The nature and evolution of Lyman-α emitters over large co-moving volumes

Sérgio Miguel da Graça Santos

Physics

Department of Physics

Lancaster University

June 2020

A thesis submitted to Lancaster University for the degree of Doctor of Philosophy in the Faculty of Science and Technology

Supervised by Dr. David Sobral

Abstract

How do young star-forming galaxies evolve with cosmic time? Extensive work has been done over the past years on the discovery and follow-up of numerous galaxies at the highest redshifts (look-back times). However, most surveys aim for deep observations in small volumes which often are not enough to detect the brightest (but also rarer) sources. Comparing multiple studies/samples in an unbiased way can also be challenging due to different selection methods. A large volume program, which probes multiple look-back times using the same method, would greatly contribute to our understanding of how galaxies evolve in the distant Universe.

In this thesis, we use 12(+4) medium(+narrow) band images to identify ~ 4000 $z \sim 2 - 6$ Lyman-α emitters (LAEs) over 2 deg^2 in the COSMOS field, producing the SC4K (Slicing COSMOS with 4k LAEs) sample. Lyman-α (Lyα) emission is typically associated with young star-forming galaxies (but also active galactic nuclei, AGNs). We use these ~ 4000 LAEs to produce a 3D map of the early Universe, in 16 individual redshift slices, providing a unique sample to explore galaxy evolution, with all sources being selected using the same self-consistent selection method. We construct Lyα luminosity functions (LFs) with our LAEs selected over a very wide co-moving volume (10^8 Mpc^3), complementing ultra-deep surveys. The Schechter component of the Lyα LF reveals a $\sim 5\times$ rise in L$^\star_{\mathrm{Ly}\alpha}$ and a $\sim 7\times$ decline in $\Phi^\star_{\mathrm{Ly}\alpha}$ from $z \sim 2$ to 6. At $z \sim 2 - 3$ we find an excess in number densities at high luminosities (L$_{\mathrm{Ly}\alpha} > 10^{43.3}\,\mathrm{erg\,s^{-1}}$) which is consistent with a higher AGN fraction at those luminosities. This excess is not detected or falls below our detection limits at $z > 4$.

We measure a Lyα luminosity density increase by a factor of ~ 2 from $z \sim 2$ to 3, which then remains constant to $z \sim 6$, which contrasts the ultraviolet (UV) luminosity density decrease at the same redshift ranges. The Lyα/UV luminosity density ratio rises from 4% to 30% from $z \sim 2.2$ to 6.

We conduct aperture photometry for individual SC4K LAEs, using 34 bands of deep multi-wavelength data in the COSMOS field from rest-frame UV to far-infrared (FIR), to measure their individual spectral energy distributions (SEDs). We find typical stellar masses $10^{9.3\pm0.6}\,\mathrm{M_\odot}$ and star formation rates (SFR) of $\mathrm{SFR_{SED}} = 4.4^{+10.5}_{-2.4}\,\mathrm{M_\odot\,yr^{-1}}$ and $\mathrm{SFR_{Ly\alpha}} = 5.9^{+6.3}_{-2.6}\,\mathrm{M_\odot\,yr^{-1}}$, combined with very blue UV slopes of $\beta = -2.1^{+0.5}_{-0.4}$, but with significant variations within the population. Overall, we measure little to no evolution of the Lyα $\mathrm{EW_0}$ and scale length parameter (w_0) which are consistently high ($\mathrm{EW_0} = 140^{+280}_{-70}\,\text{Å}$, $w_0 = 129^{+11}_{-11}\,\text{Å}$) from $z \sim 6$ to $z \sim 2$, although w_0 is anti-correlated with rest-frame UV luminosity ($\mathrm{M_{UV}}$) and stellar mass. Our results imply that sources selected as LAEs have a high Lyα escape fraction ($\mathrm{f_{esc,Ly\alpha}}$) irrespective of cosmic time, but $\mathrm{f_{esc,Ly\alpha}}$ is still higher for UV-fainter and lower mass LAEs. We also find that the least massive LAEs are typically above the star formation "Main Sequence" and thus undergoing intense star formation, which could be explained by a bursty nature.

Furthermore, we measure the evolution from $z \sim 2$ to $z \sim 6$ of the rest-frame UV luminosity function (LF) and the stellar mass function (SMF) of the SC4K sample. We explore a range of 6 dex in $\mathrm{M_{UV}}$ and 5 dex in $\mathrm{M_\star}$, which is unprecedented for such a large sample of LAEs, covering such a redshift range. For both the LFs and SMFs, we find that the Lyα luminosity limit significantly affects the shape and Schechter parameters of the distributions. As such, to probe for evolution in an unbiased way, we study a subset of the SC4K sample, with $\mathrm{L_{Ly\alpha}} > 10^{43.0}\,\mathrm{erg\,s^{-1}}$, which is a luminosity regime probed at all redshift ranges. For the UV LF of these LAEs, we find a characteristic

number density (Φ^*) increase from $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3}) \sim -5.2$ at $z = 2.5$ to ~ -4.6 at $z \sim 3$, remaining constant up to $z \sim 5-6$ and a characteristic UV luminosity (M^*_UV) brightening from -21.1 at $z \sim 3$ to -22.0 at $z \sim 5-6$. We find no significant evolution of the SMF of these LAEs with redshift, with $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3})$ staying constant at ~ -5.5 from $z \sim 2.5$ to $z \sim 6$ and the characteristic stellar mass staying constant at $\log_{10}(\mathrm{M}^*_\star/\mathrm{M}_\odot) \sim 10.7$ for the same redshift range. We measure that the UV luminosity density (ρ_UV) changes from $10^{24.2}$ to $10^{25.0}$ erg s^{-1} Hz^{-1} Mpc^{-3} and the stellar mass density (ρ_M) remains constant at $\sim 10^{5.5}$ M$_\odot$ Mpc^{-3}, with both always being smaller than literature measurements from continuum-selected galaxies. Both ρ_UV and ρ_M of LAEs converge to the measurements of continuum-selected galaxies at $z > 6$, which suggests a key role of LAEs in the epoch of reionisation.

Overall, our results show that LAEs are a unique subset of the star-forming population, and that as we move to higher redshifts, LAEs become more and more representative of the full population of galaxies.

The SC4K sample is made fully public, together with derived physical properties, so the community can fully benefit from this work.

This thesis is dedicated to my family, for always being there for me.

Acknowledgements

This thesis would not have been possible without the unwavering support of so many.

I want to start by thanking my supervisor, David Sobral. Thank you for everything, I would have been completely lost without your guidance. You got me interested in Astrophysics research so many years ago during my BSc, and all these years would have been impossible if it was not for you.

To my friend and flatmate João Calhau, thank you for putting up with me for the past 3 years.

Thank you to everyone in the Astro group. Isobel Hook for brilliantly leading the group since its inception in 2016. All the Academic Staff: John Stott, Julie Wardlow, Brooke Simons, Matthew Pitkin and Andreas Koch. All my fellow post-graduate colleagues and office mates: Matthew Chan, Steven Williams, Jon Carrick, Nick Amos, Heather Wade, Tom Cornish, Izzy Garland, David O'Ryan, Jaime Dumayne, Emma Dodd, Jess Craig, Veronica Lopez and Danil Kuzin. The visitors and temporary office mates: Pablo Arrabal Haro and Alexandro Lumbreras-Calle. Thank you all for making Lancaster University a home to me.

Thank you to the Head of Department, Professor Roger Jones. Thank you to all the admin and support staff: Deborah Dunne, Grainne Wilkinson and Shonah Ion, as well as the IT staff Rob Lewsey and Rupert Prudom. Thank you all for ensuring I had everything I needed for a successful PhD.

Thank you to all my contemporary PhD colleagues from around the world (now postdocs): Jorryt Matthee, Ana Afonso, Andra Stroe and Ali Khostovan. For setting the example of excellence, and for showing me how to be a PhD student.

Declaration

This thesis is my own work and no portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification at this or any other institute of learning.

The research presented in this thesis has been published in the relevant scientific journals in paper format, specifically Chapter 2, 3, 4 which are published in Sobral, Santos et al. (2018, MNRAS, 476, 4), Chapters 5, 6 in Santos et al. (2020, MNRAS, 493, 1) and Chapter 7 in Santos et al. (2020, submitted to MNRAS).

The sample of galaxies that we build in Chapter 2 consists of a compilation of narrow band-selected galaxies from previous work from our group (Matthee et al., 2017b; Santos et al., 2016; Sobral et al., 2017) and a completely new medium band-selected sample. In order to obtain my medium band-selected sample of galaxies, I make use of the full mosaics of medium band and broad band data, reduced by the COSMOS team (Capak et al., 2007; Scoville et al., 2007), and made publicly available in the COSMOS archive (see Ilbert et al., 2009; Taniguchi et al., 2015a).

The data analysis conducted by me can be summarised as follows:

- Dividing the full mosaics of all the available medium bands and broad bands in smaller pointings using SWARP (Bertin et al., 2002)

- Registering all images to the same WCS frame so SExtractor (Bertin & Arnouts, 1996) can be run in dual extraction mode

- Manual production of masks, which cover low quality regions and regions affected by diffraction patterns

- Zero Point calibration of all images

- Running Source Extractor in dual extraction mode to produce catalogues of detections for each medium band, with magnitudes at the corresponding positions in all broad bands

- Computation of multiple properties for each source, such as Lyα line-flux, Lyα luminosity, Lyα equivalent width and Sigma (Σ).

- Defining the selection criteria for each medium band, to select 1) line-emitters 2) Lyα emitters

- Production of a catalogue of line-emitters and a catalogue of LAEs for each medium band

- Visual checks for all Lyα emitters

- Deriving completeness correction functions for each medium band

- Construction of Lyα luminosity functions (individual filters and redshift bins)

- Deriving best Schechter fits for the Lyα luminosity functions

- Computation of luminosity function integrals, to derive $\mathrm{SFRD_{Ly\alpha}}/\mathrm{SFRD_{UV}}$

"If I have seen further it is by standing on the shoulders of Giants."
Sir Isaac Newton

Contents

List of Figures

Relevant Publications by the Author

<u>Papers included in this thesis:</u>

- "Slicing COSMOS with SC4K: the evolution of typical Lyα emitters and the Lyα escape fraction from $z \sim 2$ to $z \sim 6$"; Sobral, D.; **Santos, S.**; Matthee, J.; Paulino-Afonso, A.; Ribeiro, B.; Calhau, J.; Khostovan, A.A.; **2018, MNRAS, 476, 4.**

- "The evolution of rest-frame UV properties, Lyα EWs and the SFR-Stellar mass relation at $z \sim 2 - 6$ for SC4K LAEs"; **Santos, S.**; Sobral, D.; Matthee, J.; Calhau, J.; da Cunha, E.; Ribeiro, B.; Paulino-Afonso, A.; Arrabal Haro, P.; Butterworth, J.; **2020, MNRAS, 493, 1.**

- "The UV luminosity function and galaxy stellar mass function of $z \sim 2 - 6$ LAEs with SC4K"; **Santos, S.**; Sobral, D.; Butterworth, J.; Paulino-Afonso, A.; Ribeiro, B.; da Cunha, E.; Calhau, J.; Khostovan, A. A.; Matthee, J; Arrabal Haro, P.; **2020, submitted to MNRAS.**

<u>Other relevant publications:</u>

- "Identification of the brightest Lyα emitters at z = 6.6: implications for the evolution of the luminosity function in the re-ionisation era"; Matthee, J; Sobral, D; **Santos, S**; Rottgering, H; Darvish, B; Mobasher, B; **2015, MNRAS, 451, 4919.**

- "Evidence for PopIII-like stellar populations in the most luminous Lyα emitters at the epoch of re-ionisation: spectroscopic confirmation"; Sobral, D; Matthee, J; Darvish, B; Daniel S; Mobasher, B; Rottgering, H; **Santos, S**; Hemmati, S; **2015, ApJ, 808, 139.**

- "The Lyα luminosity function at $z = 5.7 - 6.6$ and the steep drop of the faint end: implications for reionization"; **Santos, S**, Sobral, D., Matthee, J.; **2016, MNRAS, 463, 1678.**

- "A large Hα survey of star formation in relaxed and merging galaxy cluster environments at z $\sim$ 0.15-0.3"; Stroe, A., Sobral, D., Afonso, A., Alegre, L., Calhau, J., **Santos, S.**, Weeren, R.v.; **2017, MNRAS, 465, 3.**

- "Spectroscopic properties of luminous Lyman-α emitters at $z \sim 6 - 7$ and comparison to the Lyman break population"; Matthee, J; Sobral, D; Darvish, B; **Santos, S**; Mobasher, B; Paulino-Afonso, A; Rottgering, H; Alegre, L; **2017, MNRAS, 472, 772.**

- "On the UV compactness and morphologies of typical Lyman-α emitters from $z \sim 2$ to $z \sim 6$"; Paulino-Afonso, A., Sobral, D., Ribeiro, B., Matthee, J., **Santos, S.**, Calhau, J., Forshaw, A., Johnson, A., Merrick, J., Perez, S., Sheldon, O.; **2018, MNRAS, 476, 4.**

- "The clustering of typical Lyα emitters from $z \sim 2.5$ to $z \sim 6$: host halo masses depend on Lyα and UV luminosities"; Khostovan, A; Sobral, D; Mobasher, B; Matthee, J; Cochrane, R; Chartab Soltani, N; Jafariyazani, M; Paulino-Afonso, A; **Santos, S**; Calhau, J; **2019, MNRAS, 493, 3.**

- "The X-ray and radio activity of typical and luminous Lyα emitters from $z \sim 2$ to $z \sim 6$: evidence for a diverse, evolving population"; Calhau, J., Sobral, D., **Santos, S.**, Matthee, J., Paulino-Afonso, A., Stroe, A., Simmons, B., Barlow-Hall, C.; **2020, MNRAS, 493, 3.**

Chapter 1

Introduction

The evolution of the Universe across cosmic time is one of the most complex topics humanity has ever attempted to grasp. Over the past century, with great effort and dedication from the astrophysics community, a picture of how the Universe and its contents evolve with time has started to emerge. This chapter aims to briefly portray our current understanding of the Universe, and outline how this thesis contributes to that knowledge.

1.1 Theoretical framework: ΛCDM Universe

As we attempt to probe the physics behind galaxy formation and evolution, it is fundamental to start by outlining the cosmological framework, as it models the initial conditions and seeds necessary to form primordial galaxies, which evolve into the galaxies we will study in this thesis.

The ΛCDM cosmological model (Λ Cold Dark Matter, with Λ being the cosmological constant), regarded as the "standard model" of the Universe, provides predictions for the evolution of the Universe since its initial stage, the Big Bang. As the Universe expands, and gradually becomes less dense and cools down, it eventually reaches a stage where gravity can dominate at small scales and structure can form. ΛCDM is built on the cosmological principle, which states that at large scales the Universe is homogeneous and isotropic, and thus an observer will

not be favoured by their position or by the direction they observe. It provides good explanations for observables such as the anisotropy of the Cosmic Microwave Background (CMB, Penzias & Wilson, 1965; Planck Collaboration et al., 2018) and the accelerated expansion of the Universe (Perlmutter et al., 1999; Riess et al., 1998).

1.1.1 Cosmological parameters

According to the ΛCDM model, the energy density of the Universe can be divided into four components: matter (Ω_M), dark energy (Ω_Λ, the term that drives the expansion of the Universe), radiation (Ω_R) and curvature (Ω_K). These cosmological parameters have been constrained by measuring the power spectrum of the CMB (Planck Collaboration et al., 2018; Spergel et al., 2003), the distances of type Ia supernovae (Perlmutter et al., 1999; Riess et al., 1998) and through weak gravitational lensing (Abbott et al., 2018; Hildebrandt et al., 2017). Although there is some tension between measurements, they point towards a non-zero Ω_Λ (and thus a non-zero cosmological constant Λ) that drives an accelerated expansion of the Universe. The latest Planck Collaboration et al. (2018) results[1] are $\Omega_M = 0.3111$, $\Omega_\Lambda = 0.6889$, $\Omega_R \sim 0$ and $\Omega_K \sim 0$ (consistent with a spatially flat $\Omega_K = 0$ Universe), with the latter two being negligible at the present day. ΛCDM places the age of the Universe at 13.787 Gyr.

The matter component can be further divided into baryonic ($\Omega_b = 0.0490$) and (cold) dark matter ($\Omega_c = 0.2589$). This means that the matter we can observe directly (baryonic) only constitutes 1/6 of the total mass and $\sim 5\%$ of the total energy density of the Universe. Dark matter particles have never been directly observed, however, the existence of a non-negligible amount of matter that does not emit light but interacts gravitationally is necessary to explain multiple observational evidence. Some examples include: the velocity of galaxies in the Coma Cluster (Abell 1656) is too high for the cluster to remain bound together with the observed mass (Zwicky, 1933); movement of companions around the

---

[1]All values stated in this section are obtained from Planck Collaboration et al. (2018), Table 2, last column. Ω_b is computed from $\Omega_b h^2$, with $h = H_0/100 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$ and $H_0 = 67.66 \, \text{km} \, \text{s}^{-1} \, \text{Mpc}^{-1}$

Milky Way requires massive haloes (Einasto et al., 1974); rotation speeds of stars in the Milky Way is found to be significantly higher than the predicted estimates for the observable mass (Rubin et al., 1980); spatial offset between the centre of the total mass and the centre of the baryonic mass peaks on the Bullet Cluster (Clowe et al., 2006; Markevitch et al., 2002); simulations have shown that primordial galaxy formation in the early Universe requires dark matter haloes (Navarro et al., 1997; Tormen et al., 1997, see Section 1.1.2 for a more detailed explanation on primordial galaxy formation).

Throughout this thesis, for simplicity and direct comparison with the literature, we use a ΛCDM cosmology with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.3$ and $\Omega_\Lambda = 0.7$.

1.1.2 Formation of primordial galaxies

A Universe that follows the cosmological principle with no fluctuations would have stayed uniform and isotropic, and no structure would ever form. However, this is evidently not the case. Guth (1981) proposed a rapid period of exponential expansion right after the Big Bang, the inflation. This period allows density perturbations to originate from quantum fluctuations. These density perturbations, made of baryonic and dark matter, would become the seeds for primordial galaxy formation.

As the Universe expands and cools down, gravity can now dominate at small scales. Higher density regions will attract more matter and become denser, while voids will become emptier. White & Rees (1978) proposed a two-stage theory for galaxy formation and clustering. (Cold) dark matter, which is collisionless and will thus not heat up, will collapse and relax into so called dark matter haloes. Baryonic mass will then fall into the potential wells created by the dark matter haloes. As the gas is infalling, its gravity will dominate over the dark matter halo, and it will collapse in clumps, originating stars. Dark matter haloes can grow by merging with other dark matter haloes.

The formation of stars through gravitational collapse of gas is heavily influenced by gas cooling. As the gas compresses, it heats, with radiative cooling removing the excess energy. In this period of the Universe, only light elements

(hydrogen, helium) are present, which are very inefficient at cooling, with the main cooling mechanism being through H_2 molecules (Palla et al., 1983; Silk, 1977). This leads to the first generation of stars (Population III, Pop III) being very massive, with earlier works predicting masses up to $1000\,M_\odot$ (Nakamura & Umemura, 2001) or even higher (Bond et al., 1984; Carr et al., 1984). The shape of the Initial Mass Function (IMF)[1] of Pop III stars and its boundaries[2] is challenging to estimate, but significant progress has been made in constraining them through simulations (Nakamura & Umemura, 2001). Due to their high mass, Pop III stars are very bright and short-lived. After $\sim 100\,Myr$, 10-35 $M_\odot$ Pop III stars explode into type II supernovae (Nakamura & Umemura, 2001), releasing heavy elements into the IGM.

The first sources of light produce ioniosing photons which gradually turn a then neutral Intergalactic Medium (IGM) into an ionised one (e.g. Couchman & Rees, 1986; Fukugita & Kawasaki, 1994; Ostriker & Gnedin, 1996), with this period being aptly named the epoch of reionisation (EoR). The reionisation of the Universe was likely driven by two types of sources, quasars and star-forming galaxies. The relative contribution of each and whether their contributions are even enough to reionise the Universe remain big open questions which high redshift studies strive to answer (see e.g. Verhamme et al., 2017). Quasars do not seem to be able to reionise the Universe by themselves (e.g. Cowie et al., 2009; Fontanot et al., 2012), but a population of faint Active Galactic Nucleus (AGN) recently reported by Giallongo et al. (2015) could provide the missing contribution assuming a 100% escape of ionising photons. Similarly, star-forming galaxies could in principle reionise the Universe if enough ionising photons and thus Lyman Continuum radiation (LyC) escapes to the IGM. The escape fraction of LyC photons and the escape fraction of Lyα photons is correlated (Dijkstra, 2017), and Lyα has been shown to escape more easily from UV-faint galaxies (Dijkstra, 2017). This population of UV faint galaxies which have high escape of Lyα photons (and LyC photons) could have the necessary conditions to reionise the Universe, and the discovery of galaxies with these properties provides a better understanding of reionisation.

---

[1] empirical function that quantifies the initial mass distribution of a stellar population
[2] lower and upper mass limits at which stars are allowed to form for a specific IMF

Planck Collaboration et al. (2018) places the EoR at $z = 7.82 \pm 0.71$ (assuming an instantaneous reionisation) but studies of the Lyman forest of quasars have shown that it extends until $z \sim 6$ (Fan et al., 2006). In fact, multiple studies have proposed the EoR as being a patchy inhomogeneous period with preferred locations (e.g. Matthee et al., 2015; Santos et al., 2016), with galaxies emitting ionising radiation and carving ionised bubbles around them (e.g. Cen & Haiman, 2000; Hu et al., 2016; Mason & Gronke, 2020; Matthee & Sobral, 2020; Matthee et al., 2018).

1.2 Observational tools to explore the $z > 2$ Universe

With the basic principles of primordial galaxy formation in a ΛCDM Universe now laid out, we will cover the observational bases and tools necessary to directly (or indirectly) find galaxies in the distant ($z > 2$) and young ($< 2\,\mathrm{Gyr}$) Universe, and to measure their properties throughout cosmic time. During the past two decades, giant leaps have been taken in our understanding of early galaxy evolution, mostly due to the advent of a plethora of multi-wavelength data, with increasing depth and resolution, from state-of-the-art facilities such as the *Hubble Space Telescope* (HST), the *Very Large Telescope* (VLT) and the *Atacama Large Millimeter Array* (ALMA), among many others.

1.2.1 Overview of the spectral energy distribution of a galaxy

A galaxy consists mainly of stars, gas and dust gravitationally bound together in a dark matter halo. Each star emits approximately as a black body, with the intensity at some wavelength being a function of its temperature. Emitted radiation will interact with the gas and recombine in the form of emission lines. Dust absorbs emission as a function of wavelength and re-emits in the mid- to far-infrared

(MIR, FIR). The total emission of a galaxy will be all these individual (but interlinked) components summed together. Disentangling all this information to obtain intrinsic properties of a galaxy such as its stellar mass can be a daunting task. We show in Figure 1.1 a fiducial spectral energy distribution (SED) template of observed emission from a $z = 3$ galaxy at different wavelengths, from X-ray to radio, and the causes of emission.

The stellar component (emission from stars) directly contributes to the SED from rest-frame ultraviolet (UV) to infrared (IR). Blueward of rest-frame $912 \, \text{Å}$ (Lyman limit, with $912 \, \text{Å}$ being the photoionisation potential of the atomic hydrogen) there is a distinct characteristic break (Lyman break), where the flux abruptly drops to zero. This is a consequence of radiation more energetic than the Lyman limit being easily absorbed by neutral gas in star-forming regions. The detection and position of this break has been extensively used in high redshift searches (see Section 1.2.2.1). Nebular emission from gas that gets ionised from e.g. photons from young stars will emit from optical to FIR, most significantly in the form of emission lines, which have also been extensively used to select and follow-up high redshift galaxies (see Section 1.2.2.2). Thermal emission from dust happens when emission from e.g. stars gets absorbed by dust and re-emitted approximately as a blackbody in the FIR.

Additionally, a galaxy can have very significant contribution from an AGN, i.e. from the supermassive black hole in the centre of a galaxy in the process of accreting matter, which when active can boost emission across the full wavelength spectrum. AGN can have X-ray emission from inverse Compton effect and Synchrotron radiation, UV-NIR emission from the disk and torus, and radio Synchrotron emission from radio AGN. It is important to establish whether a galaxy has an active AGN boosting its photometry before deriving physical properties from its spectral energy distribution. While the central supermassive black hole is very small in size compared to the scale of a galaxy, it can dominate the emission at certain wavelengths. Additionally, the supermassive black hole even seems to regulate galaxy processes, with the peak of super massive black holes activity coinciding with the peak of star-formation history (e.g. Aird et al., 2010; Calhau et al., 2017; Delvecchio et al., 2014; Shankar et al., 2009).
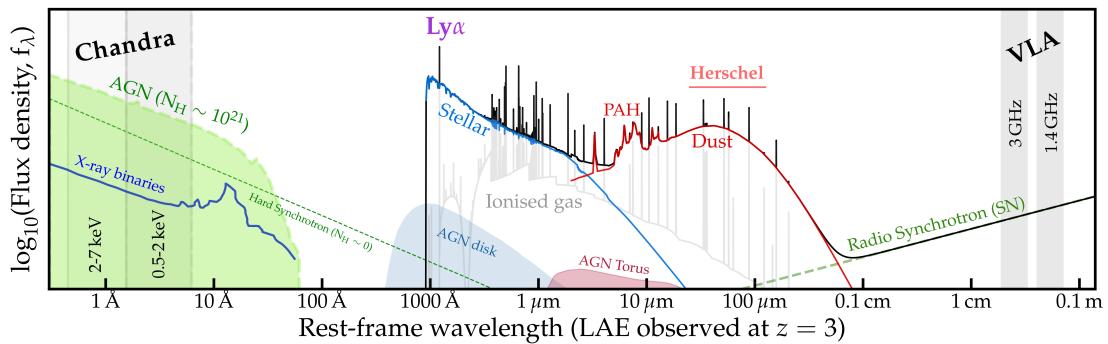
Figure 1.1: Overview of observed emission from a $z = 3$ star-forming galaxy with an AGN, identifying the key mechanisms that originate emission at different wavelengths, from X-ray to radio. We show the contribution from the stellar component (light blue line), nebular emission (including Lyα, black line) and thermal emission from dust and PAHs[a] (red line). Shaded regions show the contribution from the AGN disk (blue) and torus (red), and X-ray emission from inverse Compton effect and Synchrotron radiation (green). The dark blue line is the contribution from X-ray binaries. X-ray hard Synchrotron with no absorption (green dotted line, left) and the radio Synchrotron emission from radio AGN and supernovae (green dashed line). The telescopes Chandra, Herschel and VLA can follow up the template $z = 3$ galaxy in the X-ray, FIR and radio, respectively. The grey shaded regions mark the Chandra and VLA filters for such a multi-wavelenght follow-up in Calhau et al. (2020), credited for this figure.

---

[a]PAHs (polycyclic aromatic hydrocarbons) are associated with strong emission lines which can dominate the MIR spectrum of star-forming galaxies (e.g. Cortzen et al., 2019).

1.2.2 Observational strategies to hunt for distant galaxies

Multiple strategies have been deployed to search for young galaxies in the distant Universe. Here, we focus on two main approaches: rest-frame UV continuum searches and emission line searches.

1.2.2.1 Rest-frame UV continuum searches

UV searches have successfully used the Lyman break selection technique (Steidel et al., 1996a) to select tens of thousands of galaxies at $z \sim 2 - 10$ (e.g. Bouwens et al., 2015; Bunker et al., 2004; Finkelstein et al., 2015; Ono et al., 2018; Reddy & Steidel, 2009; Steidel et al., 1999), by searching for a characteristic feature of galaxy emission: the redshifted Lyman break (see Section 1.2.1). In Figure 1.2 we provide an example of the photometric profiles of Lyman Break galaxies (LBGs), showing the sharp break in photometry from stacked broad band imaging in 8 bands of 754 LBGs in the range $z = 5 - 6$ (McLure et al., 2009).

Typical strategies to detect LBGs focus on obtaining extremely deep observations of small areas of the sky (e.g. Bouwens et al., 2010, 2015; Finkelstein et al., 2015), targeting well-known patches of the sky such as the Hubble Ultra Deep field. Sources selected this way are typically too faint for spectroscopic follow-up and for multi-wavelength follow-up with e.g. ALMA. Multiple studies have also resorted to targeting clusters to search for gravitationally lensed high redshift galaxies (e.g. Atek et al., 2015; Bartelmann, 2010; Zitrin et al., 2015), which reveals even fainter populations that require amplification to be detected. Sources selected this way are typically also difficult to follow up and require lensing modelling to reconstruct the amplified image. Some other studies have targeted wider areas (e.g. Bowler et al., 2017; Ono et al., 2018), which has led to the discovery of brighter LBGs, pushing the observing limits of $z = 4$ LBGs from $M_{UV} \sim -23$ to as bright as $M_{UV} \sim -26$. While a significant fraction of these extremely bright M_{UV} sources are AGN (Ono et al., 2018), they reveal an important population which statistically can only be detected by probing large volumes (~ 1 source per $10^8 \, \mathrm{Mpc}^3$), which can be $\sim 2 - 3$ orders of magnitude larger than volumes probed by typical surveys.
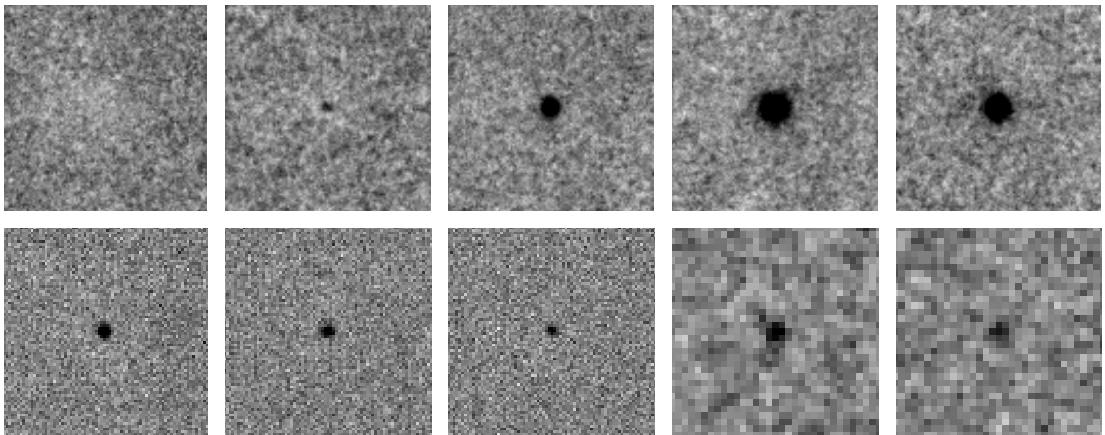
Figure 1.2: Stacked photometry in 8 broad bands of 754 galaxies at $z = 5 - 6$. A sharp break in colour, associated with the Lyman break, is visible in the top left panels. Top row: $BVRi$'. Bottom row: $z'JHK$. Credit McLure et al. (2009).

1.2.2.2 Rest-frame optical emission line surveys

Alternatively, some studies have selected large samples of galaxies by searching for their most prominent features: emission lines. Up to $z \sim 2.2$, Hα ($\lambda_{\rm rest} = 6563$ Å) has been commonly used to select and study star-forming galaxies (SFGs). Hα is a non-resonant line, requiring only a dust correction when converting luminosity to star-formation rate (see Section 1.2.3, and see also Sobral et al. (2013) for SFR measurements of Hα emitters at $z = 0.40 - 2.23$). At higher redshifts, which we aim to probe in this thesis, Hα gets redshifted into the MIR and can no longer be observed from ground-based telescopes, with alternatives being necessary until new generations of space telescopes are capable of conducting Hα surveys. Lyα ($\lambda_{\rm rest} = 1216$ Å), being intrinsically the brightest emission line in the UV-optical range due to the abundance and nature of atomic hydrogen, is an option to explore the $z > 2$ Universe. It has been very successful in selecting high redshift galaxies (see Section 1.4), albeit its complex radiative transfer and resonant nature make it challenging to interpret (for a review see Dijkstra, 2017). Constructing large samples of galaxies with Lyα emission and studying their evolution across cosmic time is the focus of this thesis.

1.2.3 Quantifying galaxy activity: star-formation rates

As we seek to measure galaxy evolution, we introduce a key parameter which quantifies how active a galaxy is: the star-formation rate (SFR), i.e. the mass of stars formed per unit of time. While stellar formation occurs with the collapse of clouds of gas in star-forming regions, resolving such regions individually within a $z > 2$ galaxy is not possible with current instrumentation and techniques (although there is significant progress in simulations to reach $\sim 10\,\mathrm{pc}$ resolutions, e.g. Agertz & Kravtsov, 2015; Muratov et al., 2015; Shimizu et al., 2019). We can thus only estimate the SFR of the entire galaxy system, and measurements should be interpreted as the total SFR of the whole galaxy. As the only observable is emission at different wavelengths, inferring the rates of stellar formation in galaxies is thus a problem of measuring mass from light (for a review see Madau & Dickinson, 2014).

The process of forming stars creates signatures, which we can use to trace and measure the SFR. These tracers of star-formation are mostly associated with the presence (or death in the form of supernovae) of very massive O and B stars, which are short-lived and emit vast amounts of ionising photons. These include direct tracers such as the UV light, which is directly emitted by these very massive stars, or more indirect tracers such as nebular emission in the form of e.g. Hα and Lyα (with the later being harder to translate directly into SFR due to the complex radiative transfer associated with this line, see Section 1.3.3) and also infrared emission due to dust absorbing UV radiation and re-emitting in the MIR-FIR. All these calibrations are based on estimating the Lyman continuum emitted by the most massive stars and inferring the contribution of the full stellar population under some assumptions of the initial mass function (within a mass boundary) and a prescription of stellar evolution.

- **UV as a SFR tracer**

  UV light traces recent star-formation in the past 100 Myr (e.g. Boselli et al., 2001; Salim et al., 2009), as very hot and massive stars that emit copious UV radiation are short-lived, with their detection being evidence of a recent episode of star-formation. Very massive stars dominate emission in UV, around $1500\,\text{Å}$, and we can use this wavelength range to estimate the SFR.

Assuming a Salpeter (1955) IMF with mass range 0.1-100 $M_\odot$[1], and the prescription from Madau et al. (1998) we can estimate SFR directly from the UV luminosity as (Kennicutt, 1998):

$$\text{SFR}_{\text{UV}}\,[\text{M}_\odot\,\text{yr}^{-1}] = 1.4 \times 10^{-28} \text{L}_\nu, \tag{1.1}$$

with L_ν being the UV luminosity in $\text{erg}\,\text{s}^{-1}\,\text{Hz}^{-1}$, computed at rest-frame $1500\,\text{Å}$. It should be noted that UV radiation is very susceptible to dust extinction, and these SFRs require a dust correction.

- **FIR as a SFR tracer**

  Dust preferentially absorbs UV photons due to the high cross section of its grains, and then radiates as a blackbody, with the intensity of the FIR emission scaling with the intensity of radiation that is absorbed. FIR can thus be used as an indirect measurement of the UV light (if we can estimate how much is absorbed) and it is thus an indirect tracer of star-formation. SFR can be estimated from the FIR luminosity (L_{FIR}) by applying the models of Leitherer & Heckman (1995) for continuous bursts of age $10 - 100\,\text{Myr}$, and adopting a Salpeter (1955) IMF, yielding the relation (Kennicutt, 1998):

  $$\text{SFR}_{\text{FIR}}\,[\text{M}_\odot\,\text{yr}^{-1}] = 4.5 \times 10^{-44} \text{L}_{\text{FIR}}, \tag{1.2}$$

  with L_{FIR} in $\text{erg}\,\text{s}^{-1}$. The FIR provides a window into obscured SFR.

- **Nebular lines as SFR tracers**

  Very massive O and B stars ionise HII regions and originate nebular emission, in the form of lines such as Hα and Lyα. These nebular lines trace star-formation activity of very massive stars and can be used to estimate the SFR. For solar abundances and the same Salpeter (1955) IMF ($0.1-100\,\text{M}_\odot$) that was used in deriving Equation 1.1, the calibrations of Kennicutt et al. (1994) and Madau et al. (1998) yield (Kennicutt, 1998):

---

[1]An assumption of the IMF and its mass boundaries is necessary to estimate the total SFR, so we can extrapolate the UV emission (traces mostly the contribution from the very massive stars) to lower masses.

$$\text{SFR}_{\text{H}\alpha} \, [\text{M}_\odot \, \text{yr}^{-1}] = 7.9 \times 10^{-42} \text{L}_{\text{H}\alpha}, \tag{1.3}$$

with $\text{L}_{\text{H}\alpha}$ in erg s^{-1}.

Due to Lyα's resonant nature (see Section 1.3.3), converting Lyα luminosity to SFR is much more challenging, as the percentage of Lyα photons that can escape a galaxy can be very uncertain. However, it is possible to infer Hα luminosity from Lyα luminosity if we can predict the escape fraction (see Section 1.3 for full details) and assuming case-B recombination (Brocklehurst, 1971). For a Salpeter (Chabrier) IMF ($0.1 - 100 \, \text{M}_\odot$) and assuming $\text{f}_{\text{esc,LyC}} = 0$, the SFR becomes (Sobral & Matthee, 2019, we provide a more detailed explanation of the derivation of this equation in Section 1.3.3):

$$\text{SFR}_{\text{Ly}\alpha} \, [\text{M}_\odot \, \text{yr}^{-1}] = \frac{\text{L}_{\text{Ly}\alpha} \times 7.9(4.4) \times 10^{-42}}{0.042 \, \text{EW}_0}, \tag{1.4}$$

with $\text{L}_{\text{Ly}\alpha}$ in erg s^{-1} and EW_0 in Å.

