)
line_variance = np.sum(sum_variance[aperture_wavelength])*header['CD1_1'] # converting to flux units
total_sigma = np.sqrt(line_variance)
print("Flux_{\AA}_.2f_{10}^{-17} \, \, erg/s/cm^2/\AA_{(3_\sigma)}" %(total_sigma*3)) # printing out results

pl.xlabel('Wavelength_{\AA}$', fontsize="10")
pl.savefig('plot1.png', format='png', dpi=600, bbox_inches='tight')
pl.show() # showing all plots

```

B Star formation rates with H α and Ly α

The first article to quantitatively estimate star formation rates for entire galaxies based on the integrated emission of H α was Kennicutt (1983). The author calculated Lyman continuum flux models for a range of stellar models. The models were integrated over the entire lifetime of each star to yield the total output. These models were used to produce a grid of galaxy models based on three different assumed initial mass functions and an exponentially declining star formation rate:

$$R(t) = R_0 e^{-\beta t} \quad (0 \leq \beta \leq \infty) \quad (18)$$

H α equivalent widths were calculated from the Lyman continuum flux models using the following expression:

$$EW(\text{H}\alpha) = \frac{1.36 \times 10^{-12} N_{\text{Lyman}}}{f_\lambda(6563 \text{ \AA})} \quad (19)$$

The author produced a plot of $(B - V)$ -indices and H α equivalent widths for the grid of galaxies. This was overplotted with real data points from earlier work. The plot clearly suggested that the most plausible IMF was the so-called 'extended' Miller-Scalo initial mass function, which is constant above 1 solar mass:

$$\psi(m) \propto m^{-1.4} \quad (0.1 < m < 1 M_\odot) \quad (20)$$

$$\psi(m) \propto m^{-2.5} \quad (1 < m < 100 M_\odot) \quad (21)$$

By adopting this IMF, Kennicutt (1983) reduced the relationship between H α luminosity and total star formation rates to the following surprisingly simple relation:

$$\text{SFR} = \frac{L(\text{H}\alpha)}{1.12 \times 10^{41} \text{ erg s}^{-1}} M_\odot \text{ yr}^{-1} \quad (22)$$

Kennicutt (1983) warns that the model doesn't take account of dust, which is the biggest contribution to uncertainty. Also, the model relies on the assumption that the galaxies are Lyman radiation-bounded. However, some ultraviolet radiation might escape meaning that the calculated SFRs should only be interpreted as lower limits. The model assumes that supermassive stars don't contribute significantly to the ionization (see equation 21). This assumption was justified in the article. The last possible limitation of equation 22 is the fact that it is strongly sensitive to variations in the IMF.

The luminosities of Ly α and H α are related by a simple scaling law, which for example can be seen from table 1 and 4 in Schaerer (2003) (noted by Pawlik et al. (2011)). Approximating this scaling as $L(\text{Ly}\alpha)/L(\text{H}\alpha) = 10$ equation 22 can be rewritten in a form including the luminosity of Ly α (Fynbo et al., 2000):

$$\text{SFR} = \frac{L(\text{Ly}\alpha)}{1.12 \times 10^{42} \text{ erg s}^{-1}} M_\odot \text{ yr}^{-1} \quad (23)$$