1.2.4 Cosmic star-formation history

The measurement of the cosmic star-formation density (ρ_{SFR}), i.e. mass of stars formed per unit of time per unit of volume at each epoch of the Universe, offers a global overview of the evolution of the Universe. Multiple studies with independent approaches have measured ρ_{SFR} over the past $\sim 12 \, \text{Gyrs}$ and we now have a well-defined picture of how the star-formation history of the Universe evolves with time. We show in Figure 1.3 ρ_{SFR} measurements from four distinct methods, using continuum UV and FIR (see Madau & Dickinson (2014) for a review) or emission lines Hα (Sobral et al., 2013), [OII] and Hβ + [OIII] (Khostovan et al., 2015). The peak of ρ_{SFR} is found to occur at $\sim 2 - 3$ (e.g. Karim et al., 2011; Lilly et al., 1996; Madau et al., 1996), indicating galaxies were on average forming more stars during this period (cosmic noon), with a subsequent decline for lower (e.g. Sobral et al., 2013) and higher redshifts (e.g. Bunker et al., 2010; Khostovan et al., 2015; Stanway et al., 2003). This results in about half of all stellar mass already being formed by $z = 1$ (e.g. Bundy et al., 2005; Mortlock et al., 2011).
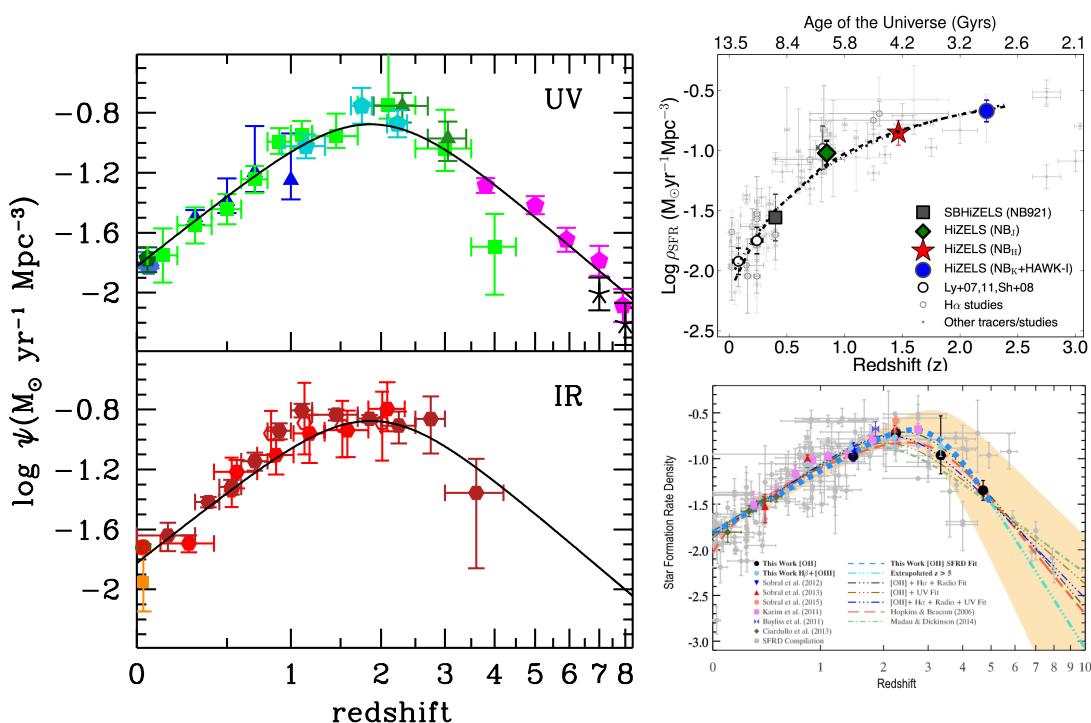
Figure 1.3: Cosmic star-formation history estimated through different approaches: UV (top left, credit Madau & Dickinson (2014)), IR (bottom left, credit Madau & Dickinson (2014)), Hα (top right, credit Sobral et al. (2013)), [OII] and Hβ + [OIII] (bottom right, credit Khostovan et al. (2015)). The evolution of $\rho_{\rm SFR}$ with redshift gives a global overview of galaxy evolution in the Universe, painting a picture where galaxies were on average forming more stars at $z \sim 2 - 3$ and are on average less active at both lower and higher redshifts.

1.3 Lyman-α: a powerful tool to probe the high-redshift Universe

Hydrogen is the most abundant element of the Universe, currently composing $\sim 75\%$ of the total baryonic mass (Planck Collaboration et al., 2018). Radiative transfers of this element are thus present everywhere, and become a powerful tool to study galaxy formation and evolution. The presence of hydrogen lines has indeed been shown to be ubiquitous in the early Universe:

> ''Nearly 100% of the sky is covered by Lyman-α emission around high redshift galaxies'' - Wisotzki et al. (2018)

We show in Figure 1.4 the detection of the Lyman-α (Lyα; $\lambda_0 = 1215.67\,\text{Å}$) transition of the hydrogen atom at $z = 3 - 6$, using data from the MUSE instrument. Lyα extended emission covers most of the sky when probing faint enough surface brightnesses, highlighting its significance when studying the early Universe.

Historically, Lyα was first predicted to be associated with young star-forming galaxies by Partridge & Peebles (1967):

> ''It seems possible that the Lyman-α line might be detected if it is a strong feature of the spectra of young galaxies.'' - Partridge & Peebles (1967)

This rests on the assumption that very young, massive and short-lived O and B stars emit large amounts of ionising photons which first ionise the gas, then the electrons recombine into atoms in the vast clouds of hydrogen that surround star-forming regions. The resulting recombination lines, including Lyα, would be indicative of the existence of on-going star-formation activity and could thus be used to find young star-forming galaxies. Ever since this prediction, Lyα (either in emission or absorption) has become a very powerful and widely used tool to explore the distant Universe. In addition to being used to select young star-forming galaxies at $z > 2$ (e.g. Cowie & Hu, 1998; Malhotra & Rhoads, 2004; Rhoads et al., 2000, 2003), it has been used to map the neutral state of the
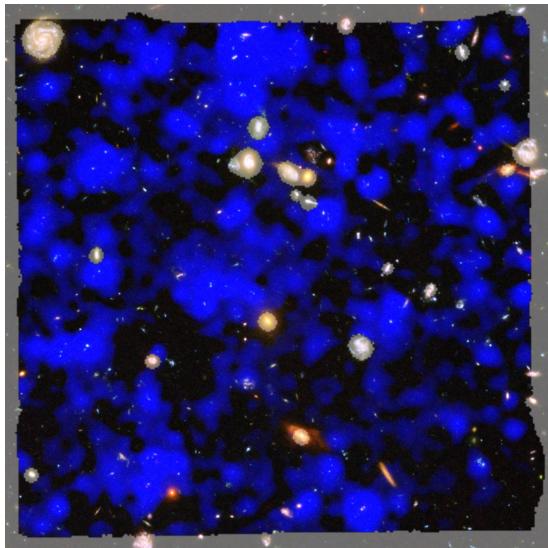
Figure 1.4: Distribution of the observed Lyα emission in the Hubble Ultra Deep Field. The extended Lyα emission (shown in blue), obtained by the Integral Field Unit MUSE, is the sum of the Lyα emission in the range $3 < z < 6$. This is superimposed over a colour composite image obtained by HST. Credit Wisotzki et al. (2018).

Universe across cosmic time through absorption features known as the Lyα forest (Fan et al., 2006). Lyα is intrinsically the brightest line in the UV-optical range, making it a staple in high redshift studies of galaxy formation and evolution, for both the selection (e.g. Cowie & Hu, 1998; Ouchi et al., 2008) and spectroscopic confirmation (e.g. Caruana et al., 2014; Ono et al., 2012; Pentericci et al., 2011; Stanway et al., 2004) of galaxies, with particular significance at $z > 2$ where it can be observed from the ground.

However, Lyα radiative transfer is a very complex process and translating observed measurements into intrinsic properties is a non-trivial task. In this section, we will describe the main physics behind Lyα emission and radiative transfer in galaxies.

1.3.1 The hydrogen atom and its transition lines

A hydrogen atom consists of one electron orbiting around one proton. The electron is described by two quantum numbers, n ($n = 1, 2, ...$) and l ($l = 0, 1, .., n-1$
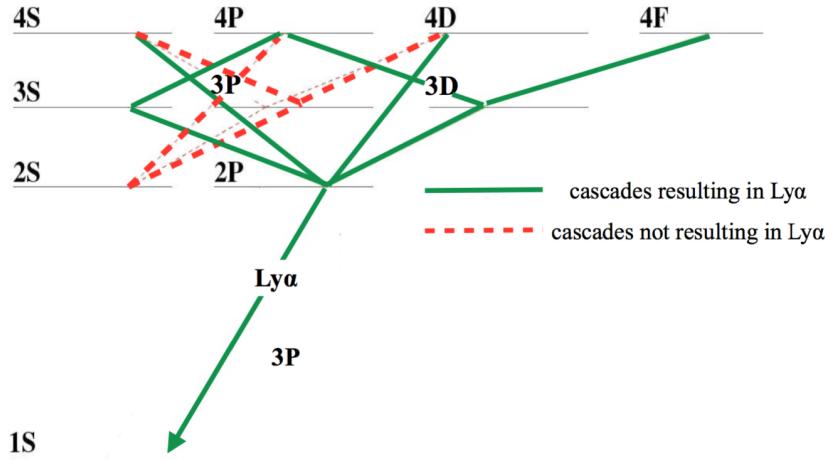
Figure 1.5: Transitions of the hydrogen atom. Cascades that result in the emission of a Lyα photon ($2p - 1s$) are shown in green and the ones that do not are shown in red. Credit Dijkstra (2017).

or equivalent $l = s, p, d, f..$). At its ground state $1s$ ($n = 1$, $l = 0$) the electron has an energy of $-13.6\,$eV, which is the energy needed to strip the electron from the atom. When hydrogen absorbs energy, the electron will move to a higher energy state and then proceed to cascade down to the ground state, obeying Δl=1 and emitting a photon at a well defined wavelength for every transition. We show in Figure 1.5 a diagram of transitions in the hydrogen atom.

When electrons cascade down from higher energy levels to the ground state, there is a probability that it will end with a $2p - 1s$ transition (Lyα). There are two main mechanisms that bring electrons to these higher excited states: collisions and recombination. Collisions occur when a hydrogen atom interacts with a free electron, with the free electron transferring some of its kinetic energy to the hydrogen atom, which will reach a higher energy state. This process will cool down the gas and is referred to as Lyα production via cooling. The recombination of a free proton and a free electron will produce a hydrogen atom with the electron at some quantum state (n, l). The electron will cascade down to the ground state with a probability of ending with a Lyα transition. The higher the temperature the lower the probability of producing Lyα.

1.3.2 Sources of Lyα emission

The most significant sources of Lyα emission are Interstellar HII regions (regions of atomic hydrogen that is ionised). These regions are powered by young, hot, massive stars that produce ionising photons that ionise the interstellar medium (ISM). Recombining protons and electrons will produce hydrogen in excited states, leading to cascades which include Lyα transitions. We can estimate the probability of a Lyα line being created, under a few assumptions.

The case-B approximation (Brocklehurst, 1971) assumes the gas is optically thick to all Lyman series photons (in contrast with case-A where it is assumed the gas is transparent to all Lyman series), and it is a viable approximation in Astrophysics as gases efficiently re-absorb very energetic photons. This means that any Lyman photon produced will get absorbed by a nearby neutral hydrogen atom and immediately re-emitted, in a process known as resonant scattering. With each scattering, there is a probability that the cascade will not end with a Lyman transition but will instead produce other higher order radiative transfers.

Assuming case-B recombination, for a hydrogen cloud with temperature T, the probability of a recombination event resulting in a Lyα transition is (Cantalupo et al., 2008):

$$\mathrm{P(Ly}\alpha) = 0.686 - 0.106 \log_{10}\left(\frac{\mathrm{T}}{10^4}\right) - 0.009\left(\frac{\mathrm{T}}{10^4}\right)^{-0.44} \qquad (1.5)$$

For a temperature of T=10^4K, the probability is P(Lyα)=0.68, while for a lower temperature of T=10^3K, it has a higher probably P(Lyα)=0.77. Other recombination lines will be produced during the cascade, such as Hα which has a probability P(Hα)=0.45 for T=10^4 K (Dijkstra, 2017). This non-resonant line can be used to calibrate Lyα as the total ratio of fluxes Lyα/Hα is 8.7 (Brocklehurst, 1971).

Lyα emission can also originate from the circumgalactic and intergalactic medium (CGM/IGM). Here, Lyα emission is triggered by sources external to these regions, usually star-forming galaxies or AGN. Emission occurs when radiation from these sources is absorbed by the CGM/IGM which, through recombination, goes through various radiative transfers which include the Lyα transition,

in a process known as fluorescence. This emission is spatially extended and can be observed as Lyα haloes around these star-forming galaxies or AGN. It is important to note that Lyα photons present in the CGM/IGM can be a combination of the two aforementioned sources of Lyα emission.

1.3.3 Complex radiative transfer physics

The journey of Lyα photons from their emission until they reach the observer is a difficult one, often shut down through resonant scattering and absorption by dust. When looking at Lyα emission originated by on-going star-formation in a star-forming galaxy, we need to consider that these photons have to escape the (dusty) ISM into the CGM/IGM and then have to travel through the CGM/IGM, which can be particularly difficult at higher redshifts where reionisation is still on-going and the medium around galaxies is not completely ionised.

Dust plays a very significant role in whether Lyα photons can escape from the ISM. Dust grains have sizes of the same order of magnitude as the wavelength of UV photons. This means that the energetic photons will get easily scattered and absorbed by dust. Thus, when a Lyα photon hits a grain of dust, it will get absorbed and radiation will be re-emitted in the FIR. The resonant nature of Lyα results in much longer path lengths, increasing the probability of Lyα photons encountering and being absorbed by grains of dust. Other redder, less energetic lines, such as Hα, will be less affected by dust, and we can use these lines to estimate intrinsic properties of lines that get destroyed by dust, after correcting for reddening.

Studies have tackled the complex problem of determining the amount of Lyα photons that escape by estimating the intrinsic Lyα luminosity and comparing it with the observed Lyα luminosity. This is quantified as the Lyα escape fraction ($f_{esc,Ly\alpha}$), which by definition is the ratio:

$$f_{esc,Ly\alpha} \equiv \frac{L_{Ly\alpha,obs}}{L_{Ly\alpha,int}}, \tag{1.6}$$

where $L_{Ly\alpha,obs}$ is the observed Lyα luminosity and $L_{Ly\alpha,int}$ is the intrinsic Lyα luminosity.

As previously mentioned, making some assumptions (case-B recombination, T=10^4K), the Hα and Lyα luminosity follow the ratios:

$$L_{H\alpha} = \frac{L_{Ly\alpha,int}}{8.7} = \frac{L_{Ly\alpha,obs}}{8.7 \times f_{esc,Ly\alpha}}, \tag{1.7}$$

From the Hα luminosity we are able to calculate the SFR (see Section 1.2.3). However, measuring $f_{esc,Ly\alpha}$ is a difficult task. Hayes et al. (2011) have estimated $f_{esc,Ly\alpha} \times \xi_{ion}$[1], where ξ_{ion} is the ionising efficiency (number of produced ionising photons per unit UV luminosity; e.g. Matthee et al., 2017a) in samples of galaxies selected by their Lyα emission (compiling different searches at $z \sim 0 - 8$) and found a significant increase of $f_{esc,Ly\alpha} \times \xi_{ion}$ with redshift, reporting $\sim$ 1 - 2% at $z \sim 0.3$, $\sim 5\%$ at $z \sim 2$ (both using direct measurements of Hα) and $\sim$ 30 - 50% at $z \sim 6$ (using dust corrected UV-continuum as Hα is no longer available at these redshifts from the ground), with an abrupt decline at $z > 6$ which can be attributed to reionisation. The most simple explanation for the increase in the escape fraction with redshift is a decrease in the dust content of galaxies, which would allow more Lyα photons to escape. Using a different approach, Wardlow et al. (2014) estimated the $f_{esc,Ly\alpha}$ of stacked $z = 2.8, 3.1$ and 4.5 Lyα emitters as the ratio between the SFR derived from Lyα luminosity (assuming case-B recombination) and the intrinsic (obscured+unobscured) SFR derived from SED templates. They find $f_{esc,Ly\alpha} > 10\%$ at all redshifts. These lower limits are consistent with the global evolution of $f_{esc,Ly\alpha}$ from Hayes et al. (2011).

Sobral & Matthee (2019) propose a simple empirical way to estimate $f_{esc,Ly\alpha}$ of a galaxy, requiring only the measurement of a much more easy to determine property, the Lyα equivalent width (EW). The EW is a measurement of strength of an emission (or absorption) line and by definition the width of continuum required to integrate in order to obtain the same area as the line (see Equation 2.6). The two properties follow the linear correlation:

---

[1] Computed from the ratio between the Lyα star-formation rate density and the UV star-formation rate density. An evolution of this ratio has historically been interpreted as an evolution of $f_{esc,Ly\alpha}$. More recently, an evolving ξ_{ion} was predicted and defined as an important factor that would contribute to the ratio (e.g. Matthee et al., 2017a). So the $f_{esc,Ly\alpha}$ measurement from Hayes et al. (2011) should be interpreted as $f_{esc,Ly\alpha} \times \xi_{ion}$.

$$f_{\text{esc,Ly}\alpha} = 0.0048 \times \text{EW}_0, \tag{1.8}$$

where EW_0 (Å) is the rest-frame EW. Combining this correlation with Equations 1.3 and 1.7, we can obtain a SFR estimate using only the observed Lyα luminosity and EW_0 (Equation 1.4).

1.4 Searches for Lyα emitters

After describing the mechanics behind Lyα emission and transmission, we now describe the strategies that have been commonly used to select galaxies with Lyα emission: Lyα emitters (LAEs). Lyα emission can be observed from ground based observatories at $z > 2$, so most Lyα studies are focused on higher redshifts. At lower redshifts, Hα is more commonly used as its non-resonant nature makes it easier to interpret. Nonetheless, some studies have searched for Lyα emitting galaxies at low redshift, such as the Lyα imaging+spectroscopy survey LARS (The Lyman-alpha reference sample Östlin et al., 2014) which used deep HST images to provide a reference sample of 14 Lyα emitting galaxies. This reference sample can be used to establish direct comparisons with its high redshift counterparts.

The search for high-redshift LAEs started with decades of unsuccessful attempts, with earlier works only being able to produce upper limits (e.g. Lowenthal et al., 1990; Pritchet & Hartwick, 1987; Thompson et al., 1995). The first LAEs were discovered in the early 1990s, with the discovery of a $z = 2.3$ LAE near a damped Lyα absorber (DLA) (e.g. Lowenthal et al., 1991), and three $z = 2.81$ candidate LAEs, also near a DLA (Møller & Warren, 1993). Additionally, the analysis of LBG samples, revealed that only $\sim 50\%$ of LBGs have Lyα emission, with a significant portion even showing absorption (Shapley et al., 2003).

On the search for LAEs in the early/distant Universe, three main approaches have been typically used: blind slit spectroscopy (e.g. Bayliss et al., 2010; Cassata et al., 2011; Martin & Sawicki, 2004; Rauch et al., 2008; Sawicki et al., 2008; Stark et al., 2007), Integral Field Unit observations (IFU, e.g. Adams et al., 2011; Bacon et al., 2015; Blanc et al., 2011; Drake et al., 2017a,b; Karman et al., 2015; van

Breukelen et al., 2005) and narrow (or medium) band surveys (e.g. Arrabal Haro et al., 2018; Cowie & Hu, 1998; Hu et al., 2010; Iye et al., 2006; Kashikawa et al., 2011; Konno et al., 2014; Malhotra & Rhoads, 2004; Matthee et al., 2014, 2015; Murayama et al., 2007; Nilsson et al., 2007; Ouchi et al., 2008, 2010; Rhoads et al., 2000, 2003; Santos et al., 2016; Shibuya et al., 2012, 2018; Shimasaku et al., 2006; Sobral et al., 2009; Taniguchi et al., 2005; Westra et al., 2006).

- **Narrow band surveys:** historically the first approach applied specifically to hunt for LAEs over large areas. In the late 1990s, Cowie & Hu (1998) successfully selected a sample of LAEs by using a custom narrow band filter centred at $\lambda = 5390$Å, sensitive to Lyα emission at $z = 3.4$. This approach requires imaging observations with one narrow band and at least one broad band at the wavelength of the narrow band to estimate the continuum of a galaxy. Sources with significant excess in the narrow band (see Bunker et al., 1995, and Section 2.3 for a more detailed explanation on the criteria for sources to be classified as line-emitters) likely have an emission line boosting the photometry at those wavelengths. Further colour-colour criteria, based on the presence of the Lyman break is applied to identify Lyα emitters from the full sample of line-emitters. Narrow band surveys are very efficient at quickly and cheaply covering wide areas/volumes, to select samples of thousands of candidate LAEs with little observing time. The wide volumes covered allow the selection of the brightest (but more rare) sources, which are prime targets for spectroscopic follow-up (e.g. Hu et al., 2016; Matthee et al., 2017c; Sobral et al., 2015). Additionally, narrow band surveys could in principle be conducted with medium bands, efficiently probing wider redshift slices at the cost of only being sensitive to higher EWs. Being based on purely photometric surveys, this approach often has an associated non-negligible interloper fraction (see Section 2.3.5).

- **Blind slit spectroscopy:** blindly placing long slits in blank areas can be an effective way to identify and obtain spectra of high-redshift galaxies, including LAEs, in an unbiased (blind) way. However, placing narrow slits in blind positions can lead to a significant loss of flux, particularly when dealing with Lyα emission which can be very extended (see Section 1.5.3).

Furthermore, the small areas/volumes and hard to estimate completeness corrections can lead to bias in measurements, particularly when attempting to probe brighter luminosities.

- **Integral field unit observations:** with the advent of new generations of integral field units such as MUSE/VLT, it became possible to obtain extremely deep data cubes in modest sized areas ($\sim 1\,\mathrm{arcmin}^2$), with the added benefit of having not only a photometric survey, but also a spectrum per pixel. This allows the immediate spectroscopic analysis of LAEs, including the study of spectral line profiles of Lyα and other emission lines. The deep observations provide an excellent probe of very faint luminosities, much fainter than the ones reached by e.g. narrow band surveys, allowing the constraint of the faint end slope of the luminosity function. However, the small volumes probed (when compared to narrow band surveys) are not enough to analyse the brightest luminosity regimes, and also are more subject to cosmic variance.

The different approaches (among others, e.g. slitless spectroscopy with *HST*) are sensitive to different luminosity ranges. Understanding their scope and limitations, and combining measurements from multiple surveys can provide a much more clear picture of the full population of LAEs.

1.5 Properties of Lyα emitters

Here, we give a brief overview of the characteristic properties of galaxies selected by the strength of their Lyα emission.

1.5.1 Nature of LAEs: star-forming galaxies vs AGN

We have discussed how Lyα emission is expected to be a strong feature of young star-forming galaxies (Partridge & Peebles, 1967) but also in AGN. The nature of LAEs can thus be very distinct, but observations can reveal the origin of Lyα emission. Very broad emission lines in the spectra of galaxies is typically indicative of AGN activity (e.g. Sobral et al., 2018b), with the broadening being caused
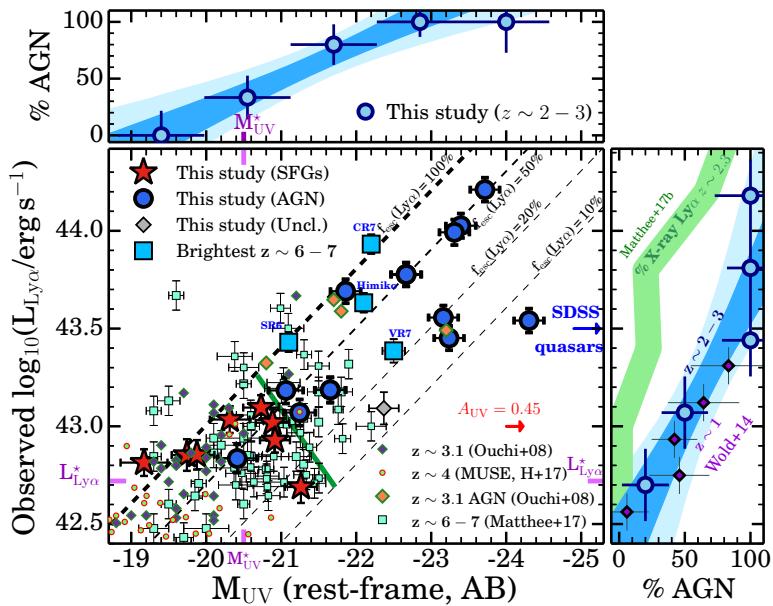
Figure 1.6: UV luminosity dependence on Lyα luminosity for a sample of $z \sim 2-3$ LAEs. There is a clear division between SFGs and AGNs, with the latter populating the most luminous regimes, both in UV and Lyα. Credit Sobral et al. (2018b).

by a rotating accretion disk. Furthermore, line ratios can also shed some light on the presence of an AGN, as AGNs typically have stronger metal lines. Aside from the differences revealed by the spectra, there is also significant photometric differences between the SFG and AGN Lyα emitting populations. In Figure 1.6, we show the UV luminosity dependence on Lyα luminosity for a sample of $z \sim 2-3$ LAEs (Sobral et al., 2018b). There is a clear distinction between SFGs and AGNs, with the latter dominating the more luminous regime, both in UV and Lyα. The transition from star-forming dominated to AGN dominated happens at $2 \times L^*$ (characteristic luminosity) and can be linked with a physical limit for the maximum observed output of an unobscured starburst (Sobral et al., 2018b).

Additionally, emission in the X-ray and radio is indicative of AGN activity (see Section 1.2.1). While X-ray and radio emission can be originated by star-forming processes due to X-ray binaries (e.g. Lehmer et al., 2016) and relativistic electrons accelerated by supernovae remnants (Condon, 1992), respectively, direct detections in the X-ray or radio at $z > 2$ should be dominated by AGN activity. In this work, LAEs are classified as candidate AGN if they are directly detected

in X-ray or radio (see Calhau et al., 2020), although we note this only gives a lower constrain on the total number of AGNs, as not all AGN have strong X-ray or radio emission.

1.5.2 Ages and metallicities

The age and metallicity of a population of stars influences the Lyman continuum output. As such, the strength of the Lyα emission will also be conditioned by intrinsic properties of the population. Schaerer (2003) used stellar synthesis models to show that high Lyα EW is associated with younger populations, lower metallicities and/or top heavy IMFs. LAEs, which are by definition galaxies with strong Lyα emission, will thus be typically associated with these properties. In Figure 1.7, we show the Lyα EW dependence on age. For any fixed metallicity, a higher Lyα EW_0 translates into a younger age. By lowering the metallicity, this dependence shifts to even higher Lyα EW_0. These trends are qualitatively the same for both a burst or constant star-formation history.

1.5.3 Size/morphology

The size or morphology of a galaxy reveal on-going physical processes. Lyα photons suffer resonant scattering when escaping a galaxy, increasing the random walks and the likelihood that a Lyα photon will get absorbed by dust, with the escape fraction being constrained by the morphology of a galaxy (see discussion in Paulino-Afonso et al., 2018). LAEs are found to be very compact in the UV (Malhotra et al., 2012; Paulino-Afonso et al., 2018). This may suggest that a compact morphology favours the escape of Lyα photons. In Figure 1.8, we show the UV sizes at different redshifts of continuum-selected galaxies and LAEs. UV sizes of continuum-selected galaxies are typically much larger but at $z > 4$ they become very similar to the sizes of LAEs, which are always small. This suggests that typical star-forming galaxies become LAE-like at high-redshift, possibly due to a bigger overlap between the two populations.

On the other hand, Lyα emission is very extended. This is a consequence of the resonant scattering that Lyα photons suffer in the gas that surrounds galaxies,
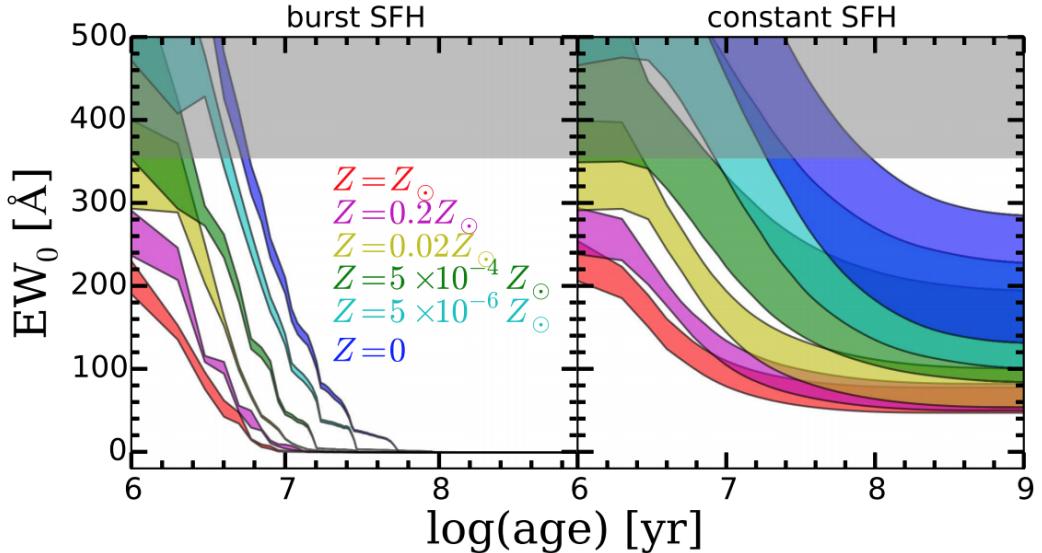
Figure 1.7: Lyα EW_0 dependence on age. For a fixed metallicity, younger populations have higher Lyα EWs. In addition, lower metallicities result in higher Lyα EWs, for both burst and constant star-formation history (SFH). Credit Hashimoto et al. (2017).

creating Lyα halos (e.g. Momose et al., 2014). In Figure 1.9, we show Lyα emission, compared to Hα and UV-continuum emission in a sample of $z = 2.23$ Hα-selected galaxies (Matthee et al., 2016). Lyα emission is extended over much larger scales than both UV continuum and the non-resonant line Hα.

1.5.4 Line profiles

The spectral profile of the Lyα emission line reveals the intrinsic conditions of a galaxy and surrounding IGM/CGM, such that the escape of Lyα photons is possible at the wavelengths given by the shape of the profile (see e.g. Hayes et al., 2020). The Lyα profile is modulated by the kinematics of the HII region where it is produced (Section 1.3.2), but also modulated by the resonant scattering that occurs within such HII regions. The constant scattering will lead to a broadening of the line and a subsequent wing scattering (Adams, 1972; Neufeld, 1990), which combined with the higher optical depth at the line centre leads to a double peaked profile. The separation of peaks will depend on the column density. Furthermore,
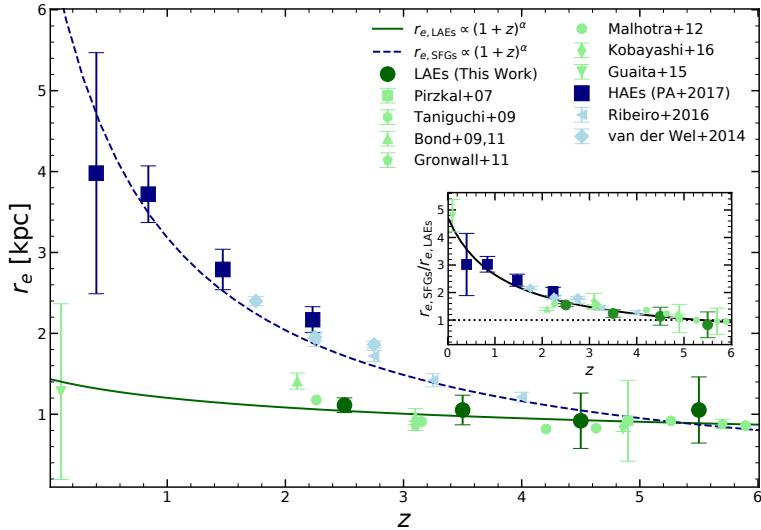
Figure 1.8: UV sizes (effective radius, r_e) of LAEs and continuum-selected samples. LAEs are found to always be compact and show little evolution with redshift. Continuum-selected galaxies have much larger r_e than LAEs at $z < 3$ but become smaller with increasing redshift and converge to the same r_e at $z \sim 5$. Credit Paulino-Afonso et al. (2018).
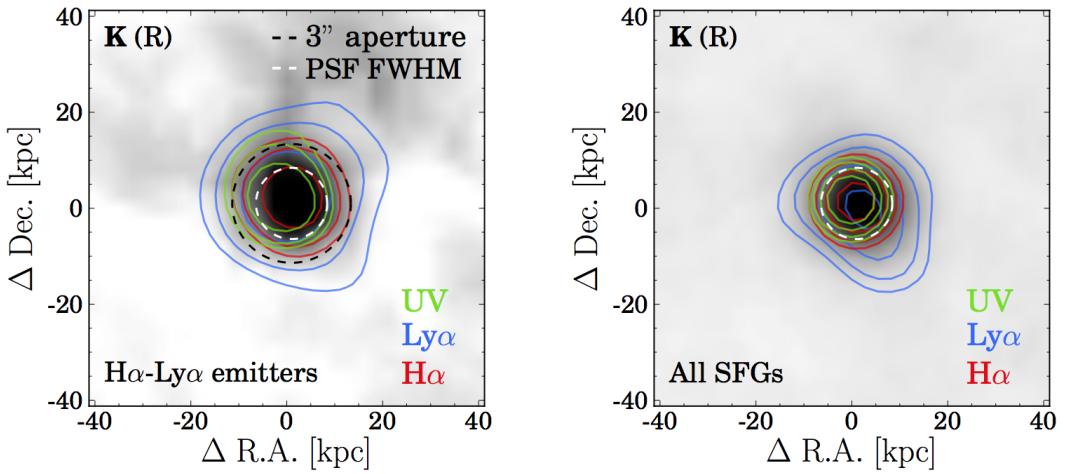
Figure 1.9: Lyα extended emission, compared to Hα and UV-continuum emission for a sample of $z = 2.23$ Hα-selected galaxies. Credit Matthee et al. (2016).

due to the preferred absorption at bluer wavelengths, the blue peak will often get attenuated, leading to a typical Lyα line profile consisting of a single asymmetric red peak. A strong blue peak is associated with a low gas column density, with the detection of a blue peak becoming exceedingly rare at $z > 6$, when the IGM is still mostly neutral. However, a detection of such a blue peak has been reported at $z = 6.6$ (Hu et al., 2016; Matthee et al., 2018), which reveals the powerful dynamics of LAEs which can carve ionising bubbles around them.

1.5.5 Luminosity functions

The luminosity function (LF), i.e. the number density of sources per luminosity bin, of LAEs and its evolution with redshift provides direct insight in the LAE population. Earlier works have obtained LFs of LAEs in individual redshift slices between $z = 3$ and $z = 6$ (e.g. Malhotra & Rhoads, 2004; Murayama et al., 2007; Ouchi et al., 2008), with Ouchi et al. (2008) reporting little redshift evolution at $z \sim 3 - 6$. On the other hand, the characteristic number density of the UV LF of LBGs decreases with redshift (Bouwens et al., 2015; Finkelstein et al., 2015), which in a first approach may not be expected since UV and Lyα are associated with the same star-forming processes (both trace emission from very massive O and B stars, see Section 1.2.3). Such distinct trends of evolution can however be explained by an evolution of the escape fraction of Lyα photons, lower dust content with increasing redshift, an evolution of the ionising efficiency and/or a combination of all. It is also important to explore whether these trends hold for larger volumes, which both reduce cosmic variance but also allow the probe of the most luminous regimes.

Matthee et al. (2015) and Santos et al. (2016) explored a large volume to select LAEs at two individual redshift slices ($z = 6.6$ and $z = 5.7$, repectively) and to probe the evolution of the LF near the epoch of reionisation. We show in Figure 1.10 the Lyα LF at $z = 5.7$ and $z = 6.6$ (and beyond). Previous works suggested that due to Lyα being easily shut down by neutral hydrogen, the number densities of LAEs would very significantly drop at at $z > 6$. However, using a very wide area approach, it was shown that even at $z = 6.6$ there are still many extremely luminous LAEs e.g. CR7, MASOSA (Sobral et al., 2015), VR7
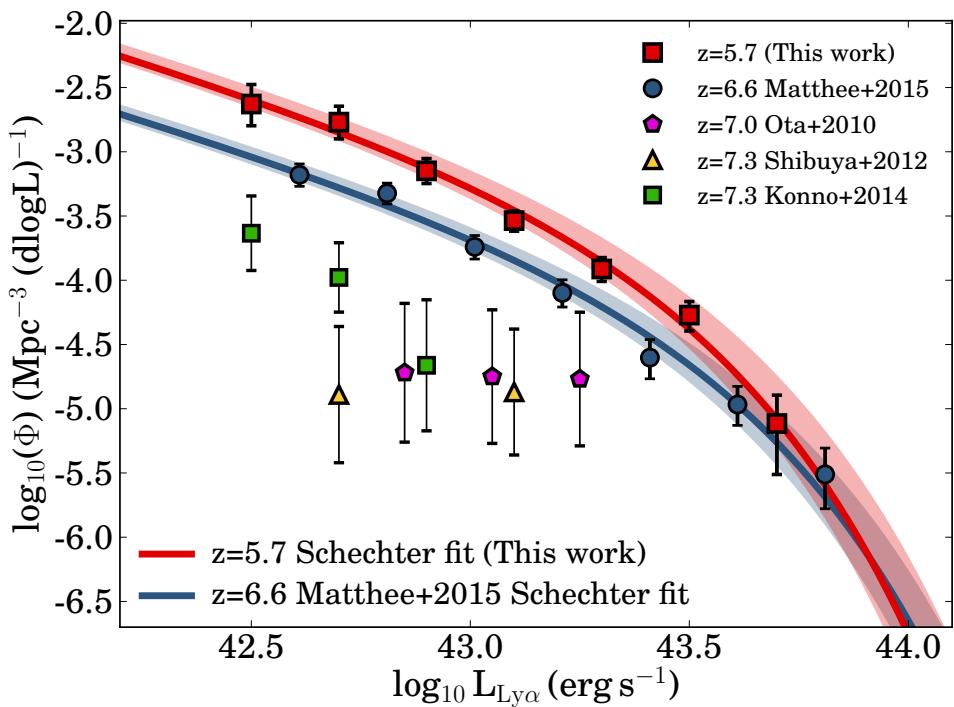
Figure 1.10: Evolution of the Lyα LF from $z = 5.7$ to $z = 6.6$. There is a strong decrease with increasing redshift in the number density of faint LAEs. However, there is little evolution in the number densities of bright LAEs, which can be explained by a preferential reionisation around the brightest sources, which can ionise the IGM around them and be observed. Credit Santos et al. (2016).

(Matthee et al., 2016) and Himiko (Ouchi et al., 2010), among others (Matthee et al., 2017c). Such sources can only be observed at those redshift ranges if they can carve ionised large bubbles around them, allowing Lyα to redshift out of the resonance wavelength and escape. Santos et al. (2016) report that the number densities of LAEs drops substantially for the faint end, but there is little evolution on the bright end. This differential evolution of the LF of LAEs, together with measurements of the extent of Lyα emission favour the line of thought that very bright sources are capable of ionising their own bubbles, and allow the escape of Lyα photons even at $z = 6.6$, while faint sources are not capable of doing it, and will only be observed if they are inside the bubbles of other sources of if they are clustered.

1.6 This Thesis

In this thesis, we aim to improve our current understanding of the evolution of galaxies properties at $z \sim 2 - 6$, particularly focusing on young star-forming galaxies. We achieve this by constructing a large sample of Lyα emitting galaxies, which we select with a compilation of narrow and medium bands over the $2 \deg^2$ of the COSMOS field. We use this sample to measure the properties of these galaxies, such as number densities, stellar masses and star formation rates and estimate how they evolve with redshift.

This thesis can be summarised in the following key points:

- constructing a large sample of ~ 4000 LAEs at 16 specific redshift slices between $z \sim 2$ and $z \sim 6$, selected over large co-moving volumes with similar selection methods (Chapter 2)

- building Lyα luminosity functions at $z \sim 2$ to $z \sim 6$ and measuring their evolution with redshift (Chapter 3)

- estimating the evolution of the Lyα escape fraction from the integrated luminosity functions (Chapter 4)

- conducting aperture photometry in 34 individual filters from rest-frame UV to FIR to measure the spectral energy distribution of our SC4K sample of LAEs (Chapter 5)

- measuring how SED-derived properties (stellar mass, UV luminosity, UV slopes) evolve with redshift (Chapter 6)

- building UV luminosity and stellar mass functions of LAEs at $z \sim 2$ to $z \sim 6$ and measuring their evolution with redshift (Chapter 7)

We present our final conclusions in Chapter 8, and also discuss possible future work and questions still unsolved.

Throughout this thesis we use a ΛCDM cosmology with $H_0 = 70 \, \mathrm{km \, s^{-1}}$ $\mathrm{Mpc^{-1}}$, $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$. All magnitudes are presented in the AB system (Oke & Gunn, 1983). In Chapters 2-4 we use a Salpeter (Salpeter, 1955) initial mass function (IMF), and in Chapters 5-7 we use a Chabrier (Chabrier, 2003) IMF.

Chapter 2

Slicing COSMOS with SC4K: constructing a sample of z ∼ 2 − 6 Lyα emitters

Abstract

We present and explore deep narrow and medium band data obtained with the Subaru and the Isaac Newton telescopes in the $\sim 2\,\mathrm{deg}^2$ COSMOS field. We use these data as an extremely wide, low-resolution ($R \sim 20 - 80$) IFU survey to slice through the COSMOS field and obtain a large sample of ~ 4000 Lyα emitters (LAEs) from $z \sim 2$ to $z \sim 6$ in 16 redshift slices (SC4K). This unique sample of star-forming galaxies (and AGNs), provides a unique opportunity to probe for galaxy evolution in a sample selected by the same selection methods.

2.1 Introduction

Understanding how galaxies form and evolve across cosmic time is a complex challenge which requires identifying and studying the inter-dependencies of key physical mechanisms over a range of environments (see e.g. Crain et al., 2015; Henriques et al., 2015; Muldrew et al., 2018; Schaye et al., 2015), informed by a variety of observations (e.g. Muzzin et al., 2013). It is now well established that the star formation rate density (SFRD) of the Universe evolves with redshift, peaking at $z \sim 2-3$ (e.g. Karim et al., 2011; Lilly et al., 1996; Madau & Dickinson, 2014; Sobral et al., 2013) and declining at even higher redshift (e.g. Bouwens et al., 2015; Khostovan et al., 2015), but several questions related to the physics of such evolution remain unanswered.

In order to unveil the evolution of physical properties of galaxies and active galactic nuclei (AGN) across time one requires self-consistent selection methods which can be applied across redshift. The Lyman Break selection (e.g. Giavalisco et al., 1996; Koo & Kron, 1980; Steidel & Hamilton, 1993) has been successfully used to produce large samples of galaxies up to $z \sim 10$ (e.g. Bielby et al., 2016; Bouwens et al., 2014a,b; Ellis et al., 2013; Finkelstein, 2016; McLure et al., 2010) through extremely deep optical to near-infrared (NIR) observations. However, UV-continuum selected samples using the Hubble Space Telescope (*HST*) are typically too faint (e.g. Atek et al., 2015) for extensive spectroscopic follow-up, particularly when probing distant look-back times (but large area surveys can still provide ideal follow-up targets e.g. Bowler et al., 2014, 2017). One alternative is to select galaxies by their hydrogen nebular recombination lines, such as Hα in the rest-frame optical (e.g. Colbert et al., 2013; Sobral et al., 2013) or Lyman-α (Lyα; $\lambda_0 = 1215.67\,\text{Å}$) in the rest-frame UV.

Lyα is intrinsically the strongest emission line in the rest-frame optical and UV (e.g. Partridge & Peebles, 1967; Pritchet, 1994) and it is routinely used to select high redshift sources ($z \sim 2-7$; see e.g. Malhotra & Rhoads 2004). Lyα is expected to be emitted by young star-forming galaxies (e.g. Charlot & Fall, 1993; Pritchet, 1994), but it is also observed around AGN (e.g. Miley & De Breuck, 2008). Searches for Lyα emitters (LAEs) have created samples of thousands of galaxies/AGN, including sources that are too faint to be detected by continuum

based searches (e.g. Bacon et al., 2015). The techniques used to detect LAEs include narrow band surveys (e.g. Hu et al., 2010; Konno et al., 2018; Matthee et al., 2015; Ouchi et al., 2008; Rhoads et al., 2000; Westra et al., 2006; Zheng et al., 2017), Integral Field Unit (IFU) surveys (e.g. Bacon et al., 2015; Drake et al., 2017a; van Breukelen et al., 2005) and blind slit spectroscopy (e.g. Cassata et al., 2011, 2015; Martin & Sawicki, 2004; Rauch et al., 2008). Galaxies selected through their Lyα emission allow for easy spectroscopic follow-up due to their high EWs (e.g. Hashimoto et al., 2017) and typically probe low stellar masses (see e.g. Gawiser et al., 2007; Hagen et al., 2016; Oyarzún et al., 2017). Narrow band and/or IFU surveys have the added benefit of being truly blind, and thus allow a good assessment of the volume and selection completeness.

Unfortunately, inferring intrinsic properties of galaxies from Lyα observations alone can be challenging due to the highly complex resonant nature of this emission line (for a review on the physics of Lyα radiative transfer see e.g. Dijkstra, 2017, and see also Section 1.3.3). A significant fraction of Lyα photons is scattered by the Inter-Stellar Medium (ISM), increasing the likelihood of being absorbed by dust, and in the Circum-Galactic Medium (CGM) as evidenced by the presence of extended Lyα halos (e.g. Momose et al., 2014; Wisotzki et al., 2016). Therefore, the Lyα escape fraction[1] (f_{esc}; see e.g. Atek et al., 2008, and see also Section 1.3.3), the ratio between the observed and the intrinsically produced Lyα luminosity from a galaxy, is still poorly understood quantitatively. New studies are now directly measuring f_{esc} of large samples of galaxies and over a range of redshifts by obtaining Hα and Lyα observations simultaneously (see Harikane et al., 2018; Matthee et al., 2016; Nakajima et al., 2012; Sobral et al., 2017). For example, f_{esc} is found to be anti-correlated with stellar mass (e.g. Matthee et al., 2016; Oyarzún et al., 2017), dust attenuation (e.g. Hayes et al., 2011; Matthee et al., 2016; Verhamme et al., 2008) and SFR (e.g. Matthee et al., 2016). Interestingly, the Lyα EW$_0$ seems to be the simplest empirical predictor of f_{esc} in LAEs with a relation that shows no evolution from $z \sim 0$ to $z \sim 2$ (Sobral et al., 2017) and that may remain the same all the way to $z \sim 5$ (Harikane et al., 2018).

---

[1]Throughout this study we use f_{esc} to quantify the escape fraction of Lyα photons, not Lyman-continuum photons.

"Typical" star-forming galaxies at $z \sim 2$ have low f_{esc} ($\sim 1-5\%$; e.g. Cassata et al. 2015; Oteo et al. 2015), likely because the dust present in their ISM easily absorbs Lyα photons (e.g. Ciardullo et al., 2014; Oteo et al., 2015; Oyarzún et al., 2017) and prevents most Lyα emission from escaping (see e.g. Song et al., 2014). However, sources selected through their Lyα emission typically have ~ 10 times higher escape fractions (e.g. Sobral et al., 2017; Song et al., 2014), with Lyα escaping over $\approx 2\times$ larger radii than Hα (e.g. Sobral et al., 2017). Furthermore, due to the sensitivity of f_{esc} to neutral hydrogen, Lyα can be used as a proxy of the ISM neutral gas (HI) content (Konno et al., 2016; Trainor et al., 2015) and the dust content (Hayes et al., 2011).

Statistically, the number density of LAEs as a function of luminosity (the luminosity function, LF), encodes valuable information on the global properties of LAEs and Lyα emission. Observations have revealed that the Lyα LF remains roughly constant at $z \sim 3-6$ (e.g. Drake et al., 2017a; Ouchi et al., 2008; Santos et al., 2016). This is in principle unexpected, as the cosmic SFRD, as traced by the UV LF, drops significantly at those redshifts (e.g. Bouwens et al., 2015; Finkelstein et al., 2015) and implies that intrinsic properties of galaxies may be evolving, on average, with redshift. Those may include lower dust content, leading to a higher f_{esc} which could compensate for a lower intrinsic production of Lyα photons (e.g. Hayes et al., 2011; Konno et al., 2016). Another possibility is that ξ_{ion}, which measures the ratio between ionising (LyC) and UV flux density increases with redshift (e.g. Duncan & Conselice, 2015; Khostovan et al., 2016; Matthee et al., 2017a). In practice, a combined increase of both ξ_{ion} and f_{esc} is also possible, which could tell us about an evolution of both the typical stellar populations/burstiness but also on the evolving physics/ISM conditions) of the escape of Lyα photons.

In this work, we use 16 different narrow and medium band filters (NB/MB) over the COSMOS field to select a large sample of LAEs in a total co-moving volume of $6.4 \times 10^7 \, \text{Mpc}^3$ and a wide redshift range of $z \sim 2-6$, addressing the current shortcomings of deep, small area surveys. Our survey can be seen as a very wide ($\approx 2 \, \text{deg}^2$), low resolution IFU survey between 400-850 nm, probing LAEs from the end of the epoch of re-ionisation at $z \sim 6$ (e.g. Fan et al., 2006) to the peak of star-formation history at $z \sim 2-3$.

We structure this Chapter as follows: Section 2.2 presents the data and the extraction of sources. Section 2.3 presents the selection of line emitters, the criteria we applied to select LAE candidates at $z \sim 2-6$ and the final SC4K sample. We present a brief summary in Section 2.4.

2.2 Data and source extraction

The COSMOS field (Capak et al., 2007; Scoville et al., 2007) is one of the most widely studied regions of extragalactic sky, with a plethora of publicly available multi-wavelength coverage. Data in COSMOS include X-ray, UV, optical, NIR, FIR and radio (see e.g. Civano et al., 2016; Ilbert et al., 2009; Laigle et al., 2016; Smolčić et al., 2017). We explore a range of narrow and medium band data (Capak et al., 2007; Matthee et al., 2017b; Santos et al., 2016; Sobral et al., 2017; Taniguchi et al., 2007, 2015a) over roughly the full COSMOS field in order to 1) create detection-catalogues for each band, 2) identify sources with strong excess emission in those bands relative to their broad band counterparts and 3) obtain dual-mode photometry on all other bands available in order to further constrain the (photometric-)redshift of each source. In Figure 2.1 we show the filter profiles of all the 12 medium bands (MBs) and the 4 narrow bands (NBs) used in this work. These filters are capable of detecting various emission lines, particularly redshifted Lyα spanning a wide redshift range, from $z \sim 2$ to $z \sim 6$.

2.2.1 Medium band data

We retrieve the publicly available fully reduced medium band data (see Table 2.1 and Figure 2.1) from the COSMOS archive (see Ilbert et al., 2009; Taniguchi et al., 2015a). All data were obtained with the Suprime-Cam (S-Cam) instrument on the Subaru Telescope (Miyazaki et al., 2002). The data were taken with seeing conditions varying from $0.6''$ to $1.0''$, with an overall FWHM of $0.8 \pm 0.1''$ (see also Muzzin et al., 2013; Taniguchi et al., 2015a). The images have a roughly similar average depth (Muzzin et al., 2013) but with some exceptions (see Table 2.1), varying from $26.2\,\mathrm{mag}$ (deepest: IA427, IA484 and IA527) to $25.4\,\mathrm{mag}$
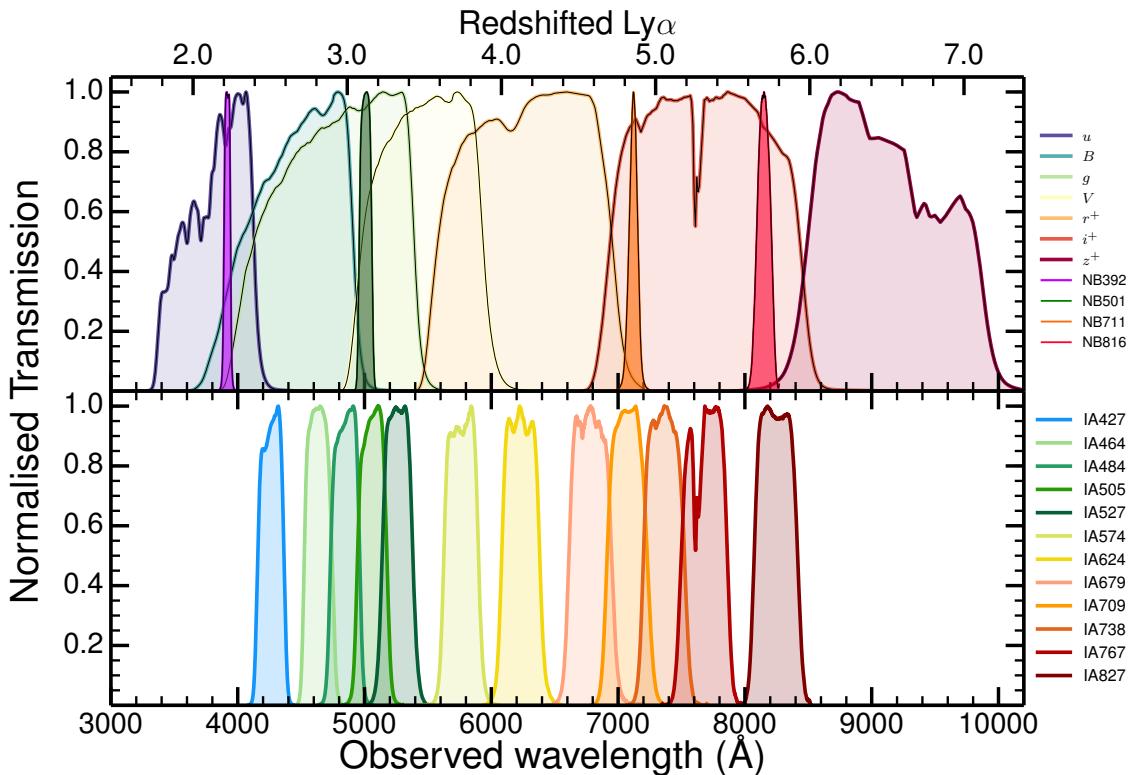
Figure 2.1: Normalised filter profiles used in this work. The top axis indicates the redshift of Lyα emission for the corresponding observed wavelength. *Top:* The broad bands (u, B, g, V, r^+, i^+ and z^+; typical FWHM $\sim$ 100 nm) which we use to estimate the continuum for candidate line-emitters and four narrow bands which we also present in our final catalogue (NB392, NB501, NB711 and NB816; typical FWHM $\sim$ 10 nm). *Bottom:* The 12 medium bands used in this study (typical FWHM $\sim$ 30 nm; see Table 2.1) which are sensitive to Lyα emission from $z \sim 2.5$ to $z \sim 5.8$. Note that some of the medium band filters overlap slightly, which can result in some sources being detected as LAEs in two consecutive medium bands, although we note that the overlapping volume is always relatively small.

Table 2.1: The medium band filters (see Taniguchi et al., 2015a) and the depth of the data obtained with them, measured directly ($3\,\sigma$; $5\,\sigma$ can be obtained by subtracting 0.55) in $2''$ apertures and by Muzzin et al. 2013 (M13, $2.1''$). We also transform our measured $3\,\sigma$ limiting magnitude ($2''$) into a flux limit (in units of $\mathrm{erg\,s^{-1}\,cm^{-2}}$) in the case of the full flux within the $2''$ medium band aperture being from an emission line.

Medium Band	Ly$\alpha\,z$	λ_c [FWHM] (Å)	$3\,\sigma$ Depth (AB)	3σ Flux ($\times 10^{-17}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$)	$5\,\sigma$ (M13) (AB)
IA427	$2.42 - 2.59$	4263.5 [207.3]	26.1	4.6	26.1
IA464	$2.72 - 2.90$	4635.1 [218.1]	26.0	4.5	25.8
IA484	$2.89 - 3.08$	4849.2 [229.1]	26.1	3.9	26.1
IA505	$3.07 - 3.26$	5062.1 [231.5]	25.8	4.8	25.9
IA527	$3.23 - 3.43$	5261.1 [242.7]	26.1	3.5	26.1
IA574	$3.63 - 3.85$	5764.8 [272.8]	25.9	4.0	25.7
IA624	$4.00 - 4.25$	6232.9 [299.9]	25.8	4.1	25.9
IA679	$4.44 - 4.72$	6781.1 [335.9]	25.7	4.3	25.6
IA709	$4.69 - 4.95$	7073.6 [316.3]	25.8	3.4	25.8
IA738	$4.92 - 5.19$	7361.5 [323.8]	25.7	3.5	25.6
IA767	$5.17 - 5.47$	7684.9 [365.0]	25.7	3.6	25.4
IA827	$5.64 - 5.92$	8244.5 [342.8]	25.7	3.0	25.4

Table 2.2: The estimated depth of broad band data used in our analysis (3σ). We measure these by placing 100,000 random $2''$ empty apertures, and computing the standard deviation of the counts and converting it to magnitudes. The $2\,\sigma$ and $4\,\sigma$ limits can be obtained by adding 0.44 and subtracting 0.31, respectively.

$u_{3\sigma}$	$B_{3\sigma}$	$V_{3\sigma}$	$g_{3\sigma}$	$r^+_{3\sigma}$	$i^+_{3\sigma}$	$z^+_{3\sigma}$
26.81	27.21	26.50	26.61	26.55	26.12	25.23

(shallowest, IA767), measured in 2.1″ apertures (5σ, see Muzzin et al., 2013). We also obtain our own depth measurements by placing 100,000 empty/random 2″ apertures in each of the (native) images and determining the standard deviation. The results are presented in Table 2.1. and, for reference, we also show the depths measured in Muzzin et al. (2013), who used PSF-matched data (making a direct comparison non-trivial).

2.2.2 Narrow band data

We complement our medium band data with four narrow band studies in the COSMOS field: the CALYMHA survey at $z = 2.2$ (Sobral et al., 2017) and a $z = 3.1$ survey (Matthee et al., 2017b) using the narrow band filters NB392 and NB501, respectively, both mounted on the 2.5 m Isaac Newton Telescope's WFC. The NB392 data ($\lambda_c = 392$ nm, $\Delta\lambda = 5.2$ nm; Sobral et al. 2017) have a 5σ depth of 23.7-24.5 AB magnitude and a typical PSF-FWHM of 1.8″. The NB501 data ($\lambda_c = 501$ nm, $\Delta\lambda = 10$ nm) were taken and reduced with a similar strategy and data-quality as the NB501 data described in Matthee et al. (2017b) and have a typical 5σ depth of 24.0 AB magnitude with 1.6″ PSF-FWHM. Limiting magnitudes for NB392 and NB501 data were measured with 3″ apertures.

In addition, we also use two narrow band surveys exploring S-Cam data: $z = 4.8$ and $z = 5.7$ (Santos et al., 2016); these have used the narrow band filters NB711 and NB816, respectively (see Figure 2.1). We note that all NB and MB selected catalogues have been obtained in similar ways, which we describe in Section 2.2.4.

2.2.3 Broad band data

We retrieve fully reduced archival broad band (BB) data from the COSMOS archive. The public archival data consists of mosaics covering the full COSMOS field, from rest-frame UV to NIR, in the BB filters u, B, g, V, r^+, i^+, z^+, Y, J, H and K (Capak et al., 2007; McCracken et al., 2012; Taniguchi et al., 2007). We provide the 3σ limiting magnitudes in Table 2.2. All data were obtained with the Suprime-Cam (S-Cam) instrument on the Subaru Telescope (Miyazaki et al.,

2002). This BB data is necessary to apply our LAE selection criteria, as fully detailed in Section 2.3.3.

2.2.4 Extraction of sources

To produce the narrow or medium band selected catalogues[1] (see e.g. Matthee et al., 2017b), we follow Santos et al. (2016). Briefly, we start by registering the u, B, g, V, r^+, i^+, z^+, Y, J, H and K broad band data and all the medium band (or narrow band) images to a common astrometric reference frame using SWARP (Bertin et al., 2002). We extract sources with a primary $2''$ aperture[2] (but we note we also extract them with multiple apertures, including MAG-AUTO, a proxy of the total magnitude) using SEXTRACTOR (Bertin & Arnouts, 1996) in dual-image mode, and with each of the medium band images as the detection image. Therefore, for each medium band, we create a catalogue with all the detections on that band, and with the broad band photometry extracted at the coordinates of each detection. We thus note that our selection is purely based on the detection of a source in a medium or narrow band, independently of its continuum strength.

Before creating our final catalogues, we investigate the need for any significant masking to remove low quality regions and diffraction patterns around bright stars. In addition to removing such regions, we also find that there is a small area in the corner of the COSMOS field (≈ 0.02 deg^2) for which there is no u-band data. Given that we require blue photometry to select LAEs and reject lower redshift sources (see Tables 2.2 and 2.3), we mask/exclude this region for filters bluer than IA574. After masking, the contiguous survey area is 1.94-1.96 deg^2 for the medium band filters and 1.96 deg^2 for the NB711 and NB816 filters, while the area covered by the NB392 and NB501 data is 1.21 deg^2 and 0.85 deg^2, respectively (Matthee et al., 2016; Sobral et al., 2017).

---

[1]We conduct the full data analysis and photometry extraction for all the medium bands (and broad bands), but the data reduction and selection of line-emitters/LAEs in narrow band images was conducted in previous works (Matthee et al., 2017b; Santos et al., 2016; Sobral et al., 2017). In this thesis, we combine the narrow band samples with the new medium band samples, and build the SC4K sample of LAEs.

[2]Because the NB392 and NB501 data have a broader PSF, photometry has been done with $3''$ apertures.
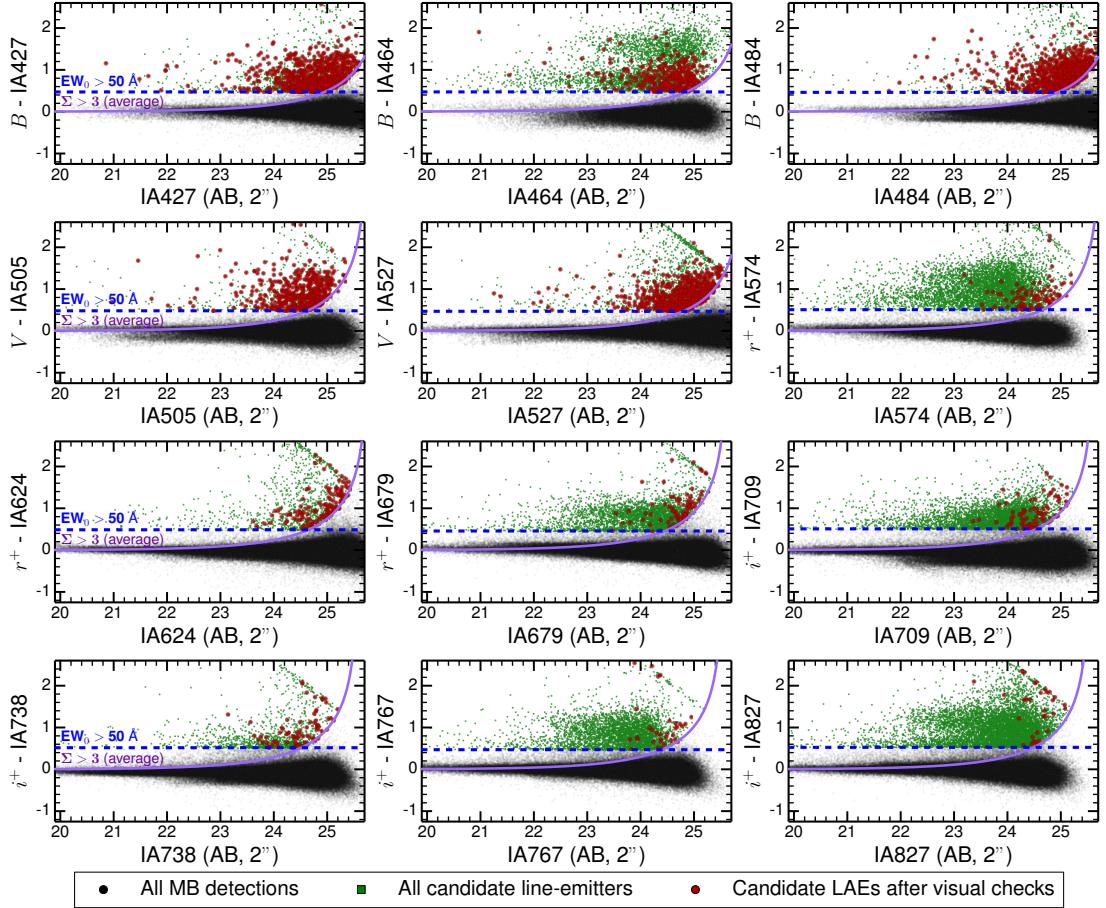
Figure 2.2: The colour-magnitude diagrams used for the selection of line-emitters for the 12 medium bands. Each medium band magnitude is plotted versus the excess colour, and we identify sources with a high enough EW (corresponding to a rest-frame EW of > 50 Å for LAEs) and with a significant excess (average $\Sigma > 3$). The selection criteria of LAEs are presented in Table 2.3. MB detections are shown in black, candidate line-emitters (prior to individual visual checks) are shown in green and candidate LAEs in red (after visually checking all of them). We assign the broad band detection limit to sources with no broad band detection. It is clear from the panels that, on average, the amount and fraction of LAEs greatly decreases from the bluer (where almost all candidate line-emitters are LAEs) to the redder filters (where only a small fraction is consistent with being a LAE). The clustering of points at the top of each panel is an artificial effect of collapsing the points with no broad band detection above 3σ.

2.3 Selection criteria: SC4K

2.3.1 Selection of candidate line-emitters

In order to identify sources with candidate emission lines out of all medium band selected sources, it is necessary to estimate the continuum of each source. As the central wavelengths of medium bands are typically offset (see Figure 2.1) from the central wavelengths of their overlapping broad band, we need to investigate and apply a correction to the medium band photometry (MB_0). This step/correction assures that a measured medium band excess is not dependent on the intrinsic slope of the continuum (estimated with two broad band magnitudes, $BB - BB_{adjacent}$), similar to corrections applied for narrow band surveys (e.g. Sobral et al., 2013; Vilella-Rojo et al., 2015). Without such correction, sources with significant colours could mimic emission lines. In practice this requires re-calibrating either MB_0 or BB photometry (or producing a new artificial BB magnitude) to assure that, on average, sources without an emission-line will have a zero colour excess ($BB - MB \approx 0$) regardless of their continuum colour ($BB - BB_{adjacent}$). We do this by evaluating the colour dependence of $BB - MB_0$ on $BB - BB_{adjacent}$ and parameterising it as (calculating m and b):

$$BB - MB_0 = m \times (BB - BB_{adjacent}) + b \tag{2.1}$$

We then use coefficients m and b to finally obtain:

$$MB = MB_0 + (m \times (BB - BB_{adjacent}) + b) \tag{2.2}$$

with the filters listed in Table 2.3. The coefficients m and b are provided in Table A.2. We note that for some filter combinations both m and b are effectively zero. For sources without $BB - BB_{adjacent}$ ($< 2\sigma$ detection in either band) we compute a median correction which we apply per medium band filter. Typical median corrections are at the 0.1 mag level and in the 0.0-0.3 range (see Table A.2).

For the selection of line-emitters we follow the same methodology used for narrow band surveys (e.g. Matthee et al., 2017b; Sobral et al., 2013, 2017), based on
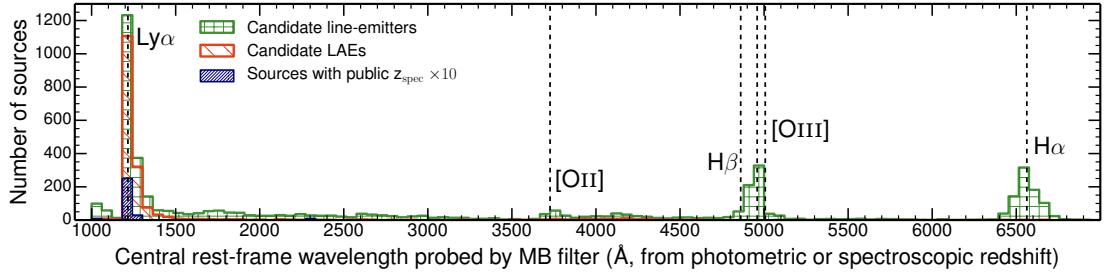
Figure 2.3: Distribution of the central rest-frame wavelengths probed by the different medium band filters (based on photometric or spectroscopic redshifts) of our continuum-bright (with accurate photometric redshifts) sample of line-emitters (green histogram) and its subset containing only our final LAEs (red histogram). The photometric redshifts used in this figure are taken from Laigle et al. (2016). The black dashed lines are the rest-frame wavelengths of the main emission lines probed. Our sample of continuum-bright line-emitters with photometric redshift is predominantly composed by Lyα emitters, followed by a population of Hα emitters, [OIII]+Hβ emitters, and finally [OII] emitters. We find that our Lyα selection criteria is able to remove the vast majority of lower redshift contaminants whilst maintaining the bulk of Lyα photometric candidates (see Section 2.3.3).

two main parameters: the emission-line equivalent width (EW), and the emission-line or excess significance (Σ; e.g. Bunker et al., 1995). In a first approximation, assuming no background noise in the images, we could define a line-emitter as being any source which follows:

$$f_{\mathrm{MB}} - f_{\mathrm{BB}} > 0, \tag{2.3}$$

where f_{MB} and f_{BB} are the flux densities[1] measured for the two filters.

As the uncertainties in the background noise cannot be neglected, f_{MB}-f_{BB} needs to be above some significance parameter. The parameter Σ is thus introduced into Equation 2.3.1:

$$f_{\mathrm{MB}} - f_{\mathrm{BB}} > \Sigma \times \sqrt{\mathrm{rms}_{\mathrm{BB}}^2 + \mathrm{rms}_{\mathrm{MB}}^2}, \tag{2.4}$$

where Σ quantifies how significantly above the noise a given BB-MB colour excess

---

[1]Note that as a consequence of the way we define/correct MB magnitudes, their flux densities (F_λ) need to be calculated with the same effective wavelength as the corresponding BB.

is (a high excess can be explained by the presence of an emission line at the MB wavelength), and $\mathrm{rms}_{\mathrm{BB}}$ and $\mathrm{rms}_{\mathrm{MB}}$ are the root mean square of the noise measured in the BB and MB images, respectively.

By converting fluxes to magnitudes, Σ can be written as (Sobral et al., 2013):

$$\Sigma = \frac{1 - 10^{-0.4(\mathrm{BB}-\mathrm{MB})}}{10^{-0.4(ZP-\mathrm{MB})}\sqrt{\mathrm{rms}_{\mathrm{BB}}^2 + \mathrm{rms}_{\mathrm{MB}}^2}}, \tag{2.5}$$

where BB and MB are the broad and the medium band magnitudes and ZP is the zero-point of the image. We estimate $\mathrm{rms}_{\mathrm{MB}}$ and $\mathrm{rms}_{\mathrm{BB}}$ by randomly placing $2''$ apertures in the appropriate images and determining the standard deviation per image. This approach takes spatially correlated noise into account. We apply an emission-line significance threshold of $\Sigma > 3$, similarly to other studies (e.g. Matthee et al., 2015). In addition to Σ, we also measure the observed EW ($\mathrm{EW}_{\mathrm{obs}}$) of potential lines as:

$$\mathrm{EW}_{\mathrm{obs}} = \Delta\lambda_{\mathrm{MB}}\frac{\mathrm{f}_{\mathrm{MB}} - \mathrm{f}_{\mathrm{BB}}}{\mathrm{f}_{\mathrm{BB}} - \mathrm{f}_{\mathrm{MB}}(\Delta\lambda_{\mathrm{MB}}/\Delta\lambda_{\mathrm{BB}})}, \tag{2.6}$$

where $\Delta\lambda_{\mathrm{MB}}$ and $\Delta\lambda_{\mathrm{BB}}$ are the FWHM of the medium (see Table 2.1) and broad band filters (Capak et al., 2007; Taniguchi et al., 2015a). We show the selection cuts in Figure 2.2.

Typical narrow band surveys apply a Lyα rest-frame EW (EW_0) cut of $\approx 25\,\text{Å}$ (e.g. Ouchi et al., 2008), mostly to avoid contamination from other line-emitters, as Lyα is typically the line with the highest observed EW (but see also other high EW contaminants in e.g. Sobral et al., 2017). Recent surveys also explored lowering this cut, showing that a few extra real Lyα sources may be recovered in those cases, which can populate the bright end (Sobral et al., 2017, see also VUDS, e.g. Le Fèvre et al. 2015), but also introduce many extra contaminants. Given that we are using wider filters in comparison to the typical narrow band filters, we are forced to use a higher observed EW cut to retrieve clean samples of line-emitters. For our analysis, we find that setting an observed EW cut from $175\,\text{Å}$ to $340\,\text{Å}$ from the bluest (narrowest) to the reddest and broader filters is able to recover clean samples of line emitters and yields an homogeneous rest-frame

Lyα equivalent width cut of $EW_0 > 50\,\text{Å}$ for all of our medium bands. Note that our EW_0 cut (for LAEs) is about twice the typical used in narrow band surveys ($25\,\text{Å}$; see e.g. Santos et al. 2016), implying we are likely less contaminated by lower redshift line-emitters, but that we may be less complete. We take this into account when deriving completeness corrections, but we note that, in practice, the vast majority of LAEs at high redshift show $EW_0 > 50\,\text{Å}$; see e.g. Ouchi et al. (2008). For an in-depth analysis of selecting LAEs with different EW_0 cuts see Sobral et al. (2017).

The full selection procedure to search for candidate emission-line sources is illustrated in Figure 2.2, which shows the medium band colour excess versus medium band magnitude for each band. It can be seen that the EW threshold is well above the scatter at bright magnitudes ($\lesssim 23$). In total, we identify 40,726 potential line-emitters, with each medium band contributing with roughly 2,000-3,000 emitters to the sample. We note, nonetheless, that we expect our full sample of $\approx 40\,\text{k}$ candidate line-emitters to still be contaminated by e.g. artefacts around bright stars, cosmic rays, and due to other image defects. In order to fully address this possibility, we visually inspect every single source in our final sample (see Section 2.3.4), but we first filter out lower redshift emitters and isolate candidate LAEs.

2.3.2 Photometric and spectroscopic redshifts

In order to test how robust our emission-line selection criteria are, we use a large compilation of photometric and spectroscopic redshifts (e.g. Ilbert et al., 2009; Laigle et al., 2016; Lilly et al., 2007). For each medium band, we look at the distribution of publicly available photometric and spectroscopic redshifts of the line-emitter sample. By converting each photometric redshift value to a rest-frame wavelength assuming the source has an emission line at the central wavelength of the corresponding medium band, we obtain Figure 2.3. We find evidence for the presence of a large population of Hβ+[OIII]$_{5007}$ and Hα emitters, although the sample is dominated by candidate LAEs. Excluding the dominating LAEs, the most common remaining sources are Hα emitters, followed by [OIII]+Hβ. [OII]$_{3727}$ emitters represent a less frequent population among all the candidate

Table 2.3: The selection of LAEs from the sample of all line-emitters, using an observed EW threshold of EW $> 50 \times (1+z)\,$Å and $\Sigma > 3$. The relevant LAE colour selection is given in the table. Numbers of LAEs are given after visually inspecting all candidate LAEs and rejecting interlopers. As described in the Section 2.3.3, colour criteria are based on the Lyman break technique and removing sources with very red colours red-wards of the emission-line (which indicates that the potential Lyman break is actually a Balmer break and that the line is not Lyα). We note that we explicitly perturb these colour selections with Markov chain Monte Carlo (MCMC) simulations and include the results in the errors when we estimate luminosity functions. We remove 20 confirmed lower redshift line emitters/contaminants, as described in Section 2.3.5 and we expect a $\approx 10-20\%$ remaining contamination. Due to a small overlap of some of the medium filters in their wings a small number (53; $\sim 1\%$) of LAEs are detected as LAEs in two adjacent MB filters; these are kept in each of the filters for the full analysis. The excess filters are the two adjacent broad bands used to estimate the continuum (see Section 2.3.1). The total number of LAEs (3908) is the sum of LAEs selected in each filter, regardless of whether they are detected in more than one filter, as the selection was done independently. [1]EW$_0 > 5\,$Å; Sobral et al. (2017) [2]EW$_0 > 25\,$Å; Santos et al. (2016); Matthee et al. (2017b).

Selection filter	Excess filter	Lyα redshift FWHM	LAE colour selection (Section 2.3.3)	# LAE candidates
IA427	B (u)	$2.42-2.59$	$(u > u_{3\sigma} \vee u - B > 0.4)$ & $(B - r^+ < 0.5)$	741
IA464	B (V)	$2.72-2.90$	$(u > u_{3\sigma} \vee u - B > 0.5)$ & $(B - r^+ < 0.8)$	311
IA484	B (V)	$2.89-3.08$	$(u > u_{3\sigma} \vee u - B > 0.5)$ & $(B - r^+ < 0.75)$	711
IA505	V (B)	$3.07-3.26$	$(u > u_{3\sigma} \vee u - V > 1.3)$ & $(B - r^+ < 0.5)$	483
IA527	V (B)	$3.23-3.43$	$(u > u_{3\sigma} \vee u - V > 1.5)$ & $(V - i^+ < 1.0)$	641
IA574	r^+ (V)	$3.63-3.85$	$(u > u_{3\sigma})$ & $(B > B_{3\sigma} \vee B - r^+ > 1.0)$ & $(V - i^+ < 0.5)$	98
IA624	r^+ (i^+)	$4.00-4.25$	$(B > B_{3\sigma})$ & $(V > V_{3\sigma} \vee V - r^+ > 0.5)$ & $(r^+ - i^+ < 1.0)$	142
IA679	r^+ (i^+)	$4.44-4.72$	$(B > B_{3\sigma})$ & $(V > V_{3\sigma} \vee V - r^+ > 0.5)$ & $(r^+ - i^+ < 1.0)$	79
IA709	i^+ (r^+)	$4.69-4.95$	$(B > B_{3\sigma})$ & $(V > V_{3\sigma})$ & $(r^+ > r^+_{3\sigma} \vee r^+ - i^+ > 0.8)$ & $(i^+ - z^+ < 1.0)$	81
IA738	i^+ (r^+)	$4.92-5.19$	$(B > B_{3\sigma})$ & $(V > V_{3\sigma})$ & $(r^+ > r^+_{3\sigma} \vee r^+ - i^+ > 0.5)$ & $(i^+ - z^+ < 1.0)$	79
IA767	i^+ (z^+)	$5.17-5.47$	$(B > B_{3\sigma})$ & $(V > V_{3\sigma})$ & $(r^+ > r^+_{3\sigma} \vee r^+ - i^+ > 0.5)$ & $(i^+ - z^+ < 1.0)$	33
IA827	i^+ (z^+)	$5.64-5.92$	$(B > B_{3\sigma})$ & $(V > V_{3\sigma})$ & $(r^+ > r^+_{3\sigma} \vee r^+ - i^+ > 0.5)$ & $(i^+ - z^+ < 1.0)$	35
NB392[1]	u (B)	$2.20-2.24$	$(z - K) > (B - z) \vee z_{phot} = 1.7 - 2.8 \vee z_{spec} = 2.20 - 2.24$	159
NB501[2]	g^{+}[2]	$3.08-3.16$	$(u > u_{3\sigma} \vee u - g^+ > 1)$ & $(g^+ - i^+ < 1.5)$	45
NB711[2]	i^+ (z^+)	$4.83-4.89$	$(B > B_{2\sigma})$ & $(V > V_{2\sigma})$ & $[(r^+ > r^+_{2\sigma} \vee (r^+ < r^+_{2\sigma} \wedge r^+ - i^+ > 1.0)]$	78
NB816[2]	i^+ (z^+)	$5.65-5.75$	$(B > B_{2\sigma})$ & $(V > V_{2\sigma})$ & $[(r^+ > r^+_{2\sigma} \vee (r^+ < r^+_{2\sigma} \wedge r^+ - i^+ > 1.0)]$	192
Full SC4K sample (This study, 12 medium band + 4 narrow band)				**Total 3908**

line-emitters, and we also find evidence for some 4000 Å and Lyman break sources making it to the sample of potential line emitters. The relative proportion of sources is not surprising, given the combination of volume and observed EW distributions of all these lines (see e.g. Hayashi et al., 2018; Khostovan et al., 2016; Sobral et al., 2014).

While Figure 2.3 shows that our sample of high EW candidate line-emitters is dominated by LAE candidates, it also reveals that many other line-emitters are expected to be in the sample. This is confirmed by spectroscopic redshifts of the full sample and stresses the importance of excluding lower redshift emitters in order to produce relatively clean and complete samples of LAEs.

2.3.3 Selection of LAEs at z $\sim$ 2.5 − 6

In order to isolate LAEs from lower redshift line-emitters (see Figure 2.3), we apply two criteria. First, we identify the presence of a colour break blue-ward of the medium band excess emission and no significant emission bluer of that (see Table 2.2). Secondly, we remove sources that have red colours (e.g. $B - r > 0.5$ for $z \sim 2.5$; $i - z > 1.0$ for $z \sim 5.5$); see Table 2.3. The first step selects the Lyman break, while the second criterion removes sources likely to be stars or red galaxies with a strong Balmer break (at a rest-frame wavelength $\sim 400\,\mathrm{nm}$) that mimics the Lyman break (see e.g. Matthee et al., 2014, 2017c). Narrow band surveys for LAEs typically apply the same/similar standard criteria (e.g. Bielby et al., 2016; Matthee et al., 2015; Ouchi et al., 2008; Santos et al., 2016), with the difference being how strict the criteria/flexible the cuts are and what bands are available to trace/identify the Lyman break. Some surveys conducted in the blue bands rely mostly on a high EW_0 cut (e.g. Ciardullo et al., 2014; Konno et al., 2016), but as discussed in e.g. Sobral et al. (2017), even in the bluest bands it is crucial to filter lower redshift contaminants out of the sample of line-emitters due to bright, high EW lines such as CIII] and CIV (see Stroe et al., 2017a,b), particularly in wide-field surveys. We note that our colour criteria to exclude very red sources only removes extreme red sources and is based on current spectroscopic samples that show that essentially no real LAE will be removed by our cuts. However, it is possible that a handful of even more extremely red LAEs (which are interesting

on themselves; see e.g. Matthee et al., 2016; Ono et al., 2010; Taniguchi et al., 2015b) may be rejected in this way.

We apply our LAE selection by taking full advantage of the deep available broad band photometry (see Table 2.2), which covers the wavelengths of the Lyman break and the Lyman continuum for our entire redshift range (see Figure 2.1). Our colour selection criteria (see Table 2.3) are defined such that a candidate LAE is required to either have no detection blue-ward of the medium band (i.e. being a drop-out galaxy), or, if the continuum is bright enough, to have a strong colour break between the two broad bands adjacent to the Lyα break expected wavelength. By not applying too strict colour-criteria, we ensure that sources with Lyman-Werner radiation or Lyman continuum leakage are not removed from our sample, as long as they have a Lyman break. We note that such sources are typically AGN, with high spectroscopic completeness in currently available spectroscopic surveys in COSMOS (e.g. Lilly et al., 2007).

The exact values for each criterion are determined empirically using the large compilation of spectroscopic and reliable photometric redshifts discussed in Section 2.3.2, but we also perturb these in Section 3.1.4. When we apply our LAE selection criteria to all line-emitters (before any visual inspection and regardless of whether they have public photometric/spectroscopic redshifts), up to $\sim 50\%$ are selected as LAEs for the lower redshift slices ($z \sim 2-3$) but only $\sim 2-5\%$ of line emitters as LAEs for the highest redshift slices ($z \sim 5-6$). This is a consequence of the differences in luminosity depth in Lyα, but even more so due to the volumes and redshifts of the other main emission lines such as Hα, [OIII]+Hβ and [OII] which become more prominent for redder filters. Our results show that even with a high EW cut, we expect that about 50% of sources will not be Lyα in the bluest bands, while about 95-97% of sources in the red bands will be lower redshift line emitters[1] (see Figure 2.3). After the LAE selection, we retrieve a total of 6,156 potential Lyα emitters out of the full 40,726 potential line emitters (15%).

---

[1] Due to the Lyman break criteria, our survey (and all similar Lyα surveys) is strongly biased against galaxy-galaxy lensed LAEs, as any lower redshift galaxy lensing a distant LAE will be classed as a lower-redshift interloper and the lensing system rejected.

2.3.4 Visual inspection of all LAE candidates

In order to obtain a clean sample of LAE candidates, we visually inspect all the candidates for spurious detections in their corresponding medium band. In practice, we remove i) fake sources due to diffraction patterns, ii) fake sources which are selected close to the borders of images where the local noise is higher, iii) sources that are clear artefacts and iv) sources which are real but that clearly have their fluxes boosted in the medium bands due to bright halos or diffraction of nearby stars. This is the same approach taken in the large-area narrow band surveys which we also explore, namely Santos et al. (2016) and Sobral et al. (2017). From a total of 6,156 LAE candidates, we conservatively reject/exclude 2,703 sources, and end up with a sample of 3,453 LAEs. We note that due to very different local noise properties, artefacts and image quality/depth, some bands (e.g. IA574 and IA827) have very high spurious fractions of $\approx 90\%$ in the initial LAE candidate sample, while other bands such as IA427 and IA679 have lower spurious fractions of $\sim 15 - 25\%$. It is worth noting that due to the strict selection criteria in terms of non-detection in the optical in many bands, along with the high excess observed in the medium bands, we easily select every single spurious/artefact in the full COSMOS images/catalogue. We thus stress the importance of visually checking all sources for such wide area surveys or, at least, to visually check a representative sub-sample and apply statistical corrections. We take into account the removal of spurious sources when computing the total areas and volumes, but we note that these only remove up to $\approx 0.03\%$ of the total area and thus they are completely negligible.

2.3.5 Spectroscopic completeness, contamination and the final sample of LAEs

Figure 2.3 reveals that our sample of line-emitters (with available photometric redshifts) is mostly composed of LAEs, and we expect that our photometric selection will further remove contaminants. This can nonetheless be quantified/investigated by using a relatively large number of spectroscopic redshifts of i) the full set of line emitters and ii) our samples of LAEs. Ideally, a sample

that is highly complete will show that essentially all spectroscopically confirmed LAEs in i) will be contained in sample ii), while a highly clean sample will see most of contaminants in i) not be selected for ii).

We compile a large sample of spectroscopic redshifts in the COSMOS field (e.g. Cassata et al., 2015; Kriek et al., 2015; Le Fèvre et al., 2015; Lilly et al., 2007; Shioya et al., 2009) to find that 132 sources within our sample of LAE candidates have a spectroscopic redshift. Out of the 132 sources, we confirm 112 as LAEs in the appropriate band. This suggests a contamination of real line-emitters which are incorrectly selected as LAEs of about 15%, well within the range of what is typically found for large area Lyα narrow band surveys at similar redshifts (e.g. Harikane et al., 2018; Santos et al., 2016; Shibuya et al., 2018; Sobral et al., 2017). We also investigate whether there is any significant dependence of this contamination rate on redshift, Lyα luminosity or EW$_0$. We find that within the Poissonian errors the contamination is found to be relatively constant and to be between 10-20%, similar to those found for narrow band surveys of LAEs (e.g. Bielby et al., 2016; Ouchi et al., 2008). In Appendix A.2 we provide further evidence of low contamination in typical $H-K_\mathrm{s}$ colours of $z \sim 3$ LAEs. There are only mild indications that the higher redshift and the highest luminosity samples may be slightly more contaminated (similarly to what has been found/discussed in e.g. Harikane et al., 2018; Matthee et al., 2015, 2017c), but such trends require further spectroscopic follow-up of our sample.

Reliable redshift identifications can also be obtained through the dual narrow band technique (e.g. Matthee et al., 2017b; Nakajima et al., 2012; Sobral et al., 2012), where multiple unique combinations of strong emission lines can be observed in specific combinations of narrow or medium band filters[1]. Within the SC4K sample of LAEs, we have already identified 27 Lyα-Ciii] emitters at $z = 2.7 - 3.3$, one Lyα-Civ emitter at $z = 4.3$ (an X-Ray AGN) and 22 Lyα-[Oiii] emitters at $z = 3.3$ (three of these [Oiii] emitters are also Ciii] emitters). One dual-emitter already had a spectroscopic redshift. Hence, we obtain 51 additional reliable redshifts confirming all these sources as LAEs.

---

[1] Here we use line-emitters identified in NB711, NB816 (Santos et al., 2016), NB921 (Matthee et al., 2015; Sobral et al., 2013), NBJ, NBH and NBK (Khostovan et al., 2015; Sobral et al., 2013) to search for another line, in addition to Lyα detected in our MBs.

We also note that some of the contaminants are not easy to isolate by using broad band colours. For example, SC4K-IA767-43371, with a redshift of $z = 5.441$, is selected as a LAE candidate in both IA767 and IA827. While this source is a confirmed LAE (in IA767), the emission line in IA827 is Nv (1240 Å). As such, we remove this source from being a IA827 LAE. There are a further 19 LAE candidates which are lower redshift interlopers and thus are removed from the final sample, either due to archival redshifts or from follow-up with AF2/WHT (see Section 8.2 about future work, where more information about the on-going WHT/AF2 data reduction and analysis is given). We find that the confirmed interlopers/contaminants have a diverse nature. At lower redshift most are Ciii] and Civ (Sobral et al., 2017; Stroe et al., 2017a), while at higher redshift there is a mix of [Oiii]+Hβ and [Oii]. We stress that neither of these class of sources could easily be removed by adjusting our selection criteria and certainly not without compromising our completeness, which we currently estimate to be at the level of $\sim 85 - 90\,\%$. After removing the 19 spectroscopically confirmed interlopers, our final sample of medium band selected sources contains 3434 candidate LAEs.

2.3.6   UV continuum properties of SC4K LAEs

In the process of selecting LAEs we find sources which have no continuum counterpart in the COSMOS data. These are typically found in very deep narrow band or IFU studies (e.g. Ouchi et al., 2008; Oyarzún et al., 2017; Wisotzki et al., 2016), but here we also find them in shallower data. In our samples, $\approx 10\%$ of LAEs have no continuum detection around the narrow or medium band. These LAEs likely occupy the lower stellar mass range of our sample and may have higher escape fractions due to their very high EWs (see e.g. Sobral et al., 2017; Verhamme et al., 2017). We note that due to the fixed broad band depths, the fraction of candidate LAEs without rest-frame UV detections becomes larger with redshift, from just $\sim 1 - 2\%$ at $z \sim 2.5$ to $\sim 10\,\%$ at $z \sim 5$ and reaching 30% for our highest redshift sources. For sources without a rest-frame UV detection, we assume that the continuum flux is an upper limit based on the measured $\mathrm{rms_{BB}}$ and derive lower limits for their EWs. We note that by stacking our LAEs in the rest-frame UV (F814W, *HST*), Paulino-Afonso et al. (2018) find
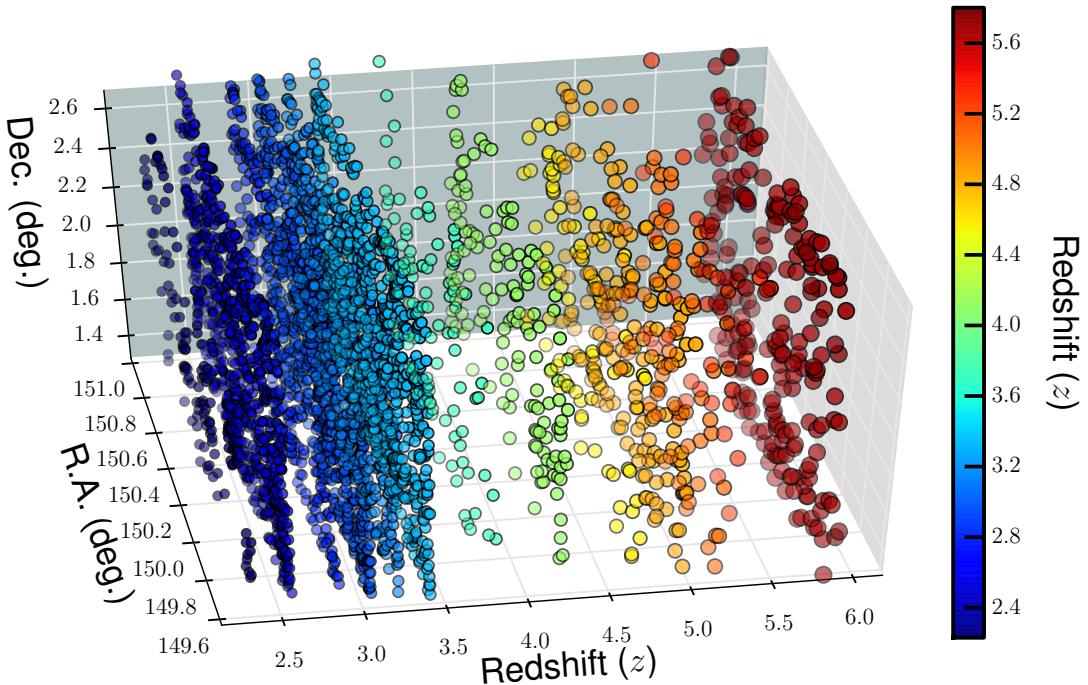
Figure 2.4: The 3D distribution of the SC4K sample presented in this work in the full $2\,\mathrm{deg}^2$ COSMOS field (see Table A.1), showing all LAEs from the 16 different redshift slices, colour coded by redshift (blue to red from lower to higher redshift). The redshift is computed using the central wavelength of the medium or narrow band filter. SC4K probes roughly 4,000 LAEs selected over a total volume close to $\sim 10^8\,\mathrm{Mpc}^3$ (see Table 3.1 for volumes probed per filter).

they have a typical rest-frame UV luminosity of $M_{UV} \sim -20$, which ranges from $M_{UV} = -19.2 \pm 0.2$ for our lowest redshift sample (the deepest in Lyα) to up to $M_{UV} \sim -21$ at higher redshift (see Paulino-Afonso et al., 2018).

2.3.7 Final sample: SC4K

Our sample of medium band selected LAEs consists of 3434 sources (see Table 2.3), visually inspected for spurious detections. We complement our medium band LAEs with four narrow band studies (Table 2.3) in the COSMOS field which follow the same methodology as in this work. We add 159 LAEs at $z \sim 2.23$ (CALYMHA survey; Sobral et al., 2017) and 45 sources at $z \sim 3.1$ (Matthee et al., 2017b), selected with narrow bands NB392 and NB501. In addition, we

also include 78 LAEs at $z \sim 4.8$ and 192 LAEs at $z \sim 5.7$ (Santos et al., 2016), selected with the narrow bands NB711 and NB816, respectively. Our final sample of LAEs contains 3908 sources. We name this sample of $\sim 4,000$ (4k) LAEs, obtained by "slicing" the COSMOS field (Figure 2.4), as SC4K. For an example and description of the catalogue, see Table A.1. Our survey is roughly equivalent to a very wide, low resolution ($R \sim 20 - 80$) IFU Lyα survey covering all the way from $z \sim 2.2$ to $z \sim 6$. A 3D view (showing the full COSMOS field and redshift as a depth dimension) of SC4K is shown in Figure 2.4.

2.4 Summary

We have conducted a wide search for LAEs, using 12 medium bands and a compilation of 4 narrow band filters, covering a wide redshift range ($z \sim 2 - 6$) over the full $\sim 2 \, \mathrm{deg}^2$ COSMOS field. We use these data as an extremely wide, low-resolution ($R \sim 20 - 80$) IFU survey to slice through the COSMOS field.

- We identify $\sim 40,000$ potential line-emitters (before visual inspection), with available photometric and spectroscopic redshifts being consistent with populations of Lyα, [OII], [OIII]+Hβ and Hα emitters.

- We construct a new sample of ~ 4000 typical ($\gtrsim L^\star_{\mathrm{Ly}\alpha}$) LAEs (Slicing COSMOS with 4K LAEs, SC4K) from $z \sim 2$ to $z \sim 6$ in 16 individual redshift slices.

- We make the SC4K sample of LAEs publicly available[1] (see Table A.1) so the community can fully benefit from this work.

This large sample of LAEs, selected with similar selection criteria over a wide area, will allow unprecedented constrains of the evolution of the Lyα LF with redshift, as well as other properties such as stellar mass and star-formation rates.

---

[1]https://academic.oup.com/mnras/article/476/4/4725/4858393#supplementary-data

Chapter 3

The evolution of the Lyα luminosity functions of SC4K LAEs from z $\sim$ 2 to z $\sim$ 6

Abstract

We present new Lyα luminosity functions (LFs) covering a co-moving volume of $\sim 10^8$ Mpc^3. SC4K extensively complements ultra-deep surveys, jointly covering over 4 dex in Lyα luminosity, with the measurements of all surveys being combined into a global ($2.5 < z < 6$) synergy LF. This synergy LF is best fit by $\alpha = -1.93^{+0.12}_{-0.12}$, $\log_{10} \Phi^*_{\mathrm{Ly}\alpha} = -3.45^{+0.22}_{-0.29}$ Mpc^{-3} and $\log_{10} \mathrm{L}^*_{\mathrm{Ly}\alpha} = 42.93^{+0.15}_{-0.11}$ erg s^{-1}. The Schechter component of the Lyα LF reveals a factor ~ 5 rise in $\mathrm{L}^*_{\mathrm{Ly}\alpha}$ and a $\sim 7\times$ decline in $\Phi^*_{\mathrm{Ly}\alpha}$ from $z \sim 2$ to $z \sim 6$. The data reveal an extra power-law (or Schechter) component above $\mathrm{L}_{\mathrm{Ly}\alpha} \approx 10^{43.3}$ erg s^{-1} at $z \sim 2.2 - 3.5$ and we show that it is partially driven by X-ray and radio AGN, as their Lyα LF resembles the excess. The power-law component vanishes and/or is below our detection limits above $z > 3.5$, likely linked with the evolution of the AGN population.

We structure this Chapter as follows: we present the methods in Section 3.1, including all the steps and corrections in determining Lyα LFs. Results are presented in Section 3.2, including the evolution of the Lyα LF with redshift, comparisons with other surveys, the synergy LF (S-SC4K) and the evolution of the Lyα luminosity density. We present our conclusions in Section 3.3.

3.1 Methods and corrections

3.1.1 Lyα luminosities and survey volumes

We compute Lyα luminosities for each of our LAE candidates per filter/redshift slice by using i) their estimated Lyα fluxes in 2$''$ apertures[1] ($F_{\mathrm{Ly}\alpha}$; see e.g. Sobral et al. 2017) and ii) the luminosity distance (D_L) corresponding to Lyα lines detected at the central wavelength of each filter. Luminosity distances (D_L) range from 20×10^3 Mpc at $z \approx 2.5$ to 55×10^3 Mpc at $z \approx 5.8$. Lyα luminosities are then calculated as $\mathrm{L}_{\mathrm{Ly}\alpha} = 4\pi F_{\mathrm{Ly}\alpha} D_L^2$. We find that our "formal" $3\,\sigma$ limit MB detections correspond to Lyα luminosity limits ranging from $10^{42.4}\,\mathrm{erg\,s^{-1}}$ at $z = 2.5$ to $10^{43.0}\,\mathrm{erg\,s^{-1}}$ at $z = 5.8$ (see Table 3.1 for luminosity limits per filter).

Paulino-Afonso et al. (2018) measured the rest-frame UV sizes of our LAEs, concluding they have half-light-radii in the range $\approx 0.1 - 0.2''$ ($\approx 0.7 - 1.3\,\mathrm{kpc}$), and thus significantly smaller than our 2$''$ apertures. However, due to the use of ground-based imaging (with a larger PSF) and the fact that we are tracing Lyα and not the rest-frame UV, the 2$''$ apertures may be missing some flux. We thus study how the fluxes computed in 2$''$ apertures compare with fluxes derived from using an estimate of the full flux using e.g. MAG-AUTO. We find an average ratio (Flux$_{[\mathrm{mag-auto}]}$/Flux$_{[2'']}$) of $\approx 1.03 \pm 0.26$ (median of 1.02). There is no systematic difference in our sample as a whole nor any significant trend with redshift. Therefore, in this study we do not apply any aperture correction and base our measurements on our directly measured 2$''$ aperture quantities (see discussion in Drake et al., 2017b).

---

[1]except for NB392 and NB501 where photometry has been done with 3$''$ apertures (see Matthee et al., 2017b; Sobral et al., 2017) due to broader PSF.

Table 3.1: The Lyα survey co-moving volumes per redshift/filter slice assuming top-hat filter profiles for medium and narrow band filters. We provide the filter name and the Lyα volume corresponding to the 50% transmission points in the normalised filter profile. The two final columns on the right present the limiting luminosity limit ($\log_{10} L_{Ly\alpha}/\mathrm{erg\,s^{-1}}$) for each slice, by using the formal flux limits from Table 2.1 and the 30% completeness limit that we measure with our methodology (see Section 3.1.2.1).

Filter	Lyα z	Area (deg^2)	Volume ($10^6\,\mathrm{Mpc}^3$)	$L_{Ly\alpha,3\sigma}\,limit$ ($\log_{10}$)	$L_{Ly\alpha}$ 30% limit ($\log_{10}$)
IA427	$2.42 - 2.59$	1.94	4.0	42.4	42.5
IA464	$2.72 - 2.90$	1.94	4.2	42.5	42.9
IA484	$2.89 - 3.08$	1.94	4.3	42.5	42.7
IA505	$3.07 - 3.26$	1.94	4.3	42.6	42.7
IA527	$3.23 - 3.43$	1.94	4.5	42.5	42.7
IA574	$3.63 - 3.85$	1.96	4.9	42.7	43.0
IA624	$4.00 - 4.25$	1.96	5.2	42.8	42.9
IA679	$4.44 - 4.72$	1.96	5.5	43.0	43.1
IA709	$4.69 - 4.95$	1.96	5.1	42.9	43.1
IA738	$4.92 - 5.19$	1.96	5.1	43.0	43.3
IA767	$5.17 - 5.47$	1.96	5.5	43.0	43.4
IA827	$5.64 - 5.92$	1.96	4.9	43.0	43.4
NB392	$2.20 - 2.24$	1.21	0.6	42.3	42.3
NB501	$3.08 - 3.16$	0.85	0.9	42.9	43.0
NB711	$4.83 - 4.89$	1.96	1.2	42.6	42.9
NB816	$5.65 - 5.75$	1.96	1.8	42.5	42.5
Total	$2.20 - 5.92$	1.96	61.5	$42.4 - 43$	$42.5 - 43.4$

We compute the co-moving volumes probed by each of the medium bands by approximating them to top-hat filters (using the measured FWHM; Table 2.1). We find co-moving volumes within $(4.0 - 5.5) \times 10^6 \, \mathrm{Mpc}^3$ per medium band and a total co-moving volume of $5.7 \times 10^7 \, \mathrm{Mpc}^3$ over all 12 medium bands; see Table 3.1. The sum of all narrow band volumes contributes with a modest volume of $4.5 \times 10^6 \, \mathrm{Mpc}^3$, but allows us to probe fainter Lyα luminosities (see Table 3.1). The full Lyα survey volume in SC4K is therefore dominated by the medium band filter survey and amounts to $6.2 \times 10^7 \, \mathrm{Mpc}^3$. We note that while our survey is only sensitive to the more typical and bright Lyα emitters, it provides a unique opportunity to explore the bright end of the Lyα luminosity function mostly for the first time, being fully complementary to other previous surveys. For example, we probe a volume $\approx 50{,}000$ times larger than MUSE (Drake et al., 2017a) and $\approx 50 - 60$ times the volumes of typical $1 \, \mathrm{deg}^2$ narrow band surveys (Ouchi et al., 2008) and still a factor of a $\sim 2 - 3$ larger than current $\sim 10 - 20 \, \mathrm{deg}^2$ surveys with Hyper-Suprimecam (e.g. Konno et al., 2018).

3.1.2 Corrections to the Lyα luminosity function

3.1.2.1 Completeness correction

Sources with weak emission lines or with low EWs may be missed by our selection criteria, causing the measured number density of sources to be underestimated. To estimate the line flux completeness we follow Sobral et al. (2013), adapted for Lyα studies by Matthee et al. (2015). Briefly, for each medium band we obtain a sample of non-line-emitters at the redshift we intend to study from the appropriate MB catalogue. To do so, we use the sources which are not classified as line-emitters (we exclude the line-emitters) and, from these, we select sources which are consistent with being at a redshift ± 0.2 of the Lyα redshift for a given filter. We do this by i) applying the same Lyman break selection as we did for the sample of line-emitters and ii) by selecting sources with photometric redshifts within ± 0.2 (Laigle et al., 2016) of the redshift window shown in Table 3.1. We check that our method leads to a distribution in MB magnitudes of non-emitters in agreement with that of the LAEs, with a tail of $\sim 2 - 5\%$ brighter

sources. Overall, our empirical approach leads to a sample of non-line emitters that is slightly brighter than that of LAEs, and thus can be seen as a conservative approach in estimating completeness corrections that does not require making any assumptions to create fake/mock sources.

Our procedure results in samples of non-line-emitters per MB filter that are at roughly the same redshift as our LAEs and allow us to estimate our line-flux completeness with an empirical/data approach. To do so, we add emission line flux to sources in steps of $10^{-18}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$, which results in increasing the flux of the medium and broad bands depending on the filter's FWHM. For each step in flux added, we apply our emission-line selection criteria and identify those that, with the flux added, now make it into a new sample of line emitters and compare those with the total sample that was flux-boosted. By determining the fraction that we retrieve (after applying our Σ and EW cuts; see Section 2.3.1) as a function of added line-flux in comparison with the full sample, we obtain a completeness estimation for each flux, which we apply to our luminosity functions. We only calculate the Lyα luminosity function for luminosity bins in which we find a completeness of 30% or higher at the lowest luminosity limit of the bin; these are in the range $L_{\mathrm{Ly}\alpha} = 10^{42.5-43.4}\,\mathrm{erg\,s^{-1}}$ (see Table 3.1). Our lowest luminosity bin is the one affected by the largest incompleteness and thus the one with the highest completeness correction being applied, which is typically a factor of ≈ 2. We find that the completeness functions strongly depend on line flux, with an increase in completeness from 30% to 90% typically corresponding to a $\approx 0.4 - 0.5\,\mathrm{dex}$ increase in Lyα luminosity, and reaching $\approx 100\%$ with a further $\sim 0.5\,\mathrm{dex}$ increase.

3.1.2.2 Filter profile effects and corrections

As discussed in detail in e.g. Sobral et al. (2013) and Matthee et al. (2015), due to the non-top-hat shape of narrow band filters, sources can be observed at a low transmission (almost no source is observed at full transmission when a filter is well described by a Gaussian function), particularly once survey volumes are large. As a result, assuming a top-hat filter will cause a complex underestimation
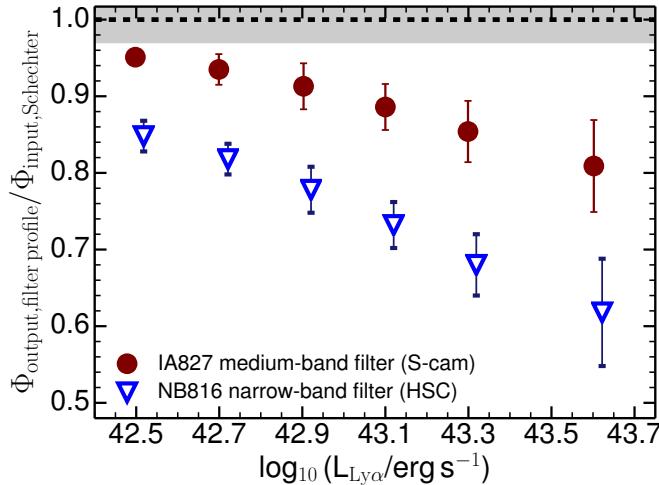
Figure 3.1: The observed ratio between 10,000 observed Lyα luminosity functions using the real filter profiles and a Schechter input simulated sample of LAEs assuming tophat filters. Our results highlight the need to correct for filter profile effects which pushes sources from intrinsically bright to observed fainter bins, and highlights that the corrections are particularly important for narrow band surveys, but are still relevant for medium band surveys.

of the flux, which is manifested in the luminosity function as a transfer of intrinsically bright sources towards observed fainter sources. For an intrinsic Schechter distribution, and particularly for the exponential regime (bright end), this effect results in an underestimation of the number density of the brightest emitters (as they can only be detected as bright over a small redshift range corresponding to the filter's peak transmission), and sometimes an overestimation of the faintest sources (as brighter sources detected away from peak transmission will look fainter). However, the necessary corrections depend on i) the filter profile, ii) the intrinsic shape of the luminosity function and iii) the depth and survey volumes.

While medium bands are broader than narrow bands and in general better fitted by a top-hat, a full investigation of the role of the filter profiles is still required. We estimate potential corrections for each filter by simulating ten million sources with an input random redshift distribution[1] which is wide enough to cover down to zero transmission by each filter on the blue and red wings. We

---

[1]Note that the output distribution is not random and follows closely the filter profile; this is what is used to study the effect of the filter profile.

generate these ten million sources with a luminosity distribution given by the observed (completeness corrected) luminosity function, following Sobral et al. (2013). By convolving the full population with i) the real filter profile and ii) the top-hat approximation we can determine the number density ratio between i) and ii) per luminosity and derive corrections based on the filter profile; an example for IA827 and the NB816 (from HSC) filters is shown in Figure 3.1 (see also Figure A.2).

Our results show that the use of medium band filters results in smaller corrections (see also Appendix A.3.1) than those derived for typical narrow band filters (Figure 3.1). This is because fluxes are only significantly underestimated at the wings of the medium band filters, which correspond to a much smaller fractional volume than for narrow band filters. We also note that the input shape of the luminosity function is crucial for the estimated filter profile effect: while an observed Schechter function leads to a large correction in the exponential part, a bright end which is observationally described by a much slower decline with luminosity (e.g. a power-law with a shallow slope) results in smaller corrections (see full discussion in Appendix A.3.1). Our results thus mean that while for previous deep surveys mostly tracing the faint-end power-law component of the Schechter function the corrections could be relatively small, for the bright end (under a Schechter assumption) the corrections can be large, close to a factor of 2-3 for narrow band filters at the highest luminosities, while they can be a factor of 1.2-1.3 for medium bands (see Figure A.2).

While the filter profile effects can be small for medium bands, we still take them into account by applying a statistical correction to each luminosity bin. This produces our final luminosity function (LF). We provide a more detailed analysis and discussion of the effects, assumptions and corrections due to the various filter profiles when contrasted to top-hat approximations in Appendix A.3.1. We also note that indirect statistical tests for our corrections can be obtained when comparing our results with e.g. MUSE (Drake et al., 2017b) and other IFU surveys which are not affected by filter profile effects (see Section 3.2.1).

3.1.3 Flux robustness and errors: random and systematic

Due to errors in the photometry, both in the MB and BB magnitudes, estimated Lyα fluxes will be subject to errors and, in some cases, also prone to potential systematic effects. We study these errors and their potential role in the derivation of the Lyα LF. Briefly, we assume that each MB and BB magnitudes and their uncertainties are described by normal distributions centred on the measured value and with the width that is equal to its associated $1\,\sigma$ error. We then perturb each galaxy magnitude 10,000 times by randomly picking values from their individual probability distributions. We use these perturbed magnitudes to compute the Lyα LF. We do not find any systematic difference, showing that the errors on the MB and BB photometry have no systematic effect in our methodology. We use the difference between the median value and the 16th and 84th percentiles of the perturbed number density distribution as the lower and upper errors on the number density for each luminosity bin. We find small variations due to this effect, with a median error of ≈ 0.03 error in $\log_{10}(\Phi)$. This is particularly sub-dominant when compared to other sources of uncertainty, but we still add it (0.03 dex) at the end in quadrature (see Figure 3.2).

3.1.4 Completeness-contamination errors in the LAE selection and final errors

While the flux and EW selection/limits can be taken into account for corrections and accounted for in errors (see Section 3.1.2.1), there are other sources of uncertainties that are linked with the photometric or colour-colour criteria applied to select LAEs/filter lower redshift sources (Section 2.3.3). While no single cut is perfect (even more so due to photometric errors), it is possible to perturb the selection and conduct a Markov chain Monte Carlo (MCMC) analysis in order to estimate the effects of varying the selection in the derivation of the Lyα LF and propagated quantities. Here we implement such an analysis. We perturb the LAE selection criteria described in Table 2.3 in a i) ±0.2 dex interval around each colour-colour and photometric cut, independently and ii) by randomly varying by $^{+0.31}_{-0.44}$ the $3\,\sigma$ magnitudes (corresponding to varying non-detection limits in
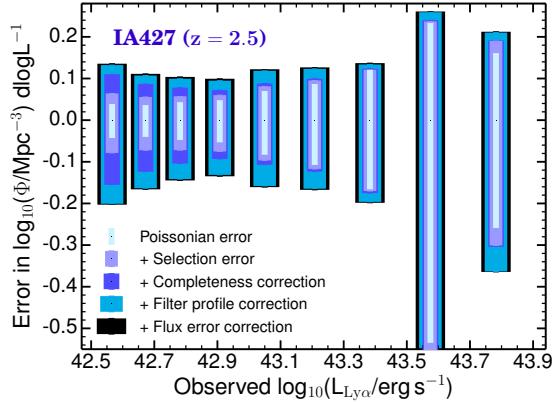
Figure 3.2: An example of the different error contributions to the Lyα luminosity bins from different sources of uncertainty which are taken in series in our analysis. We show bins with bin widths that increase with increasing Lyα luminosity. These include: Poissonian, perturbations to the selection criteria (from line emitters to LAEs; see Table 2.3), line flux completeness, filter profile corrections and flux errors. We find that selection perturbation errors are most important at the faintest luminosities, but they still contribute to the brightest bins. The Poissonian errors are relatively very large at the highest luminosities, where the number of sources per bin is the lowest.

the range $2 - 4\,\sigma$, from the least to the most conservative cuts) of bands tracing bluer than the Lyman-limit, used to reject interlopers. We run a MCMC simulation, with 10,000 iterations for each filter, randomly picking sets of values inside the full explored range, assuming all have an equal probability (flat prior). We then calculate the selection criteria errors as the difference between the median value and the 16th and 84th percentiles within each luminosity bin for all realisations.

An example of the estimation of the full errors affecting $\log_{10}(\Phi)$ can be found in Figure 3.2 for $z = 2.5$ (IA427). We find that the perturbations result in standard deviations of 0.03 to 0.1 dex per luminosity bin at $z \sim 2.5$, representing up to 50% of the total error. The perturbation error is larger in absolute terms at the brightest bins, but it becomes a much more significant fractional contribution to the faintest bins where the Poissonian errors are very small (see Figure 3.2). The errors from perturbing the selection criteria are roughly a factor of up to 2.5 the Poissonian error per bin at the bins probing the faintest luminosities (with the largest number of sources), while they can be as low as 0.2-0.8 of the Poissonian

errors if the bin is populated with only 5-10 sources (where the Poissonian error is already large). We also find that perturbation errors are more important (larger) at $z \sim 2.5 - 2.8$ and $z > 4$ than they are at $z \sim 3 - 3.3$. This roughly coincides with jumps in the selection criteria and whenever different colours/bands are used (see Table 2.3).

We add our estimated perturbation errors in quadrature to the Poissonian errors, noting that they are particularly important for the faint end where the Poissonian errors are an underestimation of the full uncertainties. We then scale the errors by the line flux completeness correction and the filter profile correction, which we assume we know with 30% accuracy (and thus add an extra 30% of such corrections to the errors, taking a conservative approach). We note that we do not add any errors due to cosmic variance, but that given the very large volumes and the multiple redshift slices, we expect these to be just a small fraction of our full errors that are much larger than the formal Poissonian errors. Finally, even though our samples are expected to be contaminated by interlopers at the 10-15% level, similarly to other narrow band surveys, our LAE selection-completeness implies we may be missing 10-15% of real LAEs (when we transform the sample of line emitters into candidate LAEs), and thus in our analysis we do not apply any corrections for this contamination or completeness, as they should roughly cancel out.

3.1.5   Redshift binning

Our multiple redshift slices allow to trace LAEs across well defined cosmic times from $z \sim 2$ to $z \sim 6$. We can also combine the slices to produce a global Lyα LF or obtain slightly broader redshift slices which are populated by a much larger number of sources, and that overcome even more cosmic variance. We bin all our $z \sim 2.5 - 6$ slices (IA427 through to IA827) in order to produce a global high redshift LF and compare it with similar measurement made with the MUSE instrument (e.g Drake et al., 2017b) or with slit observations (e.g. Cassata et al., 2011). We also split the sample into 5 different redshift bins in the following way:

- $z \sim 2.2$ ($z = 2.22 \pm 0.02$; NB392)

- $z \sim 2.5$ ($z = 2.5 \pm 0.1$; IA427)

- $z \sim 3.1$ ($z = 3.1 \pm 0.4$; IA464, IA484, IA505, IA527)

- $z \sim 3.9$ ($z = 3.9 \pm 0.3$; IA574, IA624)

- $z \sim 4.7$ ($z = 4.7 \pm 0.2$; IA679, IA709)

- $z \sim 5.4$ ($z = 5.4 \pm 0.5$; IA738, IA767, IA827)

When producing redshift binned LFs, we only use the volumes associated with a given medium band if that specific filter provides the necessary depth for a completeness above 30%.

3.1.6 Schechter, power-law and combined fits

The Schechter function (Schechter, 1976) is a widely used parametrization of the LF, defined by three parameters: the power-law slope α, the characteristic number density $\Phi^\star$ and the characteristic luminosity $L^\star$. Observations down to extremely low luminosities are necessary to accurately constrain the power-law slope α (e.g Drake et al., 2017b; Dressler et al., 2015). Our medium bands cover "typical" luminosities and higher, thus not probing much fainter than $L^\star$, and do not allow to measure α on their own. However, several studies have been able to obtain good constraints on α from $z \sim 2$ to $z \sim 6$ (e.g. Drake et al., 2017b; Dressler et al., 2015; Konno et al., 2016; Santos et al., 2016), which has been shown to be very steep (< -1.5) and potentially varying from $\alpha \approx -1.7$ at $z = 2.2$ (Konno et al., 2016; Sobral et al., 2017) to $\alpha \approx -2$ (or even steeper; see Drake et al. 2017b) by $z \sim 6$ (Drake et al., 2017b; Dressler et al., 2015; Santos et al., 2016). We therefore fit Schechter functions by varying α between -1.6 and -2.0, but we also explore fits with α fixed to -1.8 at all redshifts in order to investigate the potential redshift evolution of $L^\star_{\mathrm{Ly}\alpha}$ and $\Phi^*_{\mathrm{Ly}\alpha}$ at fixed α. Finally, we also fit α explicitly by combining our results with ultra-deep observations.

In addition to fitting Schechter functions, we also fit power-laws of the form $\log_{10}\Phi = A \log_{10} L_{\mathrm{Ly}\alpha} + B$ to the full LFs and compare these with Schechter fits. Finally, we also explore combinations of a Schechter for lower luminosities and a

power-law at higher luminosities when a single function yields a very bad fit on its own (see Section 3.2). For all LFs, we follow a MCMC approach for the fits, perturbing each bin independently within its asymmetric Gaussian error probability distribution and re-fitting for 10,000 realisations per LF. We take the median of all the best fits as the most likely combination of parameters and estimate the errors by computing the 16th and 84th percentiles for all 10,000 realisations per LF estimation. We note that due to degeneracies in the parameters, these errors can sometimes exaggerate the errors on individual parameters (i.e., parameters are linked and only allowed to vary within some specific relation and not independently), but they are generally well constrained. For each best-fit we also compute the corresponding $\chi^2_{\rm red}$ and use these to obtain the median $\chi^2_{\rm red}$ and the 16th and 84th percentiles of all realisations.

3.1.7 X-ray and radio properties: AGN candidates within our LAEs

We explore *Chandra* X-ray (e.g. Civano et al., 2016) and VLA radio data (e.g. Smolčić et al., 2017) within COSMOS to identify AGN in our sample. Full details are given in Calhau et al. (2020). Briefly, we use the publicly available *Chandra* Cosmos Legacy survey (Civano et al., 2016; Elvis et al., 2009; Puccetti et al., 2009) to select sources with X-ray counterparts, within the overlap region with SC4K (1.86 deg^2). Out of the full SC4K sample of 3908 LAEs presented in this work, 3707 have *Chandra* X-ray coverage. From those, we identify 109 sources with X-ray emission in the Civano et al. (2016) catalogue, making them strong candidates of being X-ray AGN ($L_X > 10^{42.5}\,{\rm erg\,s^{-1}}$). Calhau et al. (2020) presents a detailed analysis on the X-ray activity of our full sample of LAEs. We find a global X-ray AGN fraction among our SC4K LAEs of $2.9 \pm 0.3\%$ for $z = 2 - 6$. The AGN fraction shows evidence for a decline with increasing redshift for typical to bright LAEs, with this trend not being driven by X-ray luminosity limits (Calhau et al., 2020). At $z \sim 2.2 - 2.7$ the X-ray AGN fraction is $3.9 \pm 0.6\%$, declining to $3.5 \pm 0.4\%$ and $0.4 \pm 0.2\%$ for redshifts $2.7 < z < 3.5$ and $3.5 < z < 6$, respectively. We also identify a clear relation between the X-ray AGN fraction
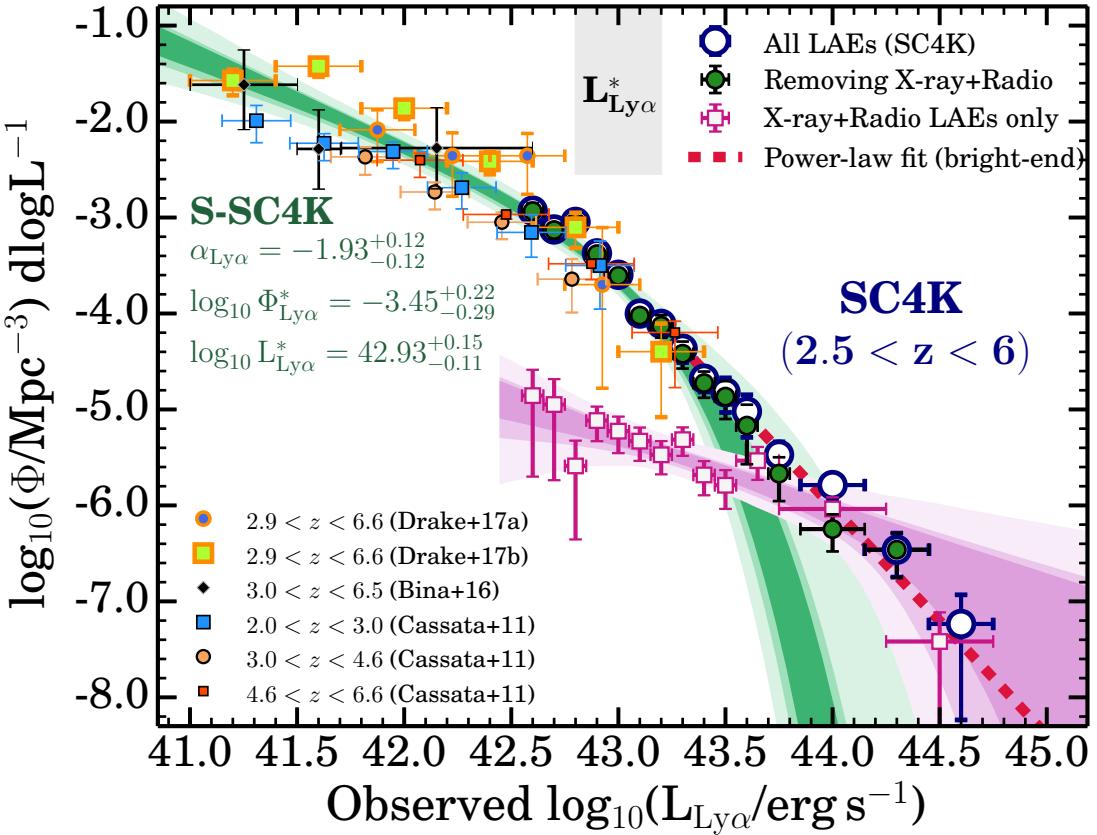
Figure 3.3: The combined global Lyα LF at $2.5 < z < 6$. Our large samples of luminous LAEs, obtained over a co-moving volume of $0.6 \times 10^8 \, \mathrm{Mpc}^3$ are able to constrain the bright end of the Lyα LF for the first time down to number densities of $\sim 10^{-7} \, \mathrm{Mpc}^{-3}$ and Lyα luminosities of $\sim 10^{44.5} \, \mathrm{erg \, s}^{-1}$. We find a significant excess of bright LAEs at the highest luminosities when compared to a single Shechter function and show that is likely driven by a population where Lyα is AGN-driven. We compute a proxy for the AGN Lyα LF with X-ray and radio AGN among our sample (1 and $2\,\sigma$ contours shown for Schechter function fits). We also compare our results with recent MUSE/VLT (Bina et al., 2016; Drake et al., 2017a,b) and VIMOS/VLT observations (e.g. Cassata et al., 2011), showing a very good agreement in the $L^*_{\mathrm{Ly}\alpha}$ range where all studies overlap. Deeper, smaller volume studies from the literature allow to cover the sub-$L^\star_{\mathrm{Ly}\alpha}$ luminosity regime, being perfectly complementary to our approach. Overall, we show the Lyα LF determined over 4 orders of magnitude in Lyα luminosity and 6 orders of magnitude in number densities at $z \sim 2.5 - 6$ resulting in the 'synergy' Lyα LF (S-SC4K; $2.5 < z < 6$) and the 1, 2 and $3\,\sigma$ confidence levels when fitting a Schechter function up to $10^{43.3} \, \mathrm{erg \, s}^{-1}$ (we also show the power-law fit done for higher luminosities). Results are provided in Table 3.2.

and Lyα luminosity of LAEs (see Calhau et al., 2020), qualitatively similar to what has been found at lower redshift (Wold et al., 2014, 2017) and also found and discussed in Ouchi et al. (2008) and Matthee et al. (2017b).

In the VLA radio data (1.4 GHz and 3 GHz; see Schinnerer et al., 2010; Smolčić et al., 2017) we identify 62 individual sources, with these being dominated by the 3 GHz detections (61; 25 detected in 1.4 GHz), and we class these as AGN. Out of these, 30/62 are also X-ray sources. We therefore find a total of 141 AGN sources among the SC4K sample of LAEs, yielding a total AGN fraction of $3.6 \pm 0.3\%$. If we split the sample in three redshift ranges, we find that the AGN fraction slowly declines from $z \sim 2.2 - 2.7$ ($4.7 \pm 0.7\%$) to $z \sim 2.7 - 3.5$ ($4.4 \pm 0.4\%$) and then drops significantly at $z \sim 3.5 - 6$ ($1.2 \pm 0.4\%$). Concentrating on radio AGN within our sample of LAEs, we find that the (radio) AGN fraction is relatively constant ($1.9 \pm 0.4\%$) at $z \sim 2.2 - 3.5$ and then drops to $0.9 \pm 0.2\%$ at $z \sim 3.5 - 6$.

3.2 Results

3.2.1 The global Lyα LF at z $\sim$ 2.5 $-$ 6

In Figure 3.3 we present the global Lyα LF at $z \sim 2.5 - 6$, which we define as the LF of our entire SC4K sample of 3908 LAEs, determined over a total volume of close to $\sim 10^8\,\mathrm{Mpc}^3$ (not an average of individual LFs at multiple redshifts). Our results probe Lyα luminosities from $\sim 10^{42.5}\,\mathrm{erg\,s}^{-1}$ to $\sim 10^{44.5}\,\mathrm{erg\,s}^{-1}$, covering 2 orders of magnitude in Lyα luminosity with a single survey. Down to our observational limits, we find that the global Lyα LF at $z \sim 2.5 - 6$ resembles a single or double power-law (or a double Schechter, but not a single Schechter function) and extends to luminosities and number densities that reach into what is expected from the quasar luminosity function (e.g. Richards et al., 2006) and follow-up of quasars in Lyα (e.g. Borisova et al., 2016). Fitting the global SC4K Lyα LF leads to a power-law of $\log_{10}(\Phi) = -2.22^{+0.08}_{-0.10} \log_{10}(\mathrm{L_{Ly\alpha}}) + 91.7^{+4.1}_{-3.6}$ (see Table 3.2), which describes the data significantly better than a single Schechter function ($\chi^2_\mathrm{Sch}/\chi^2_\mathrm{PL} \approx 8$; see Table 3.2). If we exclude X-ray AGN and radio AGN, we find that the global Lyα LF becomes steeper at the bright end. We can also derive a

Table 3.2: The global Lyα LF at $z \sim 2.5 - 6$ from SC4K only, when combined with the latest MUSE results (Drake et al., 2017b) and when using the derived consensus global Lyα LF, S-SC4K (SC4K and Bina et al., 2016; Cassata et al., 2011; Drake et al., 2017a,b, see Section 3.2.4). We also show the results when explicitly removing radio and X-ray AGN from the sample (see Section 3.1.7). The corresponding $\rho_{\mathrm{Ly}\alpha}$ have been computed by integrating Schechter functions down to $1.75 \times 10^{41} \, \mathrm{erg \, s^{-1}}$, corresponding to $0.04 \, L^{\star}_{z=3}$ from Gronwall et al. (2007); see Section 4.1. All errors are the 16th and 84th percentiles for all 10,000 realisations per LF estimation which, due to degeneracies in the parameters, can sometimes exaggerate the errors on individual parameters. We also provide a comparison (ratio) between reduced χ^2 for Schechter and power-law fits ($\chi^2_{\mathrm{Sch}}/\chi^2_{\mathrm{PL}}$) fitted over the same luminosity range for a fair comparison; values below 1 indicate that a Schechter fit performs significantly better, while a large value indicates that a simple power-law fit provides a relatively lower reduced χ^2. * Note that fits to the full LF are given for completeness and comparison, but that they fail to fit the data as a whole, as the combined faint and bright ends are not accurately described by a single Schechter or power-law functions.

Global Lyα sample (2.5 < z < 5.8)	α	$\log_{10} \mathrm{L}^*_{\mathrm{Ly}\alpha}$ ($\mathrm{erg\,s^{-1}}$)	$\log_{10} \Phi^*_{\mathrm{Ly}\alpha}$ ($\mathrm{Mpc^{-3}}$)	$\rho_{\mathrm{Ly}\alpha}/10^{40}$ Sch ($\mathrm{erg\,s^{-1}\,Mpc^{-3}}$)	Power-law (PL) (A $\log_{10}$ L+B)	$\chi^2_{\mathrm{Sch}}/\chi^2_{\mathrm{PL}}$
SC4K ($\log_{10} \mathrm{L_{Ly\alpha}} < 43.3$)	-1.8 ± 0.2 (fix)	$42.81^{+0.07}_{-0.06}$	$-3.16^{+0.13}_{-0.14}$	$0.98^{+0.22}_{-0.17}$	$-2.09^{+0.17}_{-0.17}, 86.1^{+7.3}_{-7.1}$	0.6
SC4K+MUSE ($\log_{10} \mathrm{L_{Ly\alpha}} < 43.3$)	$-1.80^{+0.11}_{-0.11}$	$42.72^{+0.07}_{-0.06}$	$-2.92^{+0.14}_{-0.16}$	$1.32^{+0.12}_{-0.12}$	$-1.36^{+0.05}_{-0.05}, 55.1^{+2.2}_{-2.4}$	0.4
S-SC4K ($\log_{10} \mathrm{L_{Ly\alpha}} < 43.3$)	$-1.93^{+0.12}_{-0.12}$	$42.93^{+0.15}_{-0.11}$	$-3.45^{+0.22}_{-0.29}$	$0.88^{+0.09}_{-0.09}$	$-1.29^{+0.06}_{-0.06}, 52.0^{+2.6}_{-2.7}$	0.8
SC4K* (All LAEs)	-1.8 ± 0.2 (fix)	$43.59^{+0.06}_{-0.06}$	$-4.53^{+0.13}_{-0.16}$	$0.33^{+0.07}_{-0.05}$	$-2.22^{+0.08}_{-0.10}, 91.7^{+4.1}_{-3.6}$	8.0
SC4K+MUSE* (All LAEs)	$-2.55^{+0.06}_{-0.06}$	$43.92^{+0.12}_{-0.11}$	$-5.47^{+0.24}_{-0.26}$	$1.40^{+0.17}_{-0.15}$	$-1.78^{+0.04}_{-0.05}, 72.7^{+2.0}_{-1.9}$	0.7
S-SC4K* (All LAEs)	$-2.45^{+0.06}_{-0.06}$	$43.87^{+0.10}_{-0.10}$	$-5.32^{+0.21}_{-0.23}$	$1.04^{+0.12}_{-0.12}$	$-1.69^{+0.05}_{-0.05}, 68.6^{+2.0}_{-2.0}$	0.7
X-ray + radio AGN only	$-1.7^{+0.3}_{-0.2}$	$51.3^{+1.2}_{-7.3}$	$-11.0^{+5.0}_{-2.6}$	$0.027^{+0.013}_{-0.013}$	$-0.75^{+0.17}_{-0.17}, 27.1^{+7.3}_{-7.2}$	1.3
SC4K* (w/o X-ray+radio)	-1.8 ± 0.2 (fix)	$43.56^{+0.06}_{-0.05}$	$-4.56^{+0.12}_{-0.14}$	$0.29^{+0.06}_{-0.05}$	$-2.38^{+0.09}_{-0.10}, 98.7^{+4.4}_{-4.1}$	8.2
SC4K+MUSE* (w/o X-ray+radio)	$-2.63^{+0.06}_{-0.06}$	$43.90^{+0.12}_{-0.10}$	$-5.59^{+0.25}_{-0.28}$	$1.48^{+0.18}_{-0.17}$	$-1.86^{+0.05}_{-0.05}, 76.2^{+2.2}_{-2.1}$	0.7
S-SC4K* (w/o X-ray+radio)	$-2.52^{+0.07}_{-0.07}$	$43.84^{+0.11}_{-0.09}$	$-5.40^{+0.21}_{-0.25}$	$1.09^{+0.14}_{-0.13}$	$-1.77^{+0.05}_{-0.05}, 72.0^{+2.2}_{-2.1}$	0.7

X-ray+radio AGN Lyα LF, which we also present in Figure 3.3, together with the range of Schechter fits encompassing 1 and 2σ ranges of all realisations. We find evidence for the AGN population being responsible for the 'bump' at high Lyα luminosities, which can be parameterised by a Schechter function[1] with a higher characteristic luminosity, a lower characteristic number density and potentially a shallower slope than the global population (see Table 3.2). The full cross-over between the likely two populations happens at $\approx 10^{43.5}\,\mathrm{erg\,s^{-1}}$, although given that X-ray and radio only provide a partial view of all the AGN, this transition may happen at slightly lower luminosities $\approx 10^{43.2-43.3}\,\mathrm{erg\,s^{-1}}$ (see e.g. Sobral et al., 2019). It is worth noting that the typical characteristic number density of the AGN component of the Lyα LF is close to $\approx 10^{-6}\,\mathrm{Mpc^{-3}}$, similar to the number densities of clusters in the Universe, and that may provide a natural link between bright LAEs at $z > 2.5$ (typically seen as very extended and thus called Lyα 'blobs'[2]) and the physical environments they inhabit (potential 'proto-clusters').

In Figure 3.3 we also show results obtained with much deeper surveys, including MUSE (Bina et al., 2016; Drake et al., 2017a,b) and results from slit spectroscopy using VIMOS/VLT (Cassata et al., 2011); see also Table A.4. We find excellent agreement within the error bars with the MUSE results presented by Drake et al. (2017a,b), although we note that the agreement is only possible to be tested around $\mathrm{L^*_{Ly\alpha}}$, where all studies overlap. Future results from the MUSE-wide project (see Caruana et al., 2018; Herenz et al., 2017), or a compilation of extra-galactic MUSE archival observations, may be able to extend the volume covered by MUSE and further increase the overlap, allowing for more detailed comparisons and to evaluate any systematics/differences. Extremely deep MUSE data allow to not only blindly find faint LAEs, but even more importantly to measure the full Lyα luminosity of each source without effects from narrow band filter profiles (see Drake et al., 2017b; Leclercq et al., 2017). The comparison thus provides statistical evidence that our corrections are able to recover the full Lyα LF.

---

[1]It can also be relatively well parameterised by a simpler power-law function, see Table 3.2.

[2]Morphological analysis of the SC4K sample reveals that SC4K LAEs are typically very compact in the UV but more extended in Lyα emission (Paulino-Afonso et al., 2018, see also Section 1.5.3).

Comparing our results with Cassata et al. (2011), we find a very good agreement with their $z \sim 2-3$ and $z \sim 4.6-6.6$ Lyα LFs. The $z \sim 3.0-4.6$ LF from Cassata et al. (2011) is slightly below ours in the small luminosity range overlap (we can only use one of their bins to directly compare with ours), but we note that their results are also below those from MUSE (see Drake et al., 2017a,b). Apart from cosmic variance and the large differences in selection (our selection is directly on Lyα, more similar to MUSE), there could also be some cosmic evolution. Furthermore, we note that the use of slits and potential underestimation of slit corrections may further explain the differences. Both narrow band surveys and MUSE have established that Lyα emission is significantly extended (e.g. Leclercq et al., 2017; Momose et al., 2014; Sobral et al., 2017; Wisotzki et al., 2016), thus making slit spectroscopy hard to correct. Slit corrections can be particularly challenging as they are often based on the UV continuum, but there is no simple relation between the Lyα extent and the UV extent (see e.g. Leclercq et al., 2017).

Lyα surveys from deeper (necessarily smaller) volumes are needed to cover the sub-L$^{\star}_{\mathrm{Ly}\alpha}$ luminosity regime (Bina et al., 2016; Drake et al., 2017a,b), as highlighted in Figure 3.3. Overall, we can now determine the Lyα LF over 4 orders of magnitude in Lyα luminosity at $z \sim 2.5-6$. Figure 3.3 also reveals how complementary ultra-deep MUSE and slit observations are to very wide narrow and medium band surveys (e.g. SC4K and Konno et al., 2018), allowing unique synergies and providing the first combined view all the way from the faintest Lyα sources to the brightest. We fully explore the combined strength of deep surveys[1] (to probe the faint end) and SC4K (to probe the bright end) and derive a combined Lyα LF (S-SC4K; see Section 3.2.4) presented in Figure 3.3 and Table 3.2. We obtain two cases: when combining SC4K with the latest MUSE results (Drake et al., 2017b) and when combining all studies with SC4K (Bina et al., 2016; Cassata et al., 2011; Drake et al., 2017a,b). While we note that a single Schechter function is simply not an appropriate fit to the full LF, we still

---

[1]In order to account for potential systematic differences between surveys, cosmic variance and due to the way we compute errors, we add errors of $^{+0.05}_{-0.08}$ to data bins determined with deeper observations/by other studies, as they are able to explain current differences between surveys and methods. We note, nonetheless, that these errors are very uncertain in themselves and depend on which surveys/methods are being compared.

provide those for completeness and for comparison of parameters. Restricting the fit to $L_{Ly\alpha} < 10^{43.3}\,\mathrm{erg\,s^{-1}}$ allows to fit a single Schechter which is likely tracing an overall population of SF-dominated galaxies, showing a steep slope, $\alpha = -1.93^{+0.12}_{-0.12}$ (greatly improved when using MUSE only see Drake et al. 2017b), with $L^*_{Ly\alpha} = 10^{42.93^{+0.15}_{-0.11}}\,\mathrm{erg\,s^{-1}}$ and $\Phi^*_{Ly\alpha} = 10^{-3.45^{+0.22}_{-0.29}}\,\mathrm{Mpc^{-3}}$ (see Table 3.2).

Due to the different Lyα luminosity limits, our global LF presented in Figure 3.3 is inevitably dominated by sources at different redshifts as a function of luminosity, with the lower luminosity bins being dominated by the (deeper) lower redshift data, while at higher luminosities all redshifts contribute roughly equally. This is not a problem in the case of a slow or negligible evolution in the Lyα LF with redshift from $z \sim 2.5$ to $z \sim 6$ (e.g. Ouchi et al., 2008), but this has not been fully established yet, particularly for the bright end (for the evolution of the faint-end, see Drake et al., 2017b). Our large sample of typical to bright LAEs is ideal to investigate whether that is the case and to quantify any potential evolution with redshift.

3.2.2 The evolution of the Lyα luminosity function from $z \sim 2$ to $z \sim 6$ in 12 redshift slices

After presenting the global Lyα LF for our full sample in Section 3.2.1, we now explore the multiple redshift slices in SC4K (see Table A.5). In Figure 3.4 we present the Lyα LF per redshift slice all the way from $z \sim 2.2$ to $z \sim 5.8$ by deriving them per filter/redshift. We find a mild but noticeable evolution of the bright end of the Lyα LF with redshift from $z \sim 2.2$ to $z \sim 6$. This evolution seems to be mostly visible in terms of i) an evolution of the shape and ii) an evolution in luminosity. At lower redshift ($z \sim 2.2-3.3$) there is a significant extra component (in addition to a single Schechter) to the Lyα LF above luminosities of $\approx 10^{43.3}\,\mathrm{erg\,s^{-1}}$, while such a component seems to completely disappear by $z \sim 3.7$ and to not show up in any of the Lyα LFs towards higher redshift. Interestingly, when considering only the major Schechter component of the Lyα LF, we find evidence for $L^*_{Ly\alpha}$ to be evolving with redshift towards $z \sim 6$ (see Table 3.3 and Figure 3.5).
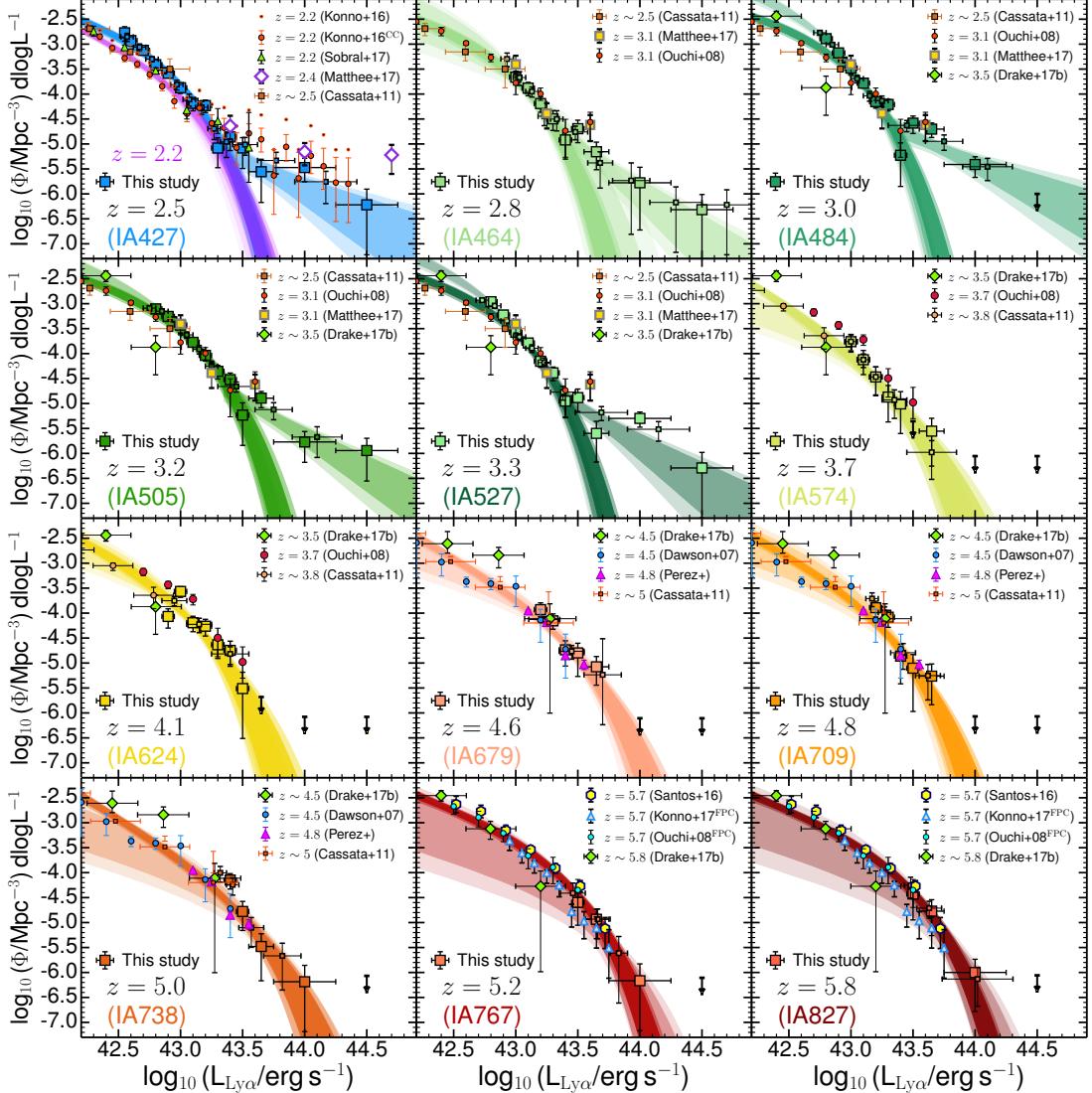
Figure 3.4: The evolution of the (bright end of the) Lyα LF from $z \sim 2.2$ to $z \sim 6$ in 13 (two in the first panel) redshift slices and comparison with a variety of surveys at roughly the same redshift as each slice. Upper limits in the number densities of our LAEs are shown as black arrows. Our results reveal a significant evolution at the bright end, with the number counts falling down as a steepening potential power-law at $z \sim 2.2 - 3.3$ which can be described as a single Schechter function at $z > 3.5$. We show two bin realisations for visualisation of binning effects, but also the much more representative range of Schechter and power-law (percentiles, corresponding to $1, 2\sigma$) fits from perturbing the data. In addition, we also show the fits and uncertainties when exploring synergies with deeper surveys (S-SC4K), which greatly reduces the uncertainties (darker contours). Note that we show both the original Konno et al. (2016) $z = 2.2$ LF in small points, and after correcting for potential contamination (see Sobral et al., 2017). At $z \sim 5.8$ we compare our measurement with NB surveys (e.g. Konno et al., 2018; Ouchi et al., 2008) corrected for filter profile effects as in Santos et al. (2016).

In order to quantify the potential redshift evolution and its significance, we use our best-fits of single power-laws, single Schechter functions and combinations of both and compare the resulting reduced χ^2 ($\chi^2_{\rm red}$, see Table A.5). We find that a single Schechter function is a particularly bad fit when including the bright end of the Lyα LF ($\chi^2_{\rm red} \sim 10 - 30$) from $z = 2.5$ to $z = 3.3$. A single power-law fits better ($\chi^2_{\rm red} \sim 3 - 7$), while a combination of a Schechter at lower luminosities and a power-law at higher luminosities with a transition around $10^{43.3}\,{\rm erg\,s^{-1}}$ provides the best fits (see Figure 3.4). The combined fit are similar to the ones applied in recent large volume Lyα studies at a variety of redshifts (e.g. Konno et al., 2016; Matthee et al., 2017b; Sobral et al., 2017; Wold et al., 2017). Interestingly, the Schechter component of the Lyα LF shows little evolution in $\rm L^*_{Ly\alpha}$ from $z \sim 2.5$ to $z \sim 3.1$, but reveals an important $\Phi^*_{\rm Ly\alpha}$ evolution from $z \sim 2.5$ to $z \sim 3.3$ (see Figure 3.5), which may be consistent with an 'extended' period of peak activity in the Universe (Madau & Dickinson, 2014). For $z > 3.5$, a single Schechter fit provides very good fits, although a single power-law could in principle also describe the bright end of the Lyα LF. From $z \sim 2.2$ to $z \sim 3.3$ the Lyα LF reveals a rise in $\Phi^*_{\rm Ly\alpha}$ by a factor ≈ 4, along with a potential steepening of the power-law component at the bright end of the LF. For $z > 3.5$, where the power-law component is not seen anymore, our results reveal a fall of $\Phi^*_{\rm Ly\alpha}$ and a rise of $\rm L^*_{Ly\alpha}$ up to $z \approx 5.8$ (see also Table A.5).

Using the redshift bins defined in Section 3.1.5 we show the overall redshift evolution of the Lyα LF in Figure 3.5 (see Table 3.3). We also use other/different filter combinations to obtain different redshift bins, and find that the results are all consistent within the error-bars, and thus not dependent on the choice of binning. The increased statistical sample from the redshift bins provides stronger constraints on the Lyα LF, and further reinforces the results already mentioned when looking at each of the individual 12 redshift slices, including the presence of a potential power-law (or extra Schechter) component at high luminosities at $z \sim 2 - 3.5$, which seems to disappear or be at too low number densities for even our survey to detect beyond $z \sim 3.5$. Focusing on the Schechter components (fitting a Schechter only up to $10^{43.3}\,{\rm erg\,s^{-1}}$ at $z < 3.3$ where a clear excess at the bright end is found), and for a fixed $\alpha = -1.8$, we find that $\rm L^*_{Ly\alpha}$ may evolve in a relatively continuous way from $10^{42.69^{+0.05}_{-0.04}}\,{\rm erg\,s^{-1}}$ at $z \sim 3.1$ to $10^{43.35^{+0.12}_{-0.11}}\,{\rm erg\,s^{-1}}$

Table 3.3: The results of fitting different Lyα LFs with a Schechter function at the appropriate luminosity range. For SC4K only we do not fit α, but instead fix it from -1.6 to -2.0 in steps of 0.05 with a uniform prior; the ± 0.2 shown therefore reflects the variation we impose on α, and not an uncertainty in fitting α. For S-SC4K we explicitly fit all three parameters. For each fit we also integrate the Lyα LF to obtain $\rho_{\mathrm{Ly}\alpha}$, derived for different redshift bins, down to $0.04 L^\star$. All errors are the 16 and 84 percentiles of all the fits, derived from our 10,000 realisations per LF. We also convert $\rho_{\mathrm{Ly}\alpha}$ to a star formation rate density by using Equation 6.4 (Kennicutt, 1998) assuming a Salpeter IMF between $0.1 - 100\,\mathrm{M}_\odot$ and affected by $\mathrm{f_{esc}}$ (note that any correction for dust extinction will also be included in the $\mathrm{f_{esc}}$ term).

Redshift bin (SC4K only)	α	$\log_{10} \mathrm{L}^*_{\mathrm{Ly}\alpha}$ (erg s^{-1})	$\log_{10} \Phi^*_{\mathrm{Ly}\alpha}$ (Mpc^{-3})	$\rho_{\mathrm{Ly}\alpha} / 10^{40}$ (erg s^{-1} Mpc^{-3})	$\mathrm{SFRD}_{\mathrm{Ly}\alpha} \times \mathrm{f_{esc}} / 10^{-2}$ (M$_\odot$ yr^{-1} Mpc^{-3})	Reference(s) (Table A.4)
$z = 2.2 \pm 0.1$ (L $< 10^{43.3}$)	-1.8 ± 0.2 (fix)	$42.69^{+0.14}_{-0.11}$	$-3.33^{+0.21}_{-0.26}$	$0.48^{+0.04}_{-0.04}$	$0.44^{+0.04}_{-0.04}$	12
$z = 2.5 \pm 0.1$ (L $< 10^{43.3}$)	-1.8 ± 0.2 (fix)	$42.76^{+0.08}_{-0.07}$	$-3.23^{+0.15}_{-0.15}$	$0.73^{+0.18}_{-0.14}$	$0.67^{+0.16}_{-0.13}$	SC4K
$z = 3.1 \pm 0.3$ (L $< 10^{43.3}$)	-1.8 ± 0.2 (fix)	$42.69^{+0.05}_{-0.04}$	$-2.73^{+0.11}_{-0.12}$	$1.90^{+0.56}_{-0.39}$	$1.73^{+0.51}_{-0.36}$	SC4K
$z = 3.9 \pm 0.2$	-1.8 ± 0.2 (fix)	$42.89^{+0.11}_{-0.10}$	$-3.71^{+0.30}_{-0.28}$	$0.34^{+0.21}_{-0.12}$	$0.31^{+0.19}_{-0.11}$	SC4K
$z = 4.7 \pm 0.1$	-1.8 ± 0.2 (fix)	$43.10^{+0.13}_{-0.12}$	$-3.82^{+0.33}_{-0.32}$	$0.48^{+0.33}_{-0.18}$	$0.43^{+0.30}_{-0.16}$	SC4K
$z = 5.4 \pm 0.4$	-1.8 ± 0.2 (fix)	$43.35^{+0.12}_{-0.11}$	$-4.18^{+0.31}_{-0.30}$	$0.41^{+0.28}_{-0.16}$	$0.37^{+0.26}_{-0.15}$	SC4K
S-SC4K: synergy Lyα LF						
$z = 2.2 \pm 0.1$ (L $< 10^{43.3}$)	$-2.00^{+0.15}_{-0.14}$	$42.82^{+0.13}_{-0.10}$	$-3.59^{+0.22}_{-0.28}$	$0.52^{+0.05}_{-0.05}$	$0.47^{+0.04}_{-0.04}$	2.1, 6.1, 12
$z = 2.5 \pm 0.1$ (L $< 10^{43.3}$)	$-1.72^{+0.15}_{-0.15}$	$42.71^{+0.09}_{-0.08}$	$-3.10^{+0.17}_{-0.21}$	$0.74^{+0.08}_{-0.07}$	$0.67^{+0.07}_{-0.07}$	2.1, 5.1
$z = 3.1 \pm 0.3$ (L $< 10^{43.3}$)	$-1.63^{+0.17}_{-0.16}$	$42.77^{+0.12}_{-0.09}$	$-3.06^{+0.21}_{-0.26}$	$0.86^{+0.10}_{-0.09}$	$0.78^{+0.09}_{-0.08}$	2.1, 5.1
$z = 3.9 \pm 0.2$	$-2.26^{+0.18}_{-0.17}$	$42.93^{+0.13}_{-0.11}$	$-3.66^{+0.30}_{-0.35}$	$1.11^{+0.19}_{-0.16}$	$1.00^{+0.17}_{-0.14}$	2.2, 5.1, 5.2
$z = 4.7 \pm 0.1$	$-2.35^{+0.19}_{-0.19}$	$43.28^{+0.20}_{-0.14}$	$-4.25^{+0.34}_{-0.49}$	$1.16^{+0.40}_{-0.27}$	$1.05^{+0.36}_{-0.25}$	2.3, 3, 5.2, 10
$z = 5.4 \pm 0.4$	$-1.98^{+0.14}_{-0.14}$	$43.28^{+0.09}_{-0.09}$	$-3.83^{+0.21}_{-0.22}$	$1.11^{+0.21}_{-0.17}$	$1.01^{+0.19}_{-0.16}$	5.3, 9.4, 11

at $z \sim 5.4$, which would imply a factor $\sim 4 - 5$ increase in the typical luminosity. This is accompanied by a strong decline of $\Phi^*_{Ly\alpha}$ of $\sim 10 - 30$ times from $z \sim 3.1$ to $z \sim 5.4$ (see Table 3.3 and Figure 3.5).

The apparent decline in $\Phi^*_{Ly\alpha}$, accompanied by a positive $L^*_{Ly\alpha}$ evolution ($L^*_{Ly\alpha}$ increase with an increasing redshift) may be linked to an evolution of the nature of the sources or changes in the conditions of the ISM and CGM, but potentially also with an evolution of the AGN population. We discuss possible explanations in Chapter 4. We note that by excluding X-ray and radio AGN we find a reduction of the number densities of LAEs at the highest luminosities, lowering and steepening the potential power-law component, but without removing it. This means that if the power-law component is fully AGN driven (Sobral et al., 2018a; Wold et al., 2017) there is still a significant component of the AGN population that is simply not detectable in the X-rays or radio (see Section 4.5), potentially because these AGN are very young and/or of very low black hole mass, but highly efficient in the production of Lyα photons which might easily escape, or due to the timescales involved in the AGN turning on and off. Our results thus highlight two potentially important/different physical mechanisms contributing to the Lyα LF at $z \sim 2-6$.

3.2.3 Comparison with other studies at z $\sim 2 - 6$

A wide range of Lyα surveys using narrow bands, slits or IFUs have derived Lyα LFs at $z \sim 2 - 6$, mostly probing at and below $L^*_{Ly\alpha}$ (e.g. Cassata et al., 2011; Dawson et al., 2007; Drake et al., 2017b; Gronwall et al., 2007; Murayama et al., 2007; Ouchi et al., 2008; Rauch et al., 2008; Shimasaku et al., 2006; Shioya et al., 2009; Westra et al., 2006); see Table A.4. These are both perfect comparisons to our results and useful extensions to fainter luminosities.

A comparison between the Lyα LFs from this work and other studies at similar redshifts from the literature is shown in Figure 3.4. We find that the $z = 2.2$ Lyα LF from Sobral et al. (2017) is in good agreement with our $z = 2.5$ measurements at the bright end, but the comparison reveals a positive $\Phi^*_{Ly\alpha}$ evolution from $z \sim 2.2$ to $z \sim 2.5$ (see also Figure 3.5). The $z \sim 2.2$ Lyα LF presented by Konno et al. (2016) is in reasonable agreement with ours, and also implies evolution from $z = 2.2$ to $z \sim 2.5$, but implies higher number densities of bright sources

(see discussion on the importance of filtering out lower redshift interlopers and how they can easily account for 50% of high EW sources in the bright end at $z \sim 2$; see Sobral et al., 2017). We also show the Konno et al. (2016) results when removing likely contaminants in Figure 3.4 as in Sobral et al. (2017), which results in an even better agreement with our results at the bright end. The $z = 2.4$ LF from Matthee et al. (2017b) is also in good agreement with our measurement at $z \sim 2.5$. We note that the number densities observed for the brightest bin in Matthee et al. (2017b) are marginally higher than ours (Figure 3.4), and that those high luminosity sources have now all been spectroscopically confirmed (see Sobral et al., 2018b), and thus contamination is not able to explain the small discrepancy. The observed lower number densities for our results based on the medium bands when compared with Matthee et al. (2017b) may be explained by some of the brightest sources having lower EWs and thus being missed by our relatively high EW cut, even after applying our completeness corrections (see full discussion in Sobral et al., 2017). Cosmic variance is another possibility. We also compare our results to Cassata et al. (2011) and find a good agreement.

The 'mild' increase from $z \sim 2$ up to $z \sim 3.3 - 3.7$ of the number density of LAEs (factor of ≈ 4) across the entire luminosity range is consistent with measurements from several studies, where a similar rise of the Schechter function is seen by comparing e.g. Sobral et al. (2017) at $z \sim 2.2$ with Ouchi et al. (2008) at $z \sim 3.1$ and $z \sim 3.7$ (Figure 3.4). In fact, at $z = 3.7$, Ouchi et al. (2008) finds higher number densities at all luminosities than ours, although by $z = 4.1$ our measurements agree very well with Ouchi et al. (2008). At $z = 4.8$, SC4K provides a unique opportunity to directly compare results from a MB and NB at roughly the same central wavelength, and we find a very good agreement at all luminosities probed by both bands, with the NB data allowing us to go deeper, while the MB allows to probe a wider volume.

As we move to even higher redshifts ($z \sim 5 - 6$), there is tentative evidence for a 'boost' in luminosity (accompanied by a decline in number density and a potential steepening of the Lyα LF; Drake et al. 2017b), which agrees with results from Santos et al. (2016), and with those at $z \sim 5.7$ from Ouchi et al. (2008) when corrected in the same way as our results (see discussion in e.g. Matthee et al., 2015; Santos et al., 2016, and also Section 3.1.2.2). Recent results from
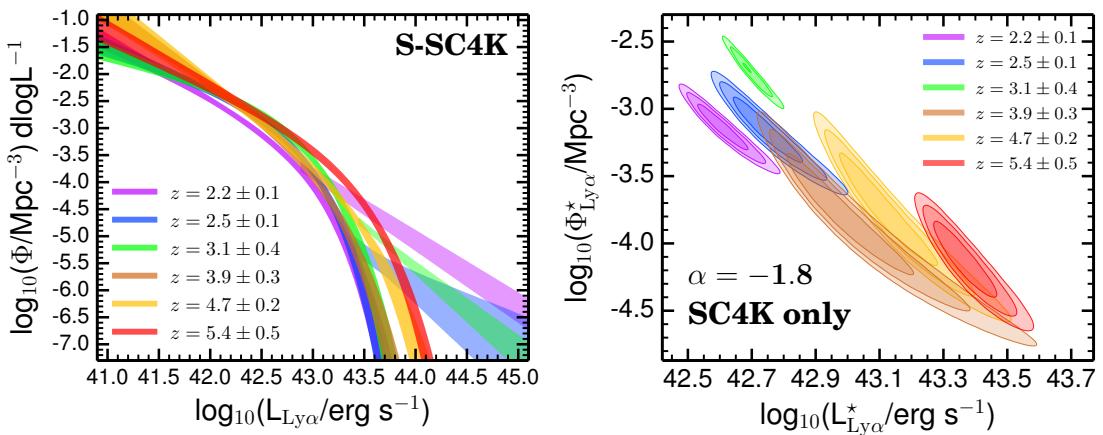
Figure 3.5: *Left:* The evolution of the Lyα LF with redshift from $z \sim 2$ to $z \sim 6$ from this study, exploring our synergy approach (S-SC4K), showing the 16th and 84th percentiles of all realisations/fits. We find a mild $L^{\star}_{Ly\alpha}$ rise with increasing redshift, at the same time that $\Phi^{\star}_{Ly\alpha}$ declines. This leads to a mild evolution in the Schechter-like component with redshift. We find that the extra power-law/Schechter component at $L_{Ly\alpha} > 10^{43.3}\,\mathrm{erg\,s^{-1}}$ declines with increasing redshift, mostly by becoming steeper and with a lower normalisation, which may be linked with the decline in the AGN population. By $z \sim 3.9$ the extra component is no longer seen at the current observational limits. *Right:* The $L^{\star}_{Ly\alpha}$-$\Phi^{\star}_{Ly\alpha}$ contours for the Schechter fits by fixing $\alpha = -1.8$ (without any perturbation) by using the SC4K MBs only. The lines are the 1σ, 2σ and 3σ contours for $L^{\star}_{Ly\alpha}$ and $\Phi^{\star}_{Ly\alpha}$ for each redshift bin. This shows the mild but significant evolution of both $L^{\star}_{Ly\alpha}$ and $\Phi^{\star}_{Ly\alpha}$ with redshift.

HSC (Konno et al., 2018) reach volumes similar to ours at $z = 5.7$ and hint for an overall lower number density of sources than those found by Ouchi et al. (2008) or Santos et al. (2016). This difference is mitigated when we apply the filter profile corrections (see e.g. Figure 3.4), but still suggests an overall lower number density of sources or systematic differences in estimating/measuring fluxes.

3.2.4 S-SC4K: the synergy Lyα LF(z)

Overall, our results show very good agreement with the literature for the range of luminosities where surveys can be directly compared. Our results also extend previous surveys not only to higher luminosities, but also to a much higher number of redshift slices, allowing us to investigate the fine redshift evolution of the Lyα LF in terms of the apparent shape change in the bright end and its positive luminosity evolution (of the main Schechter component) by a factor of about ≈ 5 from $z \sim 3$ to $z \sim 6$ and a decline in the number density of sources by a factor ≈ 10 or more. Interestingly, recent results from MUSE (Drake et al., 2017b) provide strong evidence for α being steep and tentative evidence for it steepening with increasing redshift. However, ultra-deep MUSE data on their own still suffer from an important short-coming: the uncertainty in determining the characteristic luminosity and/or number density of sources (e.g. errors on α up to $^{+1.4}_{-\infty}$ at $z \sim 3 - 6.6$ due to poor constraints on the bright end; see Drake et al., 2017b). Our SC4K survey is exactly what is needed (see Figure 3.3) to provide the extra constraints on the bright end and break the degeneracies.

We combine our SC4K results with other surveys probing to fainter luminosities than SC4K, to derive a synergy/consensus Lyα LF (S-SC4K) from the peak of star-formation into the end of re-ionisation. We present the results in Figures 3.3, 3.4, 3.5 and Tables 3.2 and 3.3. We find evidence for a steepening of the faint-end slope (see Table A.5) from $z \sim 2.5$ ($\alpha = -1.7 \pm 0.2$) to $z \sim 5$ ($\alpha = -2.5 \pm 0.2$). Most importantly, we find that α is always very steep and close to $\alpha = -2$ at all redshifts probed. The synergy LF (S-SC4K; Figure 3.5) also shows a roughly continuous increase in $L^*_{\text{Ly}\alpha}$ by a factor of $\approx 3 - 4$ from $z = 2.5$ to $z \sim 5 - 6$ (for the main Schechter component; note that at $z < 3.3$ the Lyα LF requires an extra bright component to be properly modelled). In addition,

we also find evidence of a decline in the typical number density at $L^*_{Ly\alpha}$, with $\Phi^*_{Ly\alpha}$ continuously reducing by a factor of ≈ 5. Overall, we show that there is evolution in the Lyα LF from $z \sim 2.5$ to $z \sim 6$, driven by an apparent steepening of the faint end slope, together with both a decline in number density and a positive luminosity evolution (factors of $\sim 3 - 5$). It is also worth highlighting that a single Schechter function is not capable of encompassing the full evolution of the Lyα LF at $z \sim 2 - 3.3$, due to the significant power-law or extra brighter Schechter component. We also note that it is possible that the extra population of likely AGN dominating the bright end at lower redshift (see Figure 3.5) may still contribute at higher redshift and may in principle be partially responsible for the luminosity evolution. However, as Section 4.1 shows, due the very steep faint end slope of the Lyα LF, the Lyα luminosity density is dominated by the faintest sources and thus the evolution of the bright end by itself does not dominate the luminosity budget, though it may be very important to understand the physics of sources contributing to it. We also stress that while the bright sources are not the dominant sources of Lyα luminosity density in the Universe, only the combination of ultra-deep and large volume surveys can provide the full constraints necessary to fully measure the evolution of the Lyα LF and the population of sources that contributes to it.

3.3 Conclusions

We use our large sample of ~ 4000 LAEs to construct Lyα LFs for the different redshift slices and investigate the evolution across cosmic time. We also combine SC4K with results from the literature to obtain a powerful consensus/synergy Lyα survey (S-SC4K) that spans over 4 orders of magnitude in Lyα luminosity across $z \sim 2 - 6$. Our main results are:

- SC4K extensively complements ultra-deep surveys, jointly covering over $4 \, \mathrm{dex}$ in Lyα luminosity and revealing a global $(2.5 < z < 6)$ synergy LF with a steep faint end slope $\alpha = -1.93^{+0.12}_{-0.12}$, a characteristic luminosity of $\log_{10} L^*_{Ly\alpha} = 42.93^{+0.15}_{-0.11} \, \mathrm{erg \, s^{-1}}$ and a characteristic number density of $\log_{10} \Phi^*_{Ly\alpha} = -3.45^{+0.22}_{-0.29} \, \mathrm{Mpc^{-3}}$.

- The Schechter component of the Lyα LF shows a factor ~ 5 rise in $L^*_{Ly\alpha}$, from $\approx 10^{42.7}\,\mathrm{erg\,s^{-1}}$ at $z \sim 2$ to $\approx 10^{43.35}\,\mathrm{erg\,s^{-1}}$ at $z \sim 6$ and a $\sim 7\times$ decline in $\Phi^*_{Ly\alpha}$ from $z \sim 2$ to $z \sim 6$. We also find evidence for the faint-end slope to steepen from $\alpha = -1.7 \pm 0.2$ at $z \sim 2.5$ to $\alpha = -2.5 \pm 0.2$ at $z \sim 5$. Most importantly, α is always very steep and close to $\alpha = -2$ at all redshifts probed.

- A Schechter function provides a good fit to the LF up to luminosities of $\sim 10^{43.3}\,\mathrm{erg\,s^{-1}}$, but we find a significant extra power-law (or Schechter) component above $L_{Ly\alpha} = 10^{43.3}\,\mathrm{erg\,s^{-1}}$. We show that the extra component is partially driven by X-ray and radio AGN, as their Lyα LF resembles the excess. This extra component is found to decline (steepen) significantly with redshift and/or becomes mixed with the main Schechter component beyond $z \sim 3.5$, likely linked with the evolution of the AGN population. This means that above $z \sim 3.5$ a single Schechter function becomes a good description of the Lyα luminosity function from the lowest to the highest Lyα luminosities.

Chapter 4

LAEs in the wider picture and the escape of Lyα photons

Abstract

We use our best measurements of the Lyα luminosity function at multiple redshift intervals between $z \sim 2$ and $z \sim 6$ to probe for evolution of the Lyα luminosity density. The Lyα luminosity density rises by a factor ~ 2 from $z \sim 2$ to $z \sim 3$ but is then found to be roughly constant $(1.1^{+0.2}_{-0.2} \times 10^{40}\,\mathrm{erg\,s^{-1}\,Mpc^{-3}})$ to $z \sim 6$, despite the $\sim 0.7\,\mathrm{dex}$ drop in UV luminosity density. The Lyα/UV luminosity density ratio rises from $4 \pm 1\%$ to $30 \pm 6\%$ from $z \sim 2.2$ to $z \sim 6$. Our results imply a rise of a factor of ≈ 2 in the global ionisation efficiency (ξ_{ion}) and a factor $\approx 4 \pm 1$ in the Lyα escape fraction from $z \sim 2$ to $z \sim 6$, hinting for evolution in both the typical burstiness/stellar populations and even more so in the typical ISM conditions allowing Lyα photons to escape.

We structure this Chapter as follows: we measure the Lyα luminosity density ($\rho_{\mathrm{Ly}\alpha}$) and discuss how f$_{\mathrm{esc}}$ and ξ_{ion} likely evolve with redshift in Sections 4.1 to 4.3. We discuss the nature of LAEs in a broader context in Sections 4.4 to 4.6. We provide some final remarks and conclusions in Section 4.7.

4.1 The redshift evolution of $\rho_{\mathbf{Ly}\alpha}$

We explore SC4K and S-SC4K to measure the evolution of the Lyα luminosity density ($\rho_{\mathrm{Ly}\alpha}$) from $z \approx 2.2$ to $z \sim 6$, in multiple redshift slices, with unprecedented detail. We compute $\rho_{\mathrm{Ly}\alpha}$ by integrating the LF down to different limits. For a direct comparison with Hayes et al. (2011), we integrate LFs down to $1.75 \times 10^{41}\,\mathrm{erg\,s^{-1}}$, corresponding to $0.04\,L^{\star}_{z=3}$[1] from Gronwall et al. (2007). For each LF, we calculate 10,000 integrals, each perturbing individual data-points within their asymmetric Gaussian distributions, fitting the LF and computing the integral. For SC4K-only LFs we vary α with a uniform probability distribution between -1.6 and -2.0 for a more conservative error estimation (errors are the 16 and 84 percentiles of all the integrals). The results are shown in Figure 4.1 and Tables 3.2 and 3.3.

We find evidence for $\rho_{\mathrm{Ly}\alpha}$ to increase with redshift, with a rise from $z \sim 2$ to $z \sim 3$ and then a tentative decline at $z \sim 4$ and remaining constant at $z \sim 4-6$ (Figure 4.1). These results are clear using both the individual redshift slices and also the redshift bins. We note that the decline in $\rho_{\mathrm{Ly}\alpha}$ seen from $z \sim 3$ to $z \sim 4$ with SC4K coincides with the disappearance of the bright-end excess of the Lyα LF, although we note that the potential power-law component at the highest luminosities, by itself, only represents $\sim 1-5\%$ of the Schechter luminosity density[2]. The evolution from $z \sim 3.3$ to $z \sim 4$ may be linked with a significant evolution in the nature of Lyα emitters.

When using S-SC4K, we obtain far superior constraints on $\rho_{\mathrm{Ly}\alpha}$ (much better than e.g. MUSE or SC4K on their own; see Figure 4.1). We still find that $\rho_{\mathrm{Ly}\alpha}$ increases from $z \sim 2.2$ to $z \sim 3-4$ by a factor of ≈ 2, and clear evidence for

---

[1]This corresponds to integrating down to $\approx 0.16\,\mathrm{M_\odot\,yr^{-1}}$ for a Salpeter IMF and f$_{\mathrm{esc}} = 1.0$; see Section 4.3.

[2]In our analysis we do not include the integral of the power-law component.
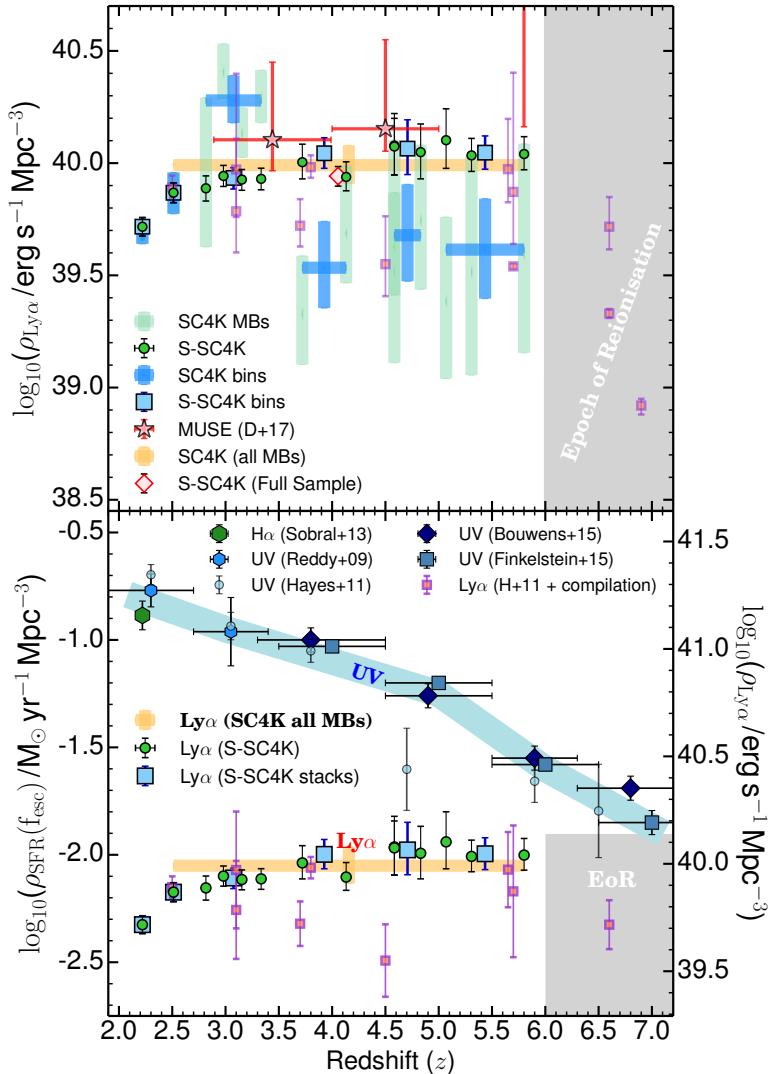
Figure 4.1: *Top:* The evolution of the Lyα luminosity density ($\rho_{\mathrm{Ly\alpha}}$). We show the measurements using SC4K only, including per filter and also redshift stacks. We find a relatively constant $\rho_{\mathrm{Ly\alpha}}$ across redshift, with the MB filters on their own suggesting a slight decline in $\rho_{\mathrm{Ly\alpha}}$ from $z \sim 3$ to $z \sim 4-6$. Combining SC4K with deep surveys (S-SC4K) reveals the importance of probing both the faint and bright ends. The combined constraints show that $\rho_{\mathrm{Ly\alpha}}$ rises from $z \sim 2$ to $z \sim 3.5$ and then stays constant with redshift all the way to $z \sim 6$. *Bottom:* We compare our results with surveys measuring the UV (Bouwens et al., 2015; Finkelstein et al., 2015; Hayes et al., 2011) and Hα (Sobral et al., 2013) luminosity densities transformed to SFRDs. While the global star formation rate density (UV luminosity density) of the Universe is falling sharply from $z \sim 2$ to $z \sim 6$ by a factor of ≈ 5, the contribution from Lyα selected sources is rising, particularly due to the steepening of the Lyα LF, accompanied by a higher typical luminosity and despite the lower typical number density.

$\rho_{\mathrm{Ly}\alpha}$ to be relatively constant with redshift from $z \sim 4$ to $z \sim 6$. Our results thus show that despite the clear evolution of the Lyα LF from $z \sim 4$ to $z \sim 6$, its integral remains roughly constant. Interestingly, we note that the global SC4K LF on its own ($2.5 < z < 6$, pink diamond in Figure 4.1) yields a value of $\rho_{\mathrm{Ly}\alpha}$ which is actually representative of the majority of the individual measurements at $z \sim 3 - 6$ (green circles in Figure 4.1). We find that the relative constancy of $\rho_{\mathrm{Ly}\alpha}$ with increasing redshift is driven by a steepening of the faint-end slope α with increasing redshift, together with an increase in $\mathrm{L}^*_{\mathrm{Ly}\alpha}$, which counter-balances the significant reduction in $\Phi^*_{\mathrm{Ly}\alpha}$ with increasing redshift. Therefore, our results show that whilst $\rho_{\mathrm{Ly}\alpha}$ stays relatively constant with redshift, there is a strong shift towards fainter LAEs becoming more and more dominant in the global $\rho_{\mathrm{Ly}\alpha}$ towards re-ionisation.

We compare our results with the literature (see e.g. Drake et al., 2017b; Hayes et al., 2011; Matthee et al., 2015; Ouchi et al., 2008; Santos et al., 2016; Zheng et al., 2017, and references therein) and find good agreement with our measurements within the errors. The scatter of individual measurements and previous studies done on single fields and/or just probing either the bright or faint regimes is also very clear in Figure 4.1. For example, MUSE data on their own suggest a potential increase in $\rho_{\mathrm{Ly}\alpha}$, while SC4K on its own would suggest a reduction. Our results highlight the importance of combining the strengths of each approach/instrument/measurement in order to truly reveal the behaviour of $\rho_{\mathrm{Ly}\alpha}$ with redshift.

On the bottom panel of Figure 4.1 we convert $\rho_{\mathrm{Ly}\alpha}$ to a star-formation rate density (SFRD$_{\mathrm{Ly}\alpha}$[1]) so we can more directly compare it with the UV luminosity density also converted to SFRD (SFRD$_{\mathrm{UV}}$; e.g. Bouwens et al., 2015; Finkelstein et al., 2015). Our results reveal the striking difference between the evolution of the UV and Lyα SFRDs with increasing redshift. While the SFRD traced by Lyman break galaxies (and Hα emitters at $z = 2.2$) is strongly declining (by a factor of about 5 from $z \sim 2.2$ to $z \sim 6$), the Lyα SFRD is increasing to

---

[1]computed by directly converting $\mathrm{L}_{\mathrm{Ly}\alpha}$ to $\mathrm{L}_{\mathrm{H}\alpha} = \mathrm{L}_{\mathrm{Ly}\alpha}/8.7$, and then computing the SFR using Equation 1.3; a likely difference between this SFRD and the one derived from LBG samples will thus be driven by $\mathrm{f}_{\mathrm{esc,Ly}\alpha}$ and the ionising efficiency; see full assumptions in Section 4.3. The SFR is computed assuming a Salpeter IMF, and thus can be converted to a Chabrier IMF by multiplying it by a scaling factor of 0.63.

$z \sim 3 - 4$ and then remaining constant all the way to the end of the epoch of re-ionisation at $z \sim 6$. Therefore, our results re-enforce the increasing importance of LAEs at higher redshift in the global SFRD, and hint for global evolution in the properties of galaxies for this to happen, including the Lyα escape fraction (which would result in a higher Lyα luminosity density for a fixed UV luminosity density) and/or the typical ionisation efficiency (which can also lead to a higher production of Lyα photons). Furthermore, the Lyα escape fraction is sensitive to a number of galaxy properties such as the dust content (e.g. Atek et al., 2008; Hayes et al., 2010b; Matthee et al., 2016; Oyarzún et al., 2017; Shibuya et al., 2014) and covering fraction of neutral hydrogen (Henry et al., 2015), and thus any of these may be evolving. The production efficiency of ionising photons is related to the nature of stellar populations, such as the metallicity and initial mass function (e.g. Erb et al., 2014; Reddy et al., 2018; Schaerer, 2003). In Section 4.3 we explore these possibilities in detail.

4.2 The evolution of the cosmic Lyα/UV ratio

Based on our results, $\rho_{\mathrm{Ly}\alpha}$ rises by a factor of about ~ 2 from $z \sim 2.2$ to $z \sim 3$ and is then relatively constant up to $z \sim 6$. However, as shown in Figure 4.1, the UV luminosity density[1] decreases by a factor ≈ 5 over the same redshift range (e.g. Bouwens et al., 2015; Finkelstein et al., 2015; Reddy & Steidel, 2009). Figure 4.2 shows that the cosmic $\mathrm{SFRD}_{\mathrm{Ly}\alpha}/\mathrm{SFRD}_{\mathrm{UV}}$ increases significantly with redshift by a factor of $\sim 7 - 8$ from $z \sim 2$ to $z \sim 6$, driven by the mild positive evolution of $\rho_{\mathrm{Ly}\alpha}$ with redshift and the sharp decline in ρ_{UV} (Figure 4.1). Our measurements follow a similar trend estimated by Hayes et al. (2011), but provide significantly better sampling in terms of redshift and further constraining both the bright (SC4K) and faint ends (S-SC4K); see Figure 4.2.

Observationally, our results mean that from $z \sim 2$ to $z \sim 6$ there is a systematic increase in the luminosity density of Lyα photons in the Universe relative to 1500 Å UV photons. Such increase should be vastly dominated by the large

---

[1]UV luminosity densities are integrated down to $0.04\,\mathrm{L}^{\star}_{\mathrm{UV},z=3}$ following Hayes et al. (2011); see also discussions in Hayes et al. (2011) and e.g. Sobral et al. (2017) on integration limits.
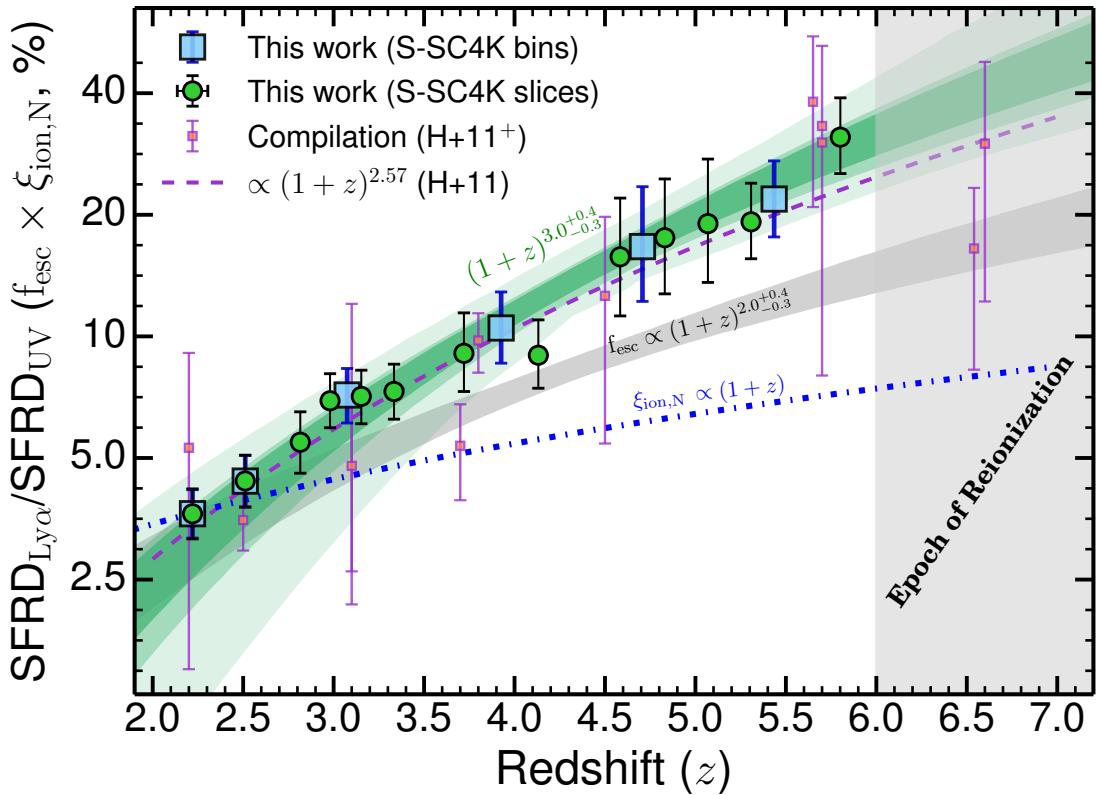
Figure 4.2: The evolution of $SFRD_{Ly\alpha}/SFRD_{UV}$ ($= \xi_{ion,N} \times f_{esc}$, Equation 4.7) from $z = 2.2$ to $z \sim 6$ with S-SC4K. We find the ratio to increase from $\approx 4\%$ at $z = 2.2$ to $\approx 30\%$ at $z \sim 6$, implying a very high Lyα to UV luminosity density ratio in the early Universe. We parameterise the rise with redshift as a power-law and find $\propto (1+z)^{3.0^{+0.4}_{-0.3}}$ (we show the 1, 2 and $3\,\sigma$ range of all fits), a slightly steeper relation than in Hayes et al. 2011 (we also include more recent measurements from the literature). Furthermore, by modelling the rise of ξ_{ion} as $\propto (1+z)$ (see Equation 10 from Matthee et al., 2017a), we infer that f_{esc} is rising as $(1 + z)^{2.0^{+0.4}_{-0.3}}$. Our results suggest a significant evolution in the typical burstiness/stellar populations (ξ_{ion}) by a factor of ≈ 2 and an even stronger evolution in the typical ISM conditions leading to an inferred f_{esc} increase of a factor ≈ 4 from $z \sim 2.2$ to $z \sim 6$.

number of faint LAEs that likely become more dominant towards higher redshift, but there is also independent evidence for a higher Lyα/UV ratio for fixed UV luminosities towards $z \sim 6$, including at high UV luminosities (see Curtis-Lake et al., 2012a; Schenker et al., 2014; Stark et al., 2017, and references therein). We explore and discuss potential explanations and interpretations for the rise of the cosmic Lyα/UV ratio in Section 4.3.

4.3 The redshift evolution of f_esc and ξ_ion

The $\rho_{\mathrm{Ly}\alpha}/\rho_{\mathrm{UV}}$ ratio is a tracer for the relative strength of Lyα to the UV. In Figure 4.2 we show that $\mathrm{SFRD}_{\mathrm{Ly}\alpha}/\mathrm{SFRD}_{\mathrm{UV}}$ ($\sim \rho_{\mathrm{Ly}\alpha}/\rho_{\mathrm{UV}}$) rises with redshift significantly. In order to fully interpret and discuss the redshift evolution of the $\rho_{\mathrm{Ly}\alpha}/\rho_{\mathrm{UV}}$ ratio, it is necessary to derive how it depends on the Lyα escape fraction and production efficiency of ionising photons. We follow Bouwens et al. (2016) and Matthee et al. (2017a), and define ξ_{ion} (see discussions in Shivaei et al., 2018), the production efficiency of hydrogen ionising photons (Lyman continuum, LyC), as:

$$\xi_{\mathrm{ion}} = \frac{\mathrm{Q}_{\mathrm{ion}}}{\mathrm{L}_{\mathrm{UV}}} \times (1 - \mathrm{f}_{\mathrm{esc,LyC}}) \; (\mathrm{Hz\,erg}^{-1}), \tag{4.1}$$

where L_{UV} is the dust-corrected UV luminosity in $\mathrm{erg\,s}^{-1}\,\mathrm{Hz}^{-1}$ at a wavelength of 1500 Å, and assuming a $\approx 0\,\%$ escape fraction of LyC photons ($\mathrm{f}_{\mathrm{esc,LyC}}$). $\mathrm{Q}_{\mathrm{ion}}$, the number of emitted ionising (LyC) photons per second, is related to the dust-corrected Hα luminosity ($\mathrm{L}_{\mathrm{H}\alpha}$) as:

$$\mathrm{Q}_{\mathrm{ion}} = \frac{\mathrm{L}_{\mathrm{H}\alpha}}{\mathrm{c}_{\mathrm{H}\alpha}} \; (\mathrm{s}^{-1}), \tag{4.2}$$

where $c_{\mathrm{H}\alpha} = 1.37 \times 10^{-12}$ erg (e.g. Kennicutt, 1998; Schaerer, 2003) is the recombination coefficient. Under the assumption of case B recombination, a temperature of 10^4 K, an electron density $350\,\mathrm{cm}^{-3}$ and a 0% escape fraction of ionising LyC photons, the Hα luminosity is related to Lyα (with $\mathrm{f}_{\mathrm{esc}}$ being the Lyα escape fraction) as:

$$\mathrm{L}_{\mathrm{H}\alpha} = \frac{\mathrm{L}_{\mathrm{Ly}\alpha}}{8.7 \mathrm{f}_{\mathrm{esc}}} \; (\mathrm{erg\,s}^{-1}). \tag{4.3}$$

With our assumptions so far, we can use $L_{H\alpha}$ to estimate the SFR[1], following Kennicutt (1998) for a Salpeter IMF $(0.1 - 100\,M_\odot)$:

$$\text{SFR}_{H\alpha} = 7.9 \times 10^{-42}\,L_{H\alpha}\ (M_\odot\,\text{yr}^{-1}). \tag{4.4}$$

We combine these equations to derive an expression for the relation between the Lyα and UV luminosities:

$$\xi_{\text{ion}} \times f_{\text{esc}} = \frac{L_{\text{Ly}\alpha}}{8.7c_{H\alpha}L_{\text{UV}}}\ (\text{Hz}\,\text{erg}^{-1}). \tag{4.5}$$

Quantitatively, both UV and Lyα luminosities are related to the SFR. The (dust-corrected) UV luminosity through direct continuum emission from young stars, and Lyα luminosity through the recombination radiation in H$\textsc{ii}$ regions from LyC photons originating from young stars. Following Kennicutt (1998), ξ_{ion} is related to the Hα and UV SFR as:

$$\xi_{\text{ion}} = 1.3 \times 10^{25}\frac{\text{SFR}_{H\alpha}}{\text{SFR}_{\text{UV}}}\ (\text{Hz}\,\text{erg}^{-1}). \tag{4.6}$$

In this equation, the constant $1.3 \times 10^{25}\,\text{Hz}\,\text{erg}^{-1}$ is dependent on the IMF and stellar spectral synthesis models. The ratio between the Hα and UV SFRs is a measure of burstiness of SF (see also Smit et al., 2016) and is equal to 1 if there is a continuous SF history for the last $100\,\text{Myr}$. Therefore, an increasing ξ_{ion} could trace both the nature of stellar populations (i.e. the hardness of the ionising spectrum) and/or the burstiness of star formation. This degeneracy can be resolved with photo-ionisation modelling when multiple emission-lines with a range of ionisation energies are observed (for example using the Helium Balmer lines). If we define $\xi_{\text{ion,N}} = \xi_{\text{ion}}/(1.3 \times 10^{25}\,\text{Hz}\,\text{erg}^{-1})$, we can write:

$$\xi_{\text{ion,N}} \times f_{\text{esc}} = \frac{\text{SFR}_{\text{Ly}\alpha}}{\text{SFR}_{\text{UV}}}, \tag{4.7}$$

allowing us to more directly interpret the ratio between $\text{SFR}_{\text{Ly}\alpha}$ and SFR_{UV}.

---

[1]For continuous SF over $10\,\text{Myr}$ timescales.

Matthee et al. (2017a) discusses how ξ_{ion} correlates with Hα EW, and how the widely agreed rise of typical Hα EWs[1] with redshift (e.g. Faisst, 2016; Fumagalli et al., 2012; Sobral et al., 2014) suggests that ξ_{ion} rises by a factor of about ~ 2 from $z \sim 2$ to $z \sim 6$ as $\propto (1+z)$, in agreement with e.g. Nakajima et al. (2016) and Harikane et al. (2018). Assuming $\xi_{ion,N} \approx 1$ at $z = 2.2$ (see Matthee et al., 2017a; Shivaei et al., 2018), we can then measure f_{esc} directly for $2 < z < 6$ by using:

$$f_{esc} = \frac{3.2}{(1+z)} \frac{SFR_{Ly\alpha}}{SFR_{UV}} \ (2 < z < 6). \tag{4.8}$$

We check with Sobral et al. (2017) that the above approach is able to roughly recover f_{esc} at $z = 2.2$ measured directly with Hα (4% with the integration limits we use and without using the power-law component of the Lyα LF). By comparing with our observations in Figure 4.2, we infer an evolution of f_{esc} of a factor ≈ 4 (from $\approx 3.8\%$ at $z \sim 2.2$ to $\approx 15\%$ at $z \sim 6$), with an increase roughly proportional to $(1+z)^{2.0\pm0.3}$ for f_{esc} (see Figure 4.2). Our results thus suggest that the strong evolution in the $SFRD_{Ly\alpha}/SFRD_{UV}$ ratio with redshift is driven by an increase in ξ_{ion} (tracing high burstiness and/or an average change in stellar populations) by a factor of ~ 2, rising as $1 + z$ and f_{esc} by a factor of $\approx 4 - 5$ from $z \sim 2$ to $z \sim 6$, rising as $(z+1)^2$. Overall, this explains the rise of $SFRD_{Ly\alpha}/SFRD_{UV}$ as $(1+z)^{3.0\pm0.3}$. Our results thus imply evolution in both ISM conditions and on the burstiness/nature of the stellar populations with increasing redshift.

4.4 The compact nature of LAEs and relation to the global increase in $\mathbf{f_{esc}}$

Paulino-Afonso et al. (2018) presents the full visual and automated morphological and structural analysis in the rest-frame UV of the SC4K sample presented in this work. They find that LAEs are systematically smaller in the rest-frame UV than the global population of star-forming galaxies, presenting sizes which are

---

[1]See also results showing a rise in typical EWs of other rest-frame optical lines such as [OIII] in Khostovan et al. (2016).

roughly constant with redshift of $\approx 1\,\mathrm{kpc}$ (see also Bond et al., 2011; Guaita et al., 2015; Malhotra et al., 2012). Paulino-Afonso et al. (2018) also points out that while "typical" star-forming galaxies at $z < 2$ are $\sim 2-4\times$ larger than LAEs (Paulino-Afonso et al., 2017; Ribeiro et al., 2016; van der Wel et al., 2014), the differences in typical sizes become smaller with increasing redshift. By $z \sim 6$, the general population of SFGs presents basically the same morphological properties as LAEs have across all redshifts.

Furthermore, Paulino-Afonso et al. (2018) also discusses how the sizes and compactness of LAEs depend on rest-frame Lyα EW$_0$. The EW$_0$ of the Lyα line has recently been shown to be the simplest/most robust empirical predictor of $\mathrm{f_{esc}}$ (Sobral et al., 2017), with the relation between EW$_0$ and $\mathrm{f_{esc}}$ showing no significant evolution at $z \sim 0 - 5$, (see Harikane et al., 2018; Sobral et al., 2017). Paulino-Afonso et al. (2018) find that LAEs with the highest EWs are the smallest and most compact at all redshifts. This suggests a relation between compactness and/or size and $\mathrm{f_{esc}}$, and may be one of the physical reasons why we find that globally $\mathrm{f_{esc}}$ seems to rise with increasing redshift. In this case, it would be because the general population of galaxies are, as a whole, compact and small enough, for Lyα photons to more easily escape. However, we note that smaller and more compact galaxies will typically be also less evolved, potentially more bursty and with lower metallicity stellar populations, which can also lead to boosting Lyα through a higher ξ_{ion}. The potentially higher $\mathrm{f_{esc}}$ at higher redshift could also be caused more directly by e.g. lower dust content and/or a more porous CGM due to strong stellar winds (e.g. Geach et al., 2014) produced in compact and highly star-forming regions, which would allow the escape of more Lyα photons.

The morphological information may be important to potentially explain the increase in $\mathrm{f_{esc}}$ with redshift, but in principle it does not tell us anything about the burstiness or the stellar populations and/or AGN activity that may be happening within LAEs across cosmic time. This is important to understand the potential evolution in ξ_{ion} (Matthee et al., 2017a), even more so as our results provide evidence that both ξ_{ion} and $\mathrm{f_{esc}}$ evolve with redshift. Further physical insight may be obtained by studying local analogues like 'green peas' or 'blueberry' galaxies (e.g. Izotov et al., 2017; Yang et al., 2017a,b). Such analogue galaxies

allow for detailed studies to be performed to make crucial measurements and test hypothesis/modelling results (e.g. Izotov et al., 2016; Verhamme et al., 2006, 2015; Yang et al., 2017a) regarding the connection between f_{esc} and the Lyα emission line peak separation, width and other properties (see e.g. Verhamme et al., 2017). Furthermore, these low redshift sources, showing essentially the same properties as SC4K galaxies at higher redshift, are ideal to further explore and test the link between LyC and Lyα photon escape (e.g. Dijkstra et al., 2016; Izotov et al., 2018; Verhamme et al., 2015, 2017) and their relation with size/compactness and other physical properties.

4.5   The bright end of the Lyα LF: AGN?

Previous studies (e.g. Konno et al., 2016; Matthee et al., 2017b; Sobral et al., 2017; Wold et al., 2017) have found evidence for a relation between the potential power-law component of the bright end of the Lyα LF and the AGN nature of sources populating it. Such evidence has been primarily driven by the detection of many of those sources in the X-rays (e.g. Konno et al., 2016; Sobral et al., 2017). With the availability of deep *Chandra* and VLA data, we have identified that $3.6 \pm 0.3\%$ of all our sources are likely AGN, with 109 ($2.9 \pm 0.3\%$) being X-ray AGN, 62 ($1.7 \pm 0.2\%$) being radio AGN and 30 ($0.8 \pm 0.1\%$) being both. While these are a very small fraction overall, as shown in Calhau et al. (2020) and Sobral et al. (2018a), AGN LAEs become more significant at the brightest Lyα luminosities, a consequence of their relatively flat Lyα LF which we have found, with a potential high $L^*_{Ly\alpha}$. Calhau et al. (2020) finds a significant correlation between the X-ray AGN fraction of LAEs and the Lyα luminosity; this fraction is consistent with $0.7 \pm 0.3\%$ below $L^*_{Ly\alpha}$, but it grows towards 100% at the highest Lyα luminosities (see also Matthee et al., 2017b; Sobral et al., 2018a). We thus find that removing the X-ray and radio AGN leads to removing sources from the bright-end of the LF, but an excess relative to a Schechter persists at $z \sim 2-3$ even after removing X-ray and radio sources. We argue that there is still a significant population of AGN sources that is undetected in the radio and X-rays, even after stacking. X-ray or radio-detected AGN only provide a lower constraint on the

total number of AGN, as not all AGN have strong X-ray or radio emission. As shown in Sobral et al. (2018a), virtually all the spectroscopically confirmed LAEs at $z \sim 2-3$ with $> 10^{43.2}\,\mathrm{erg\,s^{-1}}$ are AGN. Such AGN are revealed by deep rest-frame UV spectroscopy, even though the majority does not show any detectable X-ray or radio emission. These results indicate that the most luminous LAEs at $z \sim 2-3$ are powerful AGN that emit copious amounts of Lyα photons, boosting the bright end of the Lyα LF. Further evidence comes from the relation between X-ray and Lyα luminosities which suggests that Lyα is tracing the accretion rate for those sources, and not SF processes. AGN LAEs have X-ray luminosities in the range $\mathrm{L_{X-ray}} = 10^{43.4-45.1}\,\mathrm{erg\,s^{-1}}$, implying high black hole accretion rates of 0.1-4 $\mathrm{M_\odot\,yr^{-1}}$. AGN LAEs have radio luminosities of $\approx 10^{30.7}\,\mathrm{erg\,s^{-1}\,Hz^{-1}}$, but little relation with Lyα, probing down to lower Lyα luminosities, and potentially indicating 'bursty' AGN accretion.

Calhau et al. (2020) discusses how the relation between AGN fraction and Lyα luminosity evolves with redshift, consistent with a decline in the normalisation or an evolution towards much higher Lyα luminosities. For $3.5 < z < 6$ (where we fail to detect the power-law component), the X-ray+radio AGN fraction of LAEs remains relatively low for the entire luminosity range, although it still rises with Lyα luminosity from $0.9 \pm 0.4\%$ at the lowest Lyα luminosities to $11 \pm 7\%$ at $\approx 10^{44}\,\mathrm{erg\,s^{-1}}$. These results are consistent with those from Wold et al. (2014, 2017) at $z \sim 0-1$, but provide evidence for the AGN fraction evolving (declining) with redshift. The different X-ray and radio observational limits at different redshifts are not sufficient to explain the evolution of the AGN fraction of LAEs (Calhau et al., 2020). While we find no convincing evidence of a significant population of AGN LAEs beyond $z > 3.5$, and no detectable power-law component in the LF, it is possible that it continues to exist at $z > 3.5$, but just with number densities below our surveyed volumes and/or with a LF that is more similar to the fainter population of LAEs, thus making it indistinguishable from those. If these sources occupy the faint-end of the quasar luminosity function, one would potentially expect number densities of $10^{-9} - 10^{-10}\,\mathrm{Mpc^{-3}}$ (McGreer et al., 2013) for the most luminous $z = 5$ quasars, which would be easily below our detection limit. It is also possible that the bright end still contains AGN sources even towards $z \sim 6$, but that they are just not X-ray or radio luminous

enough to be detected either individually or by stacking (see Calhau et al., 2020). Such potential "hidden" AGN activity in luminous LAEs at higher redshift could still be driving the apparent $L^*_{Ly\alpha}$ rise towards $z \sim 6$ and might be tentatively showing up in deep spectroscopic observations of some of the most luminous LAEs at $z \sim 6 - 7$ with potential detections of HeII and/or Nv (e.g. Laporte et al., 2017; Sobral et al., 2019). In addition, we also note that while high accretion rates and relatively high black hole masses in fainter LAEs are excluded, faint LAEs may still contain young, low mass AGN that would make them currently undetectable in the X-rays and radio.

4.6 The nature and evolution of faint to bright LAEs across z $\sim$ 2 − 6: progenitors of sub-L* galaxies to proto-cluster tracers

Clustering analysis (Khostovan et al., 2018) of the SC4K sample shows a clear dependence of the clustering length and the inferred dark matter halo mass on both the Lyα luminosity and the UV luminosity or SFR. At the highest Lyα luminosities, LAEs are likely hosted by quite massive dark matter haloes of $10^{13-14}\,M_\odot$, where one expects AGN activity to be prominent. These observational results are in good agreement with modelling from e.g. Garel et al. (2016) who finds that the brightest LAEs at high redshift should reside in more massive dark matter haloes and be the progenitors of more massive haloes today, while the super faint LAEs now being found by MUSE (Drake et al., 2017b) are likely the progenitors of sub-L* galaxies today. Khostovan et al. (2018) finds similar results, with the dark matter haloes and the clustering strength of the faintest LAEs from the narrow band selected surveys being closer to $\sim 10^{11}\,M_\odot$, similar to results from e.g. Ouchi et al. (2010) and other clustering studies focusing on very faint LAEs (e.g. Kusakabe et al., 2018). The high number densities of faint LAEs at high redshift, driven by the steep ($\alpha \approx -2$) faint-end slope of the Lyα LF (S-SC4K and e.g. Drake et al., 2017b; Dressler et al., 2015) reveal that a very large number of sources are emitting Lyα photons that can escape in the early Universe. These

numerous LAEs (this study and e.g. Drake et al., 2017b) with high Lyα escape fractions and high EWs (e.g. Hashimoto et al., 2017; Sobral et al., 2017), highly ionising (Nakajima et al., 2016), compact/small sources (Malhotra et al., 2012; Paulino-Afonso et al., 2018) may play a crucial role in the early Universe. For example, our results imply that by $z \sim 6$, LAEs are likely key contributors to the global LyC photons produced in the Universe (see also the discussion about sources that can reionise the Universe in Section 1.1.2).

Overall, LAEs have low UV luminosities (which can easily make them undetected even in very deep continuum surveys), but high production of LyC photons (expressed as a high ionisation efficiency; Harikane et al. 2018; Matthee et al. 2017a; Nakajima et al. 2016). Thus, our results strongly add to current observations by pointing towards LAEs being exactly the sources that ultra-deep continuum surveys strive to detect using gravitational lensing (e.g. Atek et al., 2015). Due to the strength and high EWs of the Lyα emission line at high redshift, LAE surveys are simply much more efficient at picking the numerous, UV-faint, compact and highly ionising sources in spite of their ultra-faint UV magnitudes. Examples of such faint, strongly Lyα emitting galaxies have recently been found in e.g. Vanzella et al. (2016). Furthermore, recent results of local galaxies showing the same properties as SC4K sources (including M_{UV}, Lyα EWs and sizes e.g. Izotov et al., 2016, 2018) provide even more evidence for the importance of LAEs in the early Universe in terms of their contribution to both the SFRD and as sources that can help reionise the Universe. While the Lyα emission that we detect is mostly powered by ionising photons which are absorbed and reprocessed by neutral gas around galaxies, it has been shown that the escape fraction of LyC and the escape fraction of Lyα photons are correlated (Dijkstra, 2017). The detection of these mostly UV-faint Lyα emitting galaxies which are capable of releasing ionising photons provides a look at the type of galaxies that have the conditions to contribute to the reionisation of the Universe.

SC4K is also able to find some of the rarest, brightest LAEs across cosmic time which are likely powered by AGN. Most importantly, the brightest LAEs seem to be highly clustered, and there is convincing evidence that they trace, on average, some of the densest regions of the Universe usually classed as 'proto-clusters' (e.g. Franck & McGaugh, 2016). This is because the brightest LAEs

within SC4K across the COSMOS field are hosted by dark matter haloes of $\sim 10^{13-14}\,\mathrm{M}_\odot$ at $z > 2.5$ (Khostovan et al., 2018), which will easily result in massive clusters of $\sim 10^{14-15}\,\mathrm{M}_\odot$ in the local Universe when extrapolating to the present day using halo mass accretion growth. The number densities of these sources also agrees with our findings, being below $10^{-6}\,\mathrm{Mpc}^{-3}$. The results thus bring further context into the findings of bright LAEs in or around some of the most over-dense regions in the Universe at $z \sim 2 - 6$ (Venemans et al., 2007; Yamada et al., 2012), including e.g. Lyα 'blobs' (e.g. Kubo et al., 2013; Matsuda et al., 2004) and point towards large volume Lyα surveys as ideal ways to find these extremely over-dense regions. Given the high fraction of AGN among the population of these very high luminosity LAEs, it is not surprising that many studies also find those sources (e.g. X-ray or radio detected; see e.g. Venemans et al. 2007) to be good tracers of over-densities throughout the Universe (see Kubo et al., 2013; Lehmer et al., 2009; Matsuda et al., 2011; Overzier, 2016, and references therein).

4.7 Conclusions

We use the Lyα LFs which we derive by combining our large SC4K sample of LAEs with results from the literature (synergy Lyα survey, S-SC4K) to measure the evolution of the Lyα luminosity density and Lyα f_{esc} from $z \sim 2$ to $z \sim 6$. We also discuss LAEs in a broader context, in their nature and as progenitors of a wide range of galaxies which multiple surveys seek to target. Our main results are:

- The Lyα luminosity density rises by a factor ~ 2 from $z \sim 2$ to $z \sim 3$ but is then found to be roughly constant ($1.1^{+0.2}_{-0.2} \times 10^{40}\,\mathrm{erg\,s^{-1}\,Mpc^{-3}}$) to $z \sim 6$, despite the $\sim 0.7\,\mathrm{dex}$ drop in UV luminosity density. As a consequence, the $\mathrm{SFRD_{Ly\alpha}/SFRD_{UV}}$ ratio rises from $4 \pm 1\%$ to $30 \pm 6\%$ from $z \sim 2.2$ to $z \sim 6$. LAEs become increasingly important as SFRD contributors into the epoch of re-ionisation, and not simply a relatively minor/rare population.

- Our results are consistent with a rise of a factor of ≈ 2 in the cosmic ionisation efficiency (ξ_{ion}) and imply a factor $\approx 4 \pm 1$ increase in the cosmic

f_{esc} from $z \sim 2$ to $z \sim 6$. We find that an increase of f_{esc} with redshift as $(1+z)^{2.0\pm0.3}$ and a further increase of ξ_{ion} as $(1+z)$ can successfully model the global increase of $SFRD_{Ly\alpha}/SFRD_{UV}$ as $(1+z)^{3.0\pm0.3}$.

- Our results hint for evolution in both the typical burstiness/stellar populations and even more so in the typical ISM conditions for Lyα photons to escape more efficiently at higher redshift. These trends may well be connected with the typically younger and more metal-poor galaxies becoming more dominant – explaining the higher typical ξ_{ion} – and also typically smaller/more compact morphologies, likely linked with the rise of f_{esc}. SC4K LAEs are ideal follow-up candidates for these scenarios to be tested with current state-of-the-art and upcoming instruments/telescopes.

Chapter 5

Aperture photometry and spectral energy distribution of $z \sim 2 - 6$ SC4K LAEs

Abstract

We extend our follow-up of the SC4K sample (Chapter 2) of ~ 4000 $z \sim 2 - 6$ Lyman-α (Lyα) emitters (LAEs) by exploring deep rest-frame UV to FIR data in the COSMOS field. We measure aperture photometry of individual LAEs in 30+ bands to compute their spectral energy distributions (SED).

5.1 Introduction

The Lyman-α (Lyα, $\lambda_{0,\text{vacuum}} = 1215.67\,\text{Å}$) emission line has been predicted to be associated with young star-forming galaxies (SFGs, e.g. Partridge & Peebles, 1967) but it can also be emitted by active galaxy nuclei (AGN; e.g. Miley & De Breuck, 2008; Sobral et al., 2018b). Typical Lyα emitters (LAEs) selected with deep surveys have been found to have low stellar mass ($M_\star \lesssim 10^9\,M_\odot$), low dust content and high specific star formation rates (e.g. Gawiser et al., 2006, 2007), but LAEs can span a wide range in different properties (e.g. Hagen et al., 2016; Matthee et al., 2016). Observationally, the transition between the dominant powering source in LAEs seems to occur at $\sim 10^{43}\,\text{erg s}^{-1}$, roughly two times the characteristic Lyα luminosity ($L^\star_{\text{Ly}\alpha}$) at $z \sim 2-3$ (see Sobral et al., 2018b).

Searches using the Lyα emission line have been extremely successful at selecting young SFGs through narrow band searches (e.g. Arrabal Haro et al., 2018; Harikane et al., 2018; Hu et al., 2004; Matthee et al., 2015; Ouchi et al., 2008; Santos et al., 2016; Sobral et al., 2017) and spectroscopically confirming bright LAEs (e.g. Hu et al., 2016; Matthee et al., 2017c; Shibuya et al., 2018; Sobral et al., 2015, 2018b) due to the bright Lyα feature. Other studies have successfully selected samples of LAEs using integral field spectroscopy observations (e.g. Bacon et al., 2015; Blanc et al., 2011; Drake et al., 2017a; van Breukelen et al., 2005) and blind spectroscopy (e.g. Cassata et al., 2011; Le Fèvre et al., 2015; Martin & Sawicki, 2004; Rauch et al., 2008). LAEs typically have faint continua, and thus the study of properties of individual sources has typically only been done for extreme LAEs with $L \gtrsim L^\star_{\text{Ly}\alpha}$ (e.g. Ouchi et al., 2013; Sobral et al., 2015). For $\lesssim L^\star_{\text{Ly}\alpha}$ LAEs, studies have typically resorted to stacking of sources (e.g. Kusakabe et al., 2018; Momose et al., 2014). More commonly, large samples of high-redshift SFGs have been selected by searching for the presence of a Lyman Break (e.g. Madau et al., 1996; Steidel et al., 1996b, 1999). Currently, there are $> 10,000$s of known galaxies at $z \sim 2 - 10$ (see e.g. Bouwens et al., 2014a, 2015), mostly consisting of faint sub-$L^\star_{\text{UV}}$ galaxies found through deep small area searches, typically too faint to follow-up with current spectroscopic instrumentation.

While Lyα surveys are efficient at selecting galaxies, inferring intrinsic properties of a galaxy directly from its Lyα emission is challenging due to the complex

nature of Lyα radiative transfer. Lyα photons suffer resonant scattering from gas in the Interstellar/Circumgalactic Medium (ISM/CGM) and get easily absorbed by dust (for a review on the process of Lyα radiative transfer see Dijkstra, 2017) which can suppress Lyα emission even in young SFGs. The complex physics of Lyα radiative transfer means that the Lyα escape fraction ($f_{esc,Ly\alpha}$ - the ratio between observed and intrinsic Lyα luminosity, Equation 1.6; see Section 1.3.3) is difficult to predict. Multiple studies have taken different approaches to this problem. Observationally, $f_{esc,Ly\alpha}$ has been measured by comparing Lyα to dust-corrected Hα luminosities (Matthee et al., 2016; Oteo et al., 2015; Sobral et al., 2017). Some studies estimate $f_{esc,Ly\alpha}$ by computing the ratio between star formation rate (SFR) derived from Lyα (assuming case B recombination) and SFR derived from alternative methods such as from spectral energy distributions (SEDs, Cassata et al., 2015) or the far-infrared (FIR, Wardlow et al., 2014). Others measure the ratio between the observed Lyα luminosity density and the dust-corrected Hα luminosity density (Sobral et al., 2017). Alternatively, studies have measured the ratio between Lyα SFR density (SFRD) and UV SFRD by integrating the respective luminosity functions (Sobral et al., 2018a). Typical SFGs at $z \sim 2 - 3$ are found to have very low $f_{esc,Ly\alpha}$ ($< 5\%$, e.g. Cassata et al., 2015; Hayes et al., 2010a; Matthee et al., 2016; Oteo et al., 2015). However, sources selected due to their Lyα emission have much higher $f_{esc,Ly\alpha}$ (as high as $\sim 40\%$ at $z = 2.2$, Sobral et al., 2017).

Despite the complexity of the Lyα radiative transfer, properties of the Lyα line such as its equivalent width (EW) have been shown to hold important information. Sources selected by their Lyα emission have high rest-frame Lyα EW (EW_0)[1] $\sim 50 - 150\,$Å at $z \sim 0.3 - 6$ (see e.g. Gronwall et al., 2007; Hashimoto et al., 2017; Wold et al., 2017) which can be explained by young stellar ages, low metallicities and/or top-heavy initial mass functions (Raiter et al., 2010; Schaerer, 2003) or complex radiative transfer effects (Neufeld, 1991). The high Lyα EW_0 measured for LAEs even at low redshift ($z \sim 0.3$, Wold et al., 2017) contrasts with

---

[1]This is in part due to selection, as LAEs selected by narrow/medium band searches are selected to be above some Lyα EW threshold, and thus are, by definition, sources with high Lyα EW. Nevertheless, some narrow band searches have successfully selected LAEs down to $\sim 5\,$Å (e.g. Arrabal Haro et al., 2018; Sobral et al., 2017).

rest-frame EW measurements from other emission lines for galaxies at similar red-shifts (e.g. Hα, [OII] and Hβ + [OIII] EW$_0$) which are measured to be $\leq 25\,\text{Å}$ at $z \sim 0.3$, (e.g. SDSS: Thomas et al. 2013; HETDEX: Adams et al. 2011). Sobral & Matthee (2019) derived a simple empirical relation that estimates f$_{\text{esc,Ly}\alpha}$ from EW$_0$: f$_{\text{esc,Ly}\alpha}$ = 0.0048$\times$EW$_0$ (Equation 1.8). This relation implies a connection between the intrinsic EW and the dust attenuation. A non-evolution of typical EW$_0$ with redshift could thus imply a non-evolution of f$_{\text{esc,Ly}\alpha}$ in Lyα-selected samples. A constant typical EW$_0$ = 80 Å across redshift would result in a typical f$_{\text{esc,Ly}\alpha} \sim 40\%$ for LAEs.

With the measurement of f$_{\text{esc,Ly}\alpha}$ from EW$_0$, it is possible to derive the SFR of LAEs by translating Lyα flux into dust-corrected Hα flux with simple assumptions. This provides a SFR computation which is independent of SED fitting and provides a comparison with SED-derived SFRs for LAEs even before observations with *James Webb Space Telescope*. Exploring how LAEs, which are typically low stellar mass galaxies, fit in the star formation "Main Sequence" (Brinchmann et al., 2004; Daddi et al., 2007; Noeske et al., 2007; Schreiber et al., 2015) can shed light in a stellar mass range of the SFR-M$_\star$ relation which is still widely unconstrained at $z > 2$. Previous studies have found that LAEs occupy the low stellar mass end of the Main Sequence at $z = 2.5$ (e.g. Shimakawa et al., 2017) but are also measured to be significantly above the Main Sequence extrapolation (Whitaker et al., 2014) for low stellar masses at $z \sim 2$ (e.g. Hagen et al., 2016; Kusakabe et al., 2018) and even at $z = 4.9$ (Harikane et al., 2018). This suggests that LAEs are experiencing more intense star formation than the general population of galaxies of similar mass at similar redshifts, which may be explained by a burstier nature of star formation. We intend to expand these studies using a large sample of LAEs at $z \sim 2 - 6$.

In this work, we use a uniformly selected sample of ~ 4000 LAEs (SC4K, Sobral et al., 2018a, see 2) to measure rest-frame UV properties and their evolution from the end of reionisation at $z \sim 6$ until the peak of star formation history at $z \sim 2$. For our sample of galaxies, we measure EW$_0$, SFR, M$_\star$, UV luminosity (M$_{\text{UV}}$) and UV continuum slope (β) for individual LAEs, using photometry measurements which we conduct ourselves, including data from UltraVISTA DR4,

and by modelling SEDs using MAGPHYS (da Cunha et al., 2008, 2015). Additionally, we discuss different approaches to measure SFR and how they influence our findings and we provide all our measurements in a public catalogue.

We structure this Chapter as follows: in Section 5.2 we give a brief summary of the SC4K sample (see also Chapter 2). We describe our multi-wavelength data in Section 5.3. In Sections 5.4 and 5.5 we detail how we conduct aperture photometry for each individual SC4K LAE. We describe how we obtain SEDs and SED fits for each individual SC4K LAE in Section 5.6. We provide a small summary in Section 5.7.

5.2 The sample: SC4K

We use the public SC4K sample of LAEs (Slicing COSMOS with 4k LAEs, Sobral et al., 2018a, see Chapter 2), which contains 3908 sources selected due to their high Lyα EW at $z \sim 2 - 6$. These LAEs were selected with wide field surveys conducted with Subaru and the Isaac Newton telescopes, using 16 (12+4) medium+narrow bands (MB+NB) over $2 \, \mathrm{deg}^2$ in the COSMOS field (Capak et al., 2007; Scoville et al., 2007; Taniguchi et al., 2015a), covering a full comoving volume of $\sim 10^8 \, \mathrm{Mpc}^3$. For full details on the selection of the sample see Chapter 2. Briefly, the selection criteria applied were i) EW_0 cut of $50 \, \text{Å}$ for MBs, $25 \, \text{Å}$ for NBs and $5 \, \text{Å}$ for the NB at $z = 2.23$: see Sobral et al. 2017); ii) significant excess emission in the selection medium/narrow band, $\Sigma > 3$ (see Bunker et al., 1995; Sobral et al., 2013); iii) colour break blueward of the detected Lyα emission, due to the expected presence of a Lyman Break; iv) removal of sources with strong red colours which are typically lower redshift contaminants where the Balmer break mimics a Lyman break; v) visual inspection of all candidates to remove spurious sources and star artefacts.

Table 5.1: Overview of the SC4K sample of LAEs. We present the median of all measurements for each galaxy property, with the errors being the 16th and 84th percentile of the distribution. (1) LAE selection filter (see Section 2.2); (2) Mean redshift of the sample based on Lyα within the filter FWHM; (3) Number of LAEs (Number of LAEs after removing sources with AGN signatures, see Section 5.2.1); (4) Number of non-AGN LAEs with SEDs (percentage, see Section 5.6.1); (5) Lyα luminosity; (6) Lyα rest-frame EW; (7) SFR derived directly from $L_{Ly\alpha}$ and EW_0 (Sobral & Matthee, 2019, see Section 6.1.6.1); (8) Best likelihood SFR parameter from SED fitting; (9) Best likelihood stellar mass parameter from SED fitting; (10) UV magnitude computed by integrating the SED at $\lambda_0 = 1500\,\text{Å}$, see Section 6.1.3; (11) slope of the UV continuum measured from the SED fits, see Section 6.1.4

(1) Filter	(2) Lyα z	(3) # LAEs (no AGN)	(4) # SEDs	(5) $\log_{10} L_{Ly\alpha}$ (erg s^{-1})	(6) EW_0 (Å)	(7) $SFR_{Ly\alpha}$ (M$_\odot$ yr^{-1})	(8) SFR_{SED} (M$_\odot$ yr^{-1})	(9) M$_\star$ ($\log_{10}$ (M$_\star$/M$_\odot$))	(10) M$_{UV}$ (AB)	(11) β
NB392	2.2	159 (137)	129 (94%)	$42.55^{+0.15}_{-0.15}$	79^{+52}_{-44}	$4.7^{+4.9}_{-2.2}$	$5.5^{+20.5}_{-3.6}$	$9.5^{+0.5}_{-0.6}$	$-19.6^{+1.0}_{-0.6}$	$-1.8^{+0.9}_{-0.5}$
IA427	2.5	741 (686)	673 (98%)	$42.64^{+0.22}_{-0.14}$	128^{+220}_{-62}	$4.0^{+3.1}_{-1.8}$	$2.9^{+6.9}_{-1.5}$	$9.2^{+0.5}_{-0.5}$	$-19.7^{+0.6}_{-0.6}$	$-2.0^{+0.3}_{-0.4}$
IA464	2.8	311 (284)	283 (100%)	$42.88^{+0.22}_{-0.15}$	121^{+152}_{-52}	$6.8^{+4.6}_{-2.4}$	$4.0^{+9.1}_{-1.6}$	$9.1^{+0.6}_{-0.3}$	$-20.2^{+0.5}_{-0.5}$	$-2.1^{+0.5}_{-0.3}$
IA484	3.0	711 (636)	625 (98%)	$42.83^{+0.18}_{-0.11}$	176^{+340}_{-95}	$5.0^{+4.5}_{-2.0}$	$3.1^{+5.8}_{-1.4}$	$9.0^{+0.7}_{-0.3}$	$-20.0^{+0.6}_{-0.7}$	$-2.4^{+0.6}_{-0.0}$
NB501	3.1	45 (38)	31 (82%)	$42.92^{+0.19}_{-0.13}$	170^{+2259}_{-99}	$6.6^{+7.5}_{-3.2}$	$6.2^{+15.2}_{-3.1}$	$9.6^{+0.4}_{-0.5}$	$-20.4^{+1.1}_{-0.8}$	$-2.3^{+1.1}_{-0.2}$
IA505	3.2	483 (437)	433 (99%)	$42.89^{+0.19}_{-0.13}$	142^{+351}_{-71}	$6.3^{+4.9}_{-2.5}$	$4.5^{+6.5}_{-2.0}$	$9.4^{+0.5}_{-0.5}$	$-20.2^{+0.6}_{-0.6}$	$-2.1^{+0.4}_{-0.4}$
IA527	3.3	641 (593)	573 (97%)	$42.84^{+0.19}_{-0.10}$	149^{+245}_{-74}	$5.7^{+5.1}_{-2.3}$	$4.1^{+5.7}_{-1.9}$	$9.4^{+0.6}_{-0.6}$	$-20.2^{+0.5}_{-0.6}$	$-2.0^{+0.3}_{-0.5}$
IA574	3.7	98 (88)	87 (99%)	$42.98^{+0.14}_{-0.13}$	97^{+72}_{-39}	$10.9^{+6.4}_{-4.9}$	$6.7^{+6.9}_{-2.7}$	$9.3^{+0.7}_{-0.2}$	$-20.8^{+0.5}_{-0.4}$	$-2.4^{+0.8}_{-0.0}$
IA624	4.1	142 (139)	116 (83%)	$43.02^{+0.18}_{-0.06}$	186^{+666}_{-99}	$6.7^{+8.2}_{-1.8}$	$6.1^{+9.1}_{-2.8}$	$9.2^{+0.5}_{-0.5}$	$-20.5^{+0.5}_{-0.6}$	$-1.9^{+0.3}_{-0.5}$
IA679	4.6	79 (75)	69 (92%)	$43.25^{+0.15}_{-0.05}$	186^{+267}_{-89}	$11.6^{+12.2}_{-2.8}$	$9.3^{+18.6}_{-4.0}$	$9.5^{+0.8}_{-0.3}$	$-21.2^{+0.6}_{-0.5}$	$-2.4^{+0.8}_{-0.0}$
IA709	4.8	81 (77)	73 (95%)	$43.16^{+0.13}_{-0.10}$	124^{+200}_{-56}	$13.2^{+9.9}_{-5.5}$	$9.1^{+15.8}_{-3.8}$	$9.4^{+0.5}_{-0.3}$	$-21.1^{+0.5}_{-0.4}$	$-2.0^{+0.3}_{-0.5}$
NB711	4.8	78 (74)	56 (76%)	$42.74^{+0.28}_{-0.16}$	80^{+64}_{-42}	$7.8^{+11.2}_{-3.6}$	$14.4^{+61.0}_{-9.5}$	$9.7^{+0.6}_{-0.6}$	$-20.9^{+0.5}_{-0.8}$	$-1.9^{+0.8}_{-0.5}$
IA738	5.1	79 (75)	65 (87%)	$43.25^{+0.17}_{-0.14}$	120^{+222}_{-47}	$15.7^{+15.5}_{-7.6}$	$16.0^{+32.4}_{-9.2}$	$9.6^{+0.7}_{-0.3}$	$-21.3^{+0.4}_{-0.7}$	$-1.8^{+0.2}_{-0.6}$
IA767	5.3	33 (30)	29 (97%)	$43.37^{+0.20}_{-0.07}$	134^{+169}_{-48}	$18.7^{+16.9}_{-7.4}$	$20.6^{+50.5}_{-10.8}$	$9.7^{+0.3}_{-0.4}$	$-21.6^{+0.4}_{-0.5}$	$-2.0^{+0.3}_{-0.4}$
NB816	5.7	192 (186)	108 (58%)	$42.82^{+0.27}_{-0.11}$	235^{+547}_{-169}	$5.2^{+6.4}_{-2.4}$	$28.5^{+83.7}_{-20.8}$	$9.9^{+0.4}_{-0.5}$	$-21.4^{+0.6}_{-0.6}$	$-1.8^{+0.7}_{-0.6}$
IA827	5.8	35 (35)	27 (77%)	$43.44^{+0.19}_{-0.11}$	325^{+963}_{-266}	$22.0^{+47.5}_{-8.4}$	$25.3^{+80.1}_{-16.1}$	$9.9^{+0.6}_{-0.4}$	$-22.0^{+0.8}_{-1.0}$	$-1.8^{+0.7}_{-0.6}$
Full SC4K	4.1	3908 (3590)	3377 (94%)	$42.84^{+0.27}_{-0.17}$	138^{+281}_{-70}	$5.9^{+6.3}_{-2.6}$	$4.4^{+10.5}_{-2.4}$	$9.3^{+0.6}_{-0.5}$	$-20.2^{+0.7}_{-0.8}$	$-2.1^{+0.5}_{-0.4}$

We show an overview of the properties of the SC4K LAEs, split by selection bands, in Table 5.1. For each selection band, we provide the median of each property and the 16th (84th) percentiles of its distribution as lower (upper) uncertainties. Additionally, in Figure 5.1 we show a histogram distribution of Lyα luminosity ($L_{Ly\alpha}$), EW_0 (see Section 6.1.1) and SFR using the Sobral & Matthee (2019) calibration (see Section 6.1.6). The differences in the lower end distribution of $L_{Ly\alpha}$ are driven by an increasing luminosity distance and a roughly similar flux limit. The evolution of the Lyα luminosity function is presented in Chapter 3.

We note that extensive analysis of the SC4K public sample have already been conducted in previous works. For example, Paulino-Afonso et al. (2018) studied the UV morphologies of the sample and found that UV sizes of LAEs are constant from $z \sim 2$ to $z \sim 6$ with effective radii sizes of $r_e \sim 1.0 \pm 0.1$ kpc. Shibuya et al. (2019) analysed the radial surface brightness profiles of ~ 9000 LAEs (including SC4K) and found that LAEs typically have small sizes, similar to those presented by Paulino-Afonso et al. (2018). This means SC4K LAEs are unresolved in the continuum in ground-based data. Khostovan et al. (2019) derived clustering properties of the sample and measured typical halo masses of $\sim 10^{11}\,M_\odot$ in NB-selected LAEs and $\sim 10^{11} - 10^{12}\,M_\odot$ in MB-selected LAEs, showing the clustering and typical dark matter halo masses that host LAEs is strongly dependent on $L_{Ly\alpha}$. They find more luminous LAEs reside in more massive dark matter haloes. Calhau et al. (2020) study the X-ray and radio properties of the SC4K sample, estimating black hole accretion rates which can reach $\sim 3\,M_\odot\,yr^{-1}$ in the most extreme sources. They also find that the overall AGN fraction of LAEs is low ($<$ 10%) but dependent on $L_{Ly\alpha}$, significantly increasing with increasing luminosity and approaching 100% at $L_{Ly\alpha} > 10^{44}$ erg s^{-1}.

5.2.1 X-ray and radio AGN in SC4K

In total we have 3908 LAEs in our sample, with 254 detected in X-ray and 120 detected in radio (56 in both), resulting in 318 AGN candidates (Calhau et al., 2020). LAEs which are detected in the X-ray and/or radio are classified as AGN as star-forming processes would require SFR $\gtrsim 1000\,M_\odot\,yr^{-1}$ to be detected
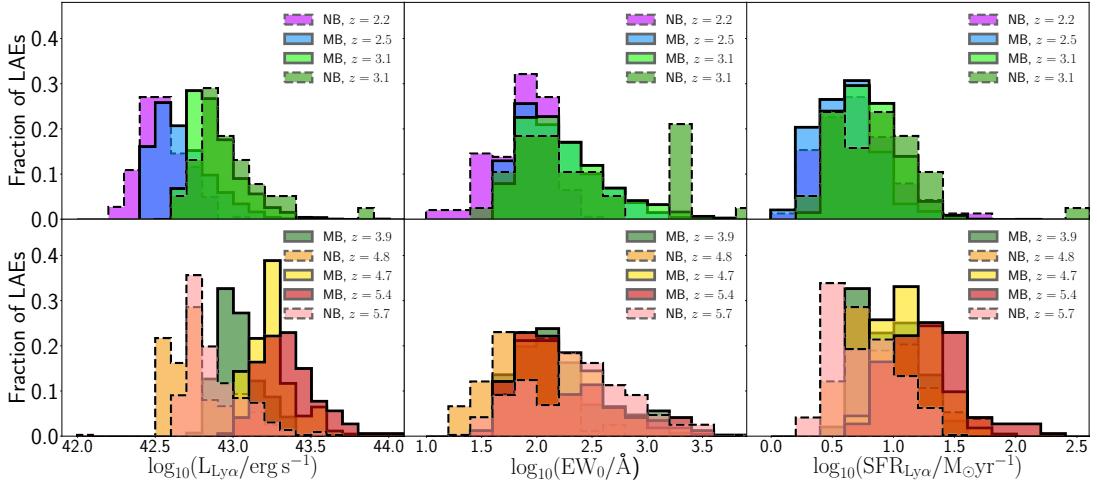
Figure 5.1: Distributions of parameters derived directly from photometry. Lyα luminosity (left panel), EW$_0$ (middle panel) and SFR derived directly from L$_{Ly\alpha}$ and EW$_0$ (Sobral & Matthee, 2019, see Section 6.1.6.1; right panel). MB (NB) data are shown as filled (dashed) lines. For each parameter, top panels show the $z \leq 3.1$ sample and bottom panels show the higher redshift LAEs. The EW$_0$ peak at $z = 3.1$ (NB) is artificial and it is the upper limit of the EW$_0$, obtained from the flux upper limit. AGN have been removed.

above the flux limit at such wavelengths and redshifts (see discussion in Calhau et al., 2020). The number of AGNs reported in this work constitutes an extra 177 sources compared to the ones originally reported in Sobral et al. (2018a), with the additional sources being identified by reaching lower S/N with deep *Chandra* data (COSMOS Chandra Legacy, Civano et al., 2016) and VLA radio data at 1.4 GHz (VLA-COSMOS Survey, Bondi et al., 2008; Schinnerer et al., 2004, 2007, 2010) and by including 3 GHz radio data (Smolčić et al., 2017). We note, however, that due to available coverage, Calhau et al. (2020) only probe 3705 SC4K LAEs with X-Ray and radio data. Throughout this work, SC4K AGNs may be shown in figures (clearly highlighted as such) but are removed from any fitting/binning and median values in tables unless stated otherwise as we focus on the properties of the star-forming population. The catalogue that is provided in this work has a flag for sources detected in X-Ray and radio (see Section 6.1.7).

5.2.2 Redshift binning

To improve the S/N in certain redshift ranges and for clearer visualisation of results, we frequently group multiple MB filters in specific redshift bins throughout this work, following the same grouping scheme as Section 3.1.5: $z = 2.5 \pm 0.1$ (IA427), $z = 3.1 \pm 0.4$ (IA464, IA484, IA505, IA527); $z = 3.9 \pm 0.3$ (IA574, IA624); $z = 4.7 \pm 0.2$ (IA679, IA709); $z = 5.4 \pm 0.5$ (IA738, IA767, IA827). We generally study the NBs separately as there are some relevant distinctions between MBs and NBs, most significantly the flux limit and EW_0 cut. Additionally, analysing the two separately provides independent results and allows checks for systematics.

5.3 Multi-wavelength data

We use the extensive archive of publicly available multi-wavelength data in the COSMOS field to conduct accurate photometric measurements in the UV, optical, near-infrared (NIR), mid-infrared (MIR) and FIR wavelengths for each SC4K LAE, individually. A summary of the filters used, effective wavelength, width and limiting magnitude is provided in Table 5.2. We use optical broad band (B, V, g^+, r^+, i^+, z^{++}), medium band (IA427, IA464, IA484, IA505, IA527, IA574, IA624, IA679, IA709, IA738, IA767, IA827) and narrow band (NB711, NB816) data taken with the Subaru/SuprimeCam (Capak et al., 2007; Taniguchi et al., 2007), retrieved from the COSMOS Archive[1]. Additionally, we use the u band from CFHT/MegaCam. We use deep NIR data (Y, J, H, Ks) from UltraVISTA DR4 (McCracken et al., 2012), taken with VISTA/VIRCAM (Sutherland et al., 2015). Data used have a 0.15" pix^{-1} pixel scale and are calibrated to a zero-point of 31.4 mag (30 mag for UltraVISTA and u images). For MIR coverage, we use data from *Spitzer*/IRAC, channels 1 ($3.6\mu m$) and 2 ($4.5\mu m$) from SPLASH (Steinhardt et al., 2014) and channels 3 ($5.6\mu m$) and 4 ($8.0\mu m$) from S-COSMOS (Sanders et al., 2007). IRAC data have a zero-point of 21.5814 mag and a pixel scale of 0.6" pix^{-1} (after drizzling; the native scale is 1.2" pix^{-1}).

---

[1]https://irsa.ipac.caltech.edu/data/COSMOS/images/

For the FIR coverage, we use $100\mu m$ and $160\mu m$ data (PEP, Lutz et al., 2011) taken with *Herschel*/PACS (Pilbratt et al., 2010) and $250\mu m$, $350\mu m$ and $500\mu m$ data (HerMES, Griffin et al., 2010; Oliver et al., 2012) taken with *Herschel*/SPIRE. The five listed FIR images have a pixel scale of 1.2" pix^{-1}, 2.4" pix^{-1}, 6" pix^{-1}, 8.3" pix^{-1} and 12" pix^{-1}, respectively. FIR images are calibrated to provide fluxes in Jansky and thus have a zero-point of 8.9 mag.

5.4 Multi-wavelength photometry

Accurate photometric measurements are essential to obtain robust SEDs and derive accurate galaxy properties, particularly for sources that are faint in the continuum. While there is a plethora of publicly available catalogues for the COSMOS field (e.g. Ilbert et al., 2009; Laigle et al., 2015), such catalogues are typically broad band selected and thus miss a significant number of line-emitters, especially faint, high EW sources. For example, 9% of our LAEs are not detected in the i band-selected catalogue from Ilbert et al. (2009) with 1" radius matching and 29% of SC4K LAEs are not detected in the NIR-selected catalogue from Laigle et al. (2015). Continuum faint sources with very blue UV continuum slopes have low fluxes in the observed optical and will fall below the detection thresholds of NIR selected catalogues (e.g. Laigle et al., 2015), particularly if they have low stellar masses. Therefore, to obtain consistent, controllable and uniform measurements for the entire sample of LAEs, we conduct our own aperture photometry and estimate errors locally using empty apertures. We also compare our photometry with measurements from the COSMOS catalogues and find a very good agreement. Furthermore, because we have measured the sizes in the rest-frame UV and found SC4K LAEs to be very compact (point-like for the data we use; $r_e = 1.0$ kpc corresponds to 0.13" at $z = 3$), we opt to conduct PSF photometry, as fully explained in Section 5.5.

Table 5.2: Overview of the photometric filters used in this work ranked from the lowest to highest wavelengths. (1) Photometric filter; (2) Effective wavelength; (3) Filter FWHM; (4) 3σ magnitude depth measured in a fixed 2" aperture (except for *Herschel* measurements, see Section 5.5.5); (5) Correction term summed to the measured magnitudes to correct for systematic offsets (*includes an additional offset to correct the systematic uncertainties Section 5.5.6; †denotes values obtained from the deblended FIR catalogue presented by Jin et al. 2018); (6) Filter dependent galactic extinction correction that is subtracted from the measured magnitudes; (7) Instrument and telescope used for the observations; (8) Source of the data.

Filter	λ_{eff} (Å)	FWHM (Å)	Depth (3σ, 2")	s_f	A_λ	Instrument, Telescope	Source
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
u	3911.0	538.0	27.8	0.054	0.0878	MegaCam, CFHT	Capak et al. (2007)
$IA427$	4256.3	206.5	27.0	0.037	0.0816	Suprime-Cam, Subaru	Capak et al. (2007)
B	4439.6	806.7	28.3	-0.242	0.0784	Suprime-Cam, Subaru	Capak et al. (2007)
$IA464$	4633.3	218.0	26.9	0.013	0.0750	Suprime-Cam, Subaru	Capak et al. (2007)
g^+	4728.3	1162.9	27.6	0.024	0.0733	Suprime-Cam, Subaru	Capak et al. (2007)
$IA484$	4845.9	228.5	27.0	0.000	0.0713	Suprime-Cam, Subaru	Capak et al. (2007)
$IA505$	5060.7	230.5	26.8	-0.002	0.0678	Suprime-Cam, Subaru	Capak et al. (2007)
$IA527$	5258.9	242.0	27.1	0.026	0.0646	Suprime-Cam, Subaru	Capak et al. (2007)
V	5448.9	934.8	27.6	0.046*	0.0616	Suprime-Cam, Subaru	Capak et al. (2007)
$IA574$	5762.1	271.5	26.8	0.078	0.0570	Suprime-Cam, Subaru	Capak et al. (2007)
$IA624$	6230.0	300.5	26.8	0.002	0.0506	Suprime-Cam, Subaru	Capak et al. (2007)
r^+	6231.8	1348.8	27.7	0.003	0.0506	Suprime-Cam, Subaru	Capak et al. (2007)
IA679	6778.8	336.0	26.7	0.039*	0.0442	Suprime-Cam, Subaru	Capak et al. (2007)
IA709	7070.7	315.5	26.8	-0.024	0.0411	Suprime-Cam, Subaru	Capak et al. (2007)
NB711	7119.6	72.5	25.9	0.014	0.0406	Suprime-Cam, Subaru	Capak et al. (2007)
IA738	7358.7	323.5	26.5	0.017	0.0383	Suprime-Cam, Subaru	Capak et al. (2007)
i^+	7629.1	1489.4	27.2	0.019	0.0360	Suprime-Cam, Subaru	Capak et al. (2007)
IA767	7681.2	364.0	26.5	0.041	0.0356	Suprime-Cam, Subaru	Capak et al. (2007)
NB816	8149.0	119.5	26.6	0.068	0.0320	Suprime-Cam, Subaru	Capak et al. (2007)
IA827	8240.9	343.5	26.5	-0.019	0.0313	Suprime-Cam, Subaru	Capak et al. (2007)
z^{++}	9086.6	955.3	26.8	-0.037	0.0265	Suprime-Cam, Subaru	Capak et al. (2007)
Y	10211.2	930.0	26.2	0.0	0.0211	VIRCAM, VISTA	McCracken et al. (2012) (DR4)
J	12540.9	172.0	25.8	0.0	0.0144	VIRCAM, VISTA	McCracken et al. (2012) (DR4)
H	16463.7	2910	26.1	0.0	0.0088	VIRCAM, VISTA	McCracken et al. (2012) (DR4)
Ks	21487.7	3090	25.8	0.0	0.0053	VIRCAM, VISTA	McCracken et al. (2012) (DR4)
$IRAC1$	35262.5	7412	25.6	0.002	0.0021	IRAC, *Spitzer*	Steinhardt et al. (2014)
$IRAC2$	44606.7	10113	25.5	0.000	0.0014	IRAC, *Spitzer*	Steinhardt et al. (2014)
$IRAC3$	56764.4	13499	22.6	0.013	0.0010	IRAC, *Spitzer*	Sanders et al. (2007)
$IRAC4$	77030.1	28397	22.5	-0.171	0.0007	IRAC, *Spitzer*	Sanders et al. (2007)
$100\mu m$	979036.1	356866	15.4	0.20†	0.0000	PACS, *Herschel*	Lutz et al. (2011)
$160\mu m$	1539451.3	749540	14.3	-0.06†	0.0000	PACS, *Herschel*	Lutz et al. (2011)
$250\mu m$	2471245.1	658930	10.9	-0.49†	0.0000	SPIRE, *Herschel*	Oliver et al. (2012)
$350\mu m$	3467180.4	937200	10.6	-0.15†	0.0000	SPIRE, *Herschel*	Oliver et al. (2012)
$500\mu m$	4961067.7	1848042	10.6	0.03†	0.0000	SPIRE, *Herschel*	Oliver et al. (2012)

5.5 Aperture photometry of SC4K LAEs

5.5.1 Overview of our aperture photometry

In order to obtain accurate aperture photometry for individual LAEs, for each band, we estimate the total magnitude by following the steps:

- conducting photometry in fixed apertures (Section 5.5.2);

- applying aperture corrections based on PSF stars around each LAE (Section 5.5.3);

- applying reddening corrections for galactic extinction (Section 5.5.4);

- introducing systematic offset corrections based on known offsets and COSMOS catalogues (Section 5.5.6);

Magnitudes per source and per band are computed as:

$$\mathrm{mag} = \mathrm{mag}_0 + \mathrm{aper}_{\mathrm{cor}} + s_f - A_\lambda, \tag{5.1}$$

where mag_0 is the magnitude calculated by converting the flux obtained in fixed apertures (typically 2" diameter for most of the data) to the AB magnitude system before any correction is applied, $\mathrm{aper}_{\mathrm{cor}}$ is the aperture correction derived per band and per source, based on PSF stars around each LAE, s_f the systematic offset correction for the filter and A_λ the reddening correction for galactic extinction computed for the effective wavelength of the filter. The error in the final magnitude is obtained by propagating the error in flux, scaling the error with the correction that was applied to the flux and then adding 30% of the total correction to the error in flux[1]. Aperture photometry in the FIR is discussed separately in Section 5.5.5.

---

[1] We note that we use 30% as a conservative approach to add unknown systematic errors.

5.5.2   Aperture photometry in fixed apertures

We conduct aperture photometry centred on the position of each SC4K LAE (Sobral et al., 2018a) over all the filters listed in Table 5.2. We do this by creating 200x200 pixel (30" ×30" for a 0.15" pix^{-1} pixel scale) cutouts, where we conduct the photometry[1]. For optical to MIR images, we use 2" diameter apertures. We estimate the background by placing 2000 2" apertures in random positions of the field where there are no detections above $2\,\sigma$ (given by the segmentation maps per filter produced by SExtractor; Bertin & Arnouts 1996) and subtract it from the counts of the aperture placed on the LAE. Upper and lower errors are measured as the 84th and 16th percentiles of all random apertures. We repeat this procedure per band per source.

5.5.3   Aperture correction

The original point-spread function (PSF) was kept across all images as we have opted for correcting the photometry with PSF stars, instead of PSF matching the data, in order to avoid modifying the data and confuse nearby sources. Fixed aperture photometry in non-PSF matched images requires correction of the PSF effect on photometry so we can obtain total fluxes and total magnitudes for point-like sources. To do this, we measure the magnitude of stars[2] in 2" apertures and with MAG_AUTO (Bertin & Arnouts, 1996)[3]. We define the correction factor (aper$_{cor}$ in Equation 5.1) as the difference between MAG_AUTO and magnitudes measured in 2" apertures. This correction is valid for point-like sources, an assumption that should be valid for our LAEs given the rest-frame UV sizes as measured by Paulino-Afonso et al. (2018) using high-resolution *HST*/ACS images. The correction term is measured for each filter, and it is the median correction of stars within a 0.3 degree radius around each LAE, accounting for spatial variations of the PSF per band. The aperture corrections are typically small, averaging ~ -0.1 to ~ -0.3.

---

[1]We use PSF stars beyond this region.

[2]Selected from Ilbert et al. (2009): photoz=0.0; stellaricity=1; detected in the point-source catalogue 2MASS Skrutskie et al. (2006); visually checked to remove binary systems or close projections.

[3]aper$_{cor}$ = MAG_AUTO−mag$_0$, where mag$_0$ is the magnitude before corrections are applied.

5.5.4 Galactic extinction correction

We correct for dust attenuation along the line-of-sight due to our Galaxy. For the COSMOS field, the median galactic extinction is measured to be $E(B-V) = 0.0195 \pm 0.006$ (Capak et al., 2007). The slope of the extinction curve with wavelength is parametrised by the factor $R(V)$:

$$R(V) \equiv \frac{A(V)}{E(B-V)}, \tag{5.2}$$

where $A(V)$ is the total extinction at the V band. For the diffuse interstellar medium, the median value of $R(V)$ is estimated to be 3.1 (e.g. Fitzpatrick, 1999) and it is the value used in this work. We use the model from Fitzpatrick & Massa (2007) where the attenuation at a wavelength (λ) becomes:

$$A_\lambda = A(V) \left(1 + \frac{k}{R(V)} \right), \tag{5.3}$$

where k is a polynomial expansion of λ^{-1} (Equation 2 from Fitzpatrick & Massa 2007) with a linear component for UV wavelengths, a curvature term for the far-UV and a Lorentzian-like bump at 2175 Å. We determine A_λ for the effective wavelength of each filter and show its value for each filter in Table 5.2.

5.5.5 FIR photometry

For FIR data, due to the large PSF of 7.2", 12", 18.15", 25.15" and 36.30" ($100\mu m$, $160\mu m$, $250\mu m$, $350\mu m$ and $500\mu m$, respectively), the usage of 2" diameter aperture photometry is not viable. We conduct aperture photometry using apertures which are the size of the PSF: radius of 6, 5, 3, 3 and 3 pixels (7.2", 12", 18", 24.9" and 36"), respectively (retrieving 67% of the total flux), with the same random empty aperture procedure to estimate background. This allows us to then apply aperture corrections of 1/0.67 to get full fluxes for point-like sources. For $100\mu m$ ($160\mu m$), we multiply the flux by the filter correction factor 1.1 (1.2) as described in the PEP public data release notes (see Lutz et al., 2011).
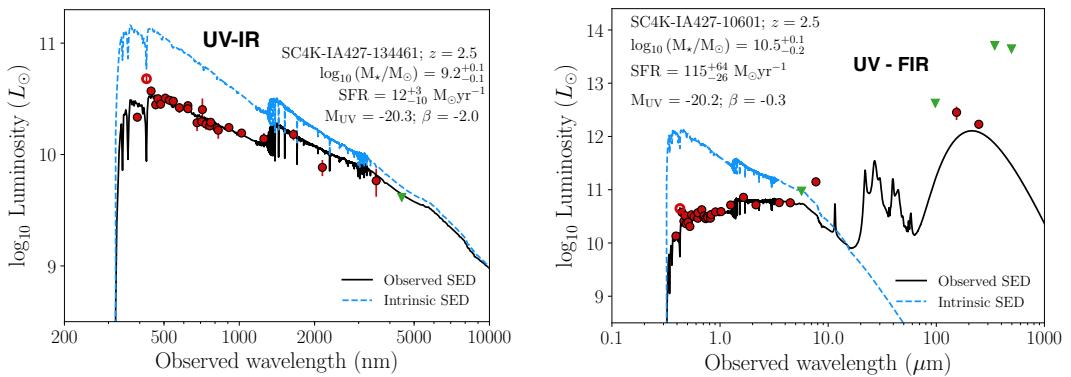
Figure 5.2: *Left:* SED of SC4K-IA427-134461 (at $z = 2.5$), for observed UV-IR wavelengths as we only obtain upper limits in the FIR. Red circles show the luminosity (in solar units) measured at the corresponding observed wavelength and green arrows show the upper limits for non-detections, where the flux is $< 3\,\sigma$. Unfilled circles are the luminosity at the NB/MB where the LAE was selected, and we note that this filter was not used to derive the SED fit. The black line is the best-fit SED to the observed photometry and the blue dashed line the intrinsic (dust-free) SED. This is an example of a very blue ($\beta = -2.0$) and low stellar mass ($M_\star = 10^{9.2}\,M_\odot$) LAE. *Right:* Same as left panel but for SC4K-IA427-10601 (at $z = 2.5$) and at a wider wavelength range, showing FIR wavelengths as this LAE is detected in $250\mu m$ and $350\mu m$ due to the presence of dust. This LAE is redder ($\beta = -0.3$) and more massive ($M_\star = 10^{10.5}\,M_\odot$). Note that this LAE is not representative of the SC4K sample as only $\sim 3\%$ (2%) non-AGN LAEs are as massive (as red).

However, the blending of sources is still a serious issue, as the large pixel scale makes it difficult to establish if a detection is produced by one of our LAEs or by a neighbouring source. To solve this, we use the FIR measurements from the publicly available deblended COSMOS catalogue (Jin et al., 2018), where FIR emission is deblended to match optical-NIR coordinates. With a 1" match to the deblended catalogue, there are 14, 11, 29, 19 and 12 SC4K LAEs with 3σ detections in $100\mu m$, $160\mu m$, $250\mu m$, $350\mu m$ and $500\mu m$, respectively. Whenever a source is undetected in the FIR, we assign the local estimate of the background as an upper limit, which we measure with 2000 empty apertures the size of the PSF. We ensure our own flux measurements are consistent with Jin et al. (2018) (see Section 5.5.6).

5.5.6 Systematic offsets

We correct for systematic offsets (s_f) in the photometry by applying the corrections derived by Ilbert et al. (2009) (we present these values in Table 5.2). After applying the systematic offsets and all previous correction terms, we compare our total magnitudes with measurements from Ilbert et al. (2009) and Laigle et al. (2015). We find no statistically significant difference with our measurements except for two filters (IA679, V) which have systematic offsets of ~ 0.5 mag. We apply a further correction (included in the s_f, Table 5.2) to our magnitudes, so the median of the magnitude difference becomes zero. For FIR magnitudes, we estimate the systematic correction term from the FIR deblended catalogue (Jin et al., 2018), also presented in Table 5.2.

5.6 Spectral Energy Distributions of SC4K LAEs

Having conducted photometry in the 34 filters listed in Table 5.2, we can now explore the SED of each individual LAE, observed from UV to FIR. We use the publicly available SED-fitting code MAGPHYS[1] (da Cunha et al., 2008, 2012) with the high-redshift extension (see da Cunha et al., 2015), to obtain SED

---

[1]http://www.iap.fr/magphys/

fits for each individual galaxy, using our rest-frame UV, optical and NIR-FIR photometric measurements. If a LAE is not detected in a specific filter above the estimated background noise, we define it as an upper limit.

MAGPHYS is based on dust attenuation models from Charlot & Fall (2000) and uses the stellar population synthesis model from Bruzual & Charlot (2003) with a Chabrier (2003) IMF (range $0.1 - 100$ M$_\odot$) to compute the emission of simple stellar populations (SSPs, populations of coeval stars with similar properties). We use the prescription of Madau (1995) to model the Intergalactic Medium (IGM). The software generates a library of model SEDs for galaxies at the mean redshift of the NB/MB filter (see Table 5.1) and for the given photometric bands. The modelled SED of a galaxy is composed by the weighted sum of SSPs, with the star formation history (SFH) being a continuously delayed exponential function with an early rise followed by a decay. Instantaneous bursts of star formation of random duration (lasting 30-300 Myr) and amplitude (forming mass between 0.1-100 times the mass formed by the continuous SFH) are superimposed. A Bayesian approach is then used to compare model SEDs with observed photometry, creating a parameter likelihood distribution for several galaxy properties such as stellar mass, SFR and dust attenuation.

As the models are purely stellar (no nebular line fitting), we do not fit photometry from filters where we expect strong nebular emission, namely Lyα at the selection NB or MB filters, as it is by definition significant in our Lyα-selected sample. While we do not remove photometry from filters which may have contribution from other emission lines such as Hα (IRAC filters at $z \sim 4 - 6$) or [O III] (H-K bands at $z \sim 2 - 3$), by removing the Lyα-contaminated filter, combined with the large number of filters used, we do not expect an overestimation of masses due to nebular line contamination. We explore this by rerunning MAGPHYS for the entire $z = 2.5$ sample (IA427) after removing the H and K bands, which may be contaminated from [O III] and Hα emission, respectively, and compare the difference of estimated stellar masses. We find that when removing both H and K, the median difference of stellar masses is $\log_{10}(\mathrm{M}_{\star,\mathrm{no\,HK}}/\mathrm{M}_\odot) - \log_{10}(\mathrm{M}_\star/\mathrm{M}_\odot) = 0$, with no dependence on mass, and the average difference -0.07. Removing H and K makes the estimation of stellar mass more uncertain as the rest-frame optical

becomes more poorly constrained. Additionally, we test the effect of only removing H, with the IR still being constrained by the other bands. We also find no significant difference in stellar masses, with the median of the difference being $\log_{10}(M_{\star,no\,H}/M_{\odot}) - \log_{10}(M_{\star}/M_{\odot}) = 0$ and the average -0.08. Overall, we find that not removing photometric bands outside Lyα does not lead to a significant overestimation of stellar masses for our sources. However, including nebular lines may still be important, particularly if we look at other parameters (e.g. ages), as there may be some systematics, particularly for the faintest sources with the highest EWs. This will be addressed in a forthcoming work with an SED-fitting code that models nebular emission (CIGALE, Boquien et al., 2019; Noll et al., 2009).

For our $z \sim 2-6$ LAEs, the optical bands are essential to fit the rest-frame UV continuum, IRAC filters can constrain fluxes redward of D_{4000} and the FIR measurements provide upper constraints in the dust emission, which can improve the SFR estimates. We note that, as explained in Section 5.2.1, while we exclude sources with evidence of AGN activity when computing median properties of the sample, we still obtain SED fits (without using any AGN SED model) for those sources.

In Figure 5.2, we show observed and intrinsic SED fits and photometric measurements/upper limits for two LAEs. The SEDs were purposely chosen to show two very distinct galaxies within the SC4K sample: one with a very blue and steep UV continuum slope, with low stellar mass that dominate the sample and one with a more red continuum, more massive and with higher dust extinction which is much more rare in the sample of LAEs. While the latter is not well representative of a typical LAE, it is still important to show that LAEs can span a large variety of physical properties. This LAE is detected in two *Herschel* bands, which shows that FIR can be important to constrain the SED fits and derive properties of high redshift LAEs.

5.6.1 Number of derived SEDs

Although all LAEs are by definition detected in the MB/NB where they were selected (Chapter 2), a small fraction of our LAEs have few to no detections

in other photometric bands. For such cases, SED-fitting may fail. Out of the 3590 non-AGN LAEs, we obtain reliable SEDs for 3377 (94%, see Table 5.1). The catalogue that we release with this work (see Section 6.1.7) has an SED flag which marks unreliable SEDs. The AGN flag indicates AGN LAEs (Section 5.2.1, Calhau et al. 2020), and we reiterate that while we compute parameters for these sources, the SED-derived parameters are not reliable and are not included in any median property estimation done in this work.

5.7 Summary

We conducted PSF photometry over 34 bands from rest-frame UV to FIR and derived the best-fit SEDs using MAGPHYS for each source in the large SC4K sample of ~ 4000 $z \sim 2 - 6$ LAEs. Out of the 3590 non-AGN LAEs, we obtain reliable SEDs for 3377 (94%, see Table 5.1). The evolution of derived properties such as M_{UV}, Lyα EWs and the SFR-stellar mass relation will be discussed in the following chapter.

Chapter 6

The evolution of rest-frame UV properties, Lyα EWs, and the SFR-stellar mass relation at $z \sim 2 - 6$ for SC4K LAEs

Abstract

For our sample of $z \sim 2 - 6$ LAEs, we find typical stellar masses of $10^{9.3\pm0.6} \, \mathrm{M}_\odot$ and star formation rates (SFR) of $\mathrm{SFR}_\mathrm{SED} = 4.4^{+10.5}_{-2.4} \, \mathrm{M}_\odot$ yr^{-1} and $\mathrm{SFR}_\mathrm{Ly\alpha} = 5.9^{+6.3}_{-2.6} \, \mathrm{M}_\odot \, \mathrm{yr}^{-1}$, combined with very blue UV slopes of $\beta = -2.1^{+0.5}_{-0.4}$, but with significant variations within the population. M_UV and β are correlated in a similar way to UV-selected sources, but LAEs are consistently bluer. This suggests that LAEs are the youngest and/or most dust-poor subset of the UV-selected population. We also study the Lyα rest-frame equivalent width (EW_0) and find 45 "extreme" LAEs with $\mathrm{EW}_0 > 240 \, \text{Å} \, (3 \, \sigma)$, implying a low number density of $(7 \pm 1) \times 10^{-7} \, \mathrm{Mpc}^{-3}$. Overall, we measure little to no evolution of the Lyα EW_0 and scale length parameter (w_0) which are consistently high ($\mathrm{EW}_0 = 140^{+280}_{-70} \, \text{Å}$, $w_0 = 129^{+11}_{-11} \, \text{Å}$) from $z \sim 6$ to $z \sim 2$ and below. However, w_0 is anti-correlated with M_UV and stellar mass. Our results imply that sources selected as LAEs have a high Lyα escape fraction ($\mathrm{f}_\mathrm{esc,Ly\alpha}$) irrespective of cosmic time, but $\mathrm{f}_\mathrm{esc,Ly\alpha}$ is still higher for UV-fainter and lower mass LAEs. The least massive LAEs ($< 10^{9.5} \, \mathrm{M}_\odot$) are typically located above the star formation "Main Sequence" (MS), but the offset from the MS decreases towards $z \sim 6$ and towards $10^{10} \, \mathrm{M}_\odot$. Our results imply a lack of evolution in the properties of LAEs across time and reveals the increasing overlap in properties of LAEs and UV-continuum selected galaxies as typical star-forming galaxies at high redshift effectively become LAEs.

We structure this Chapter as follows: we present the properties of LAEs in Section 6.1, where we show the methodology we use to derive EW_0, SFR, M_{UV} and β. We present our results in Section 6.2, looking into the M_{UV}-β and SFR-$M_\star$ relations and the potential evolution of EW_0 with redshift, along with physical interpretations. Finally, we present our conclusions in Section 6.3.

6.1 The properties of LAEs

In this section, we present our methodology and computations to derive galaxy properties for individual LAEs, using our full photometric measurements and SED fits from MAGPHYS. EW_0 and $L_{Ly\alpha}$ of all LAEs in the SC4K sample have been derived and published in Sobral et al. (2018a) (shown in Chapters 2 and 3).

6.1.1 Lyα luminosity ($L_{Ly\alpha}$)

$L_{Ly\alpha}$ is calculated from the Lyα line flux ($f_{Ly\alpha}$):

$$L_{Ly\alpha}[\mathrm{erg\,s^{-1}}] = 4\pi f_{Ly\alpha} D_L^2(z) \tag{6.1}$$

where $D_L(z)$ is the luminosity distance at the redshift of each source, computed from the redshifted Lyα at the effective wavelength of the detection NB/MB. In Figure 5.1 (left) we show the $L_{Ly\alpha}$ distribution of our LAEs, spanning a wide range of luminosities $L_{Ly\alpha} = 10^{42-44}$ erg s^{-1}.

6.1.2 Lyα rest-frame equivalent width (EW_0)

The observed EW (EW_{obs}) of an emission line is the ratio between the flux of the line and the continuum flux density and can be calculated as:

$$EW_{obs}[\mathring{A}] = \Delta\lambda_1 \frac{f_1 - f_2}{f_2 - f_2(\Delta\lambda_1/\Delta\lambda_2)}, \tag{6.2}$$

where $\Delta\lambda_1$ is the FWHM of the NB or MB, $\Delta\lambda_2$ the FWHM of the broad band filter (Table 2.3), f_1 is the flux density measured in the NB or MB and f_2 is the

flux density computed from two adjacent BB filters, which avoids assumptions of the slope of the continuum (for details see Sobral et al., 2018a). The rest-frame EW (EW_0) is calculated as:

$$EW_0[\text{Å}] = \frac{EW_{\text{obs}}}{1 + z}, \tag{6.3}$$

where z is the redshift of Lyα at the effective wavelength of the NB or MB (Sobral et al., 2018a). We provide the median EW_0 for different redshifts and for the full SC4K sample in Table 5.1.

6.1.2.1 EW$_0$ scale length (w_0)

An exponential fit of the form $N = N_0 \exp(\text{-}EW_0/w_0)$ has been widely used to describe Lyα EW$_0$ distributions (e.g. Gronwall et al., 2007; Hashimoto et al., 2017; Wold et al., 2017), with the rate of decay being determined by the scale length parameter w_0. With our sample of LAEs, we analyse EW$_0$ distributions in multiple well defined redshift ranges between $z \sim 2$ and $z \sim 6$. To estimate w_0, we define bins of 20 Å and fit the exponential function to the observed distribution (see Figure 6.1), taking into account Poissonian errors. Bins with less than two sources are excluded from the fits. To account for bin width choice, we add 10 Å (half the bin width) in quadrature to the errors of w_0. We also explore how an EW$_0$ upper cut affects w_0 as it removes sources with extreme (and more uncertain) EWs. We apply a cut of $EW_0 = 240$ Å, the theoretical limit of EW$_0$ powered by Population II star formation (e.g. Charlot & Fall, 1993) and the value which has been extensively used in Lyα emission studies to identify "extreme" EW galaxies (e.g. Cantalupo et al., 2012; Marino et al., 2018). We compute χ^2_{red} by comparing the best exponential fit to the histogram of observed counts and their associated Poisson errors.

Additionally, we fully explore how the errors on EW$_0$ influence the measurement of w_0 by using an MCMC approach. For each iteration, we perturb the EW$_0$ of each LAE in that specific sample within their asymmetric error bars (assuming a double normal probability distribution function centred at each EW$_0$ and with FWHM equal to the errors derived from photometry; Chapter 2). We
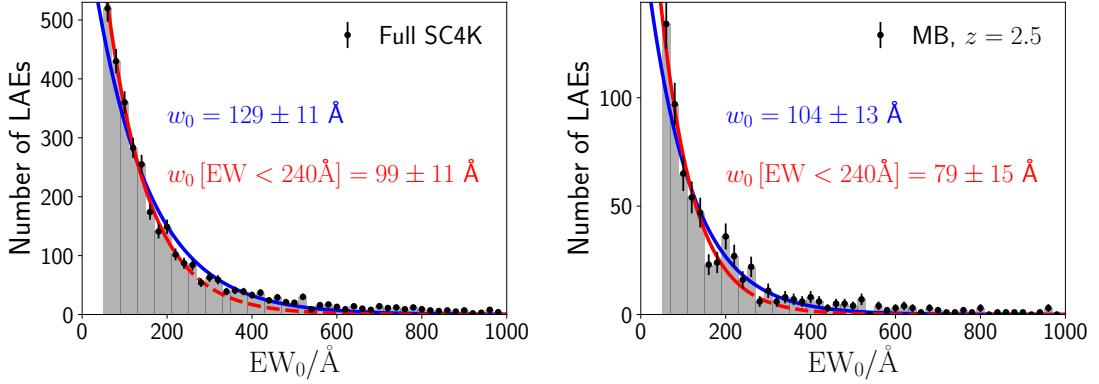
Figure 6.1: *Left:* EW_0 distribution of the full SC4K sample of LAEs. We fit an exponential function of the form $N = N_0 \exp(-EW_0/w_0)$, and derive the parameter w_0. Fit derived with the distribution of EW_0 ($EW_0 < 240\,\text{Å}$) is shown in red (blue). *Right:* Same but for an individual filter (IA427) with LAEs at $z = 2.5$.

impose a hard lower limit equal to the detection threshold (50 Å for MBs, 25 Å for NBs except for NB392 which has a lower limit of 5 Å; see Chapter 5) and an upper limit of 1000 Å, with any source outside these values not being included in a specific realisation. With the perturbed EW_0, we construct the histogram of the current iteration, using bins of 20 Å. We fit an exponential to the generated histogram bins, taking into account the associated Poissonian error ($\sqrt{N}$) of each bin. We iterate this process 200 times, and the final w_0 is the median value of all fits with error up (down) being the 84th (16th) percentile of all fits. In addition, to account for the uncertainty introduced by the bin width choice, we also add 10 Å in quadrature to the errors of w_0. We also apply the MCMC approach with a cut of $EW_0 = 240\,\text{Å}$. For the MCMC approach, where EW_0 are perturbed, χ^2_{red} is computed by comparing the best fit to the median histogram of all iterations and its Poisson errors.

In Table A.6, we show the inferred w_0 values (including perturbed estimates) for different redshift ranges and filter combinations.

Furthermore, it is important to establish how the EW_0 distribution depends on M_{UV} and $M_\star$. To understand this dependence, we measure w_0 in three M_{UV} and $M_\star$ ranges and show our measurements in Table A.6. For the faintest and the lowest mass ranges, we are significantly incomplete to the low EW_0 end of the

EW distribution, resulting in a peak at $\sim 100\,\text{Å}$. Thus, we only fit $EW_0 > 100\,\text{Å}$ to accurately estimate the exponential decay of the distribution for these two cases.

6.1.3 Rest-frame UV luminosity (M_{UV})

The UV luminosity of a galaxy is associated with continuum emission from massive stars and traces SFR in the past 100 Myr (e.g. Boselli et al., 2001; Salim et al., 2009). A priori, sources selected by their strong Lyα emission could be expected to have strong M_{UV} as both trace recent star formation (neglecting AGN contribution), although Lyα can trace slightly more recent star formation because stars dominating the ionising photon budget have lifetimes of ~ 10 Myr. However, as shown by e.g. Matthee et al. (2017c) and Sobral et al. (2018a) more factors come into play as Lyα and M_{UV} do not necessarily correlate with each other, due to e.g. highly ISM dependent $f_{esc,Ly\alpha}$ (which can result in most Lyα emission being absorbed by dust particles or scattered off neutral hydrogen) or an ionising efficiency which is evolving with redshift.

We compute M_{UV} by integrating the best-fit SEDs at rest-frame $\lambda_0 = 1400 - 1600\,\text{Å}$. We show the M_{UV} histogram distribution in Figure 6.2 (centre). Due to the magnitude limits, at higher redshift we are only sensitive to more luminous M_{UV} sources. We detect SC4K LAEs as bright as $M_{UV} = -23$ and as faint as $M_{UV} = -17$.

6.1.4 UV continuum slope (β)

The slope of the UV continuum can be parametrised in the form $f_\lambda \propto \lambda^\beta$ (e.g. Meurer et al., 1999). The slope β is sensitive to the age, metallicity and dust content of a galaxy. Bruzual & Charlot (2003) models used by MAGPHYS have a hard limit to how negative (blue) β can be ($\beta = -2.44$), a natural consequence of an upper limit in the IMF. While β may be intrinsically even bluer for more "extreme" stellar populations, in this study, we do not explore those.

We measure β directly from the best-fit as the slope of the continuum at rest-frame $\lambda_0 = 1300 - 2100\,\text{Å}$. We apply a conservative approach and only use β
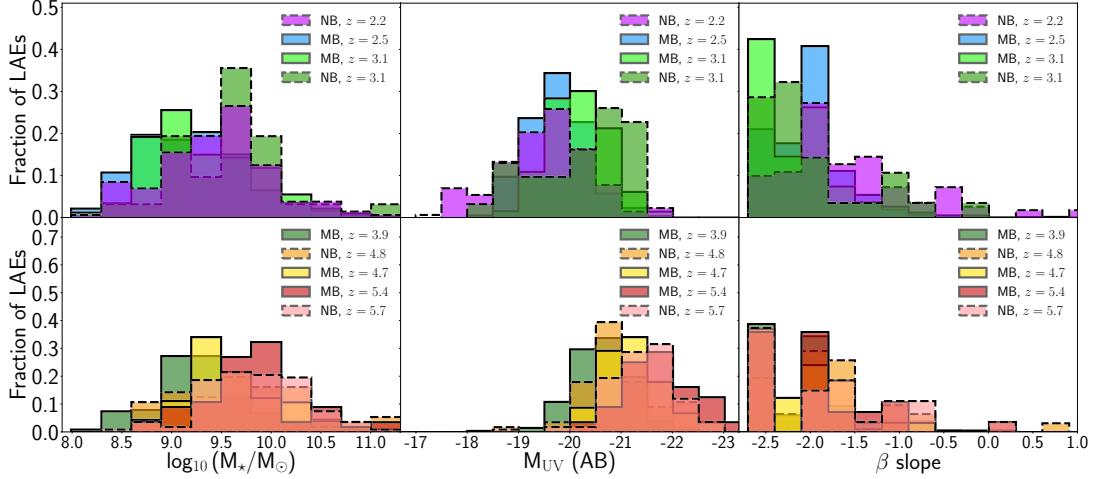
Figure 6.2: Distribution of properties derived from the SED fitting (MAGPHYS, see Section 5.6). We show the stellar mass, $M_\star$ (left), rest-frame UV luminosity, M_{UV} (middle) and rest-frame UV slope, β (right). Top panels show the $z \leq 3.1$ sample and the bottom panels show the higher redshift LAEs. AGNs have been removed.

measurements from sources with at least two detections in this wavelength range. This ensures the β slope is directly constrained and not a direct consequence of assumed SED templates. As expected, due to an increasing luminosity distance, combined with rest-frame $\lambda_0 = 1300 - 2100\,\text{Å}$ moving into IR wavelengths, there are fewer β measurements at higher redshift. In addition, we also compute β by fitting a power-law (β_{pl}) to the photometric measurements (similar to e.g. Bouwens et al., 2014a), with no SED fitting assumptions. We fit β_{pl} in the range $\lambda_0 = 1400 - 2100\,\text{Å}$, which is smaller than the range used to compute β from the SED fit to avoid broad band filters at $\sim 1300\,\text{Å}$, which can be contaminated by the Lyα break. Only sources with at least three 3σ detections in that range are considered for the power-law measurement. For the full SC4K, we measure a median $\beta_{pl} = -1.8^{+0.8}_{-0.7}$, which is redder ($+0.3$) than β from the SED fit (0.2 when only considering sources with β_{pl} measurements), but still within the error bars. Overall, β is better constrained through SED fitting as it uses a prior - the SED models included in MAGPHYS. Furthermore, the SED models take into account ~ 30 filters over the full UV-FIR wavelength range, preventing it from being as sensitive to individual filter measurements in the smaller $\lambda_0 =$

125

$1400 - 2100$ Å range. Thus, throughout this work, we use β computed from the SED fits.

We show the histogram distributions of β in Figure 6.2 (right). LAEs tend to be very blue across all redshift ranges (median $\beta = -2.1^{+0.5}_{-0.4}$, Table 5.1). LAEs at $z = 2.2$ are found to have the reddest β slopes, albeit still very blue and comparable to the Lyman Break Galaxy (LBG) population (see further discussion in 6.2.1). We note, nonetheless, that the $z = 2.2$ sample has some key differences compared to other LAEs in SC4K sample, as it selects LAEs down to 5 Å EW_0 in addition to reaching the faintest $L_{Ly\alpha}$. This allows redder sources to be picked up, while the much higher EW_0 LAEs tend to have much bluer β slopes.

6.1.5 Stellar Mass ($M_\star$)

The total mass of stars in a galaxy (stellar mass, $M_\star$) is a fundamental galaxy property which is a reflection of its star formation history. We use $M_\star$ derived from the likelihood parameter distribution from MAGPHYS modelling.

We show the histogram distribution of $M_\star$ in our sample in Figure 6.2 (left). Most LAEs (88%) have stellar masses $< 10^{10}\,M_\odot$, although it is important to stress there are some more massive galaxies, which shows a significant diversity. We observe a slight shift to higher masses as we move to higher redshifts (see also Table 5.1) but this is a natural consequence of only being sensitive to intrinsically more luminous galaxies at higher redshift. We find that typical LAEs are low stellar mass galaxies, with the median of the SC4K sample of LAEs being $M_\star = 10^{9.3^{+0.6}_{-0.5}}\,M_\odot$.

6.1.6 Star Formation Rates (SFRs)

6.1.6.1 Emission line-based SFRs with Lyα

We estimate the SFR directly from $L_{Ly\alpha}$ and EW_0, using the recipe from Sobral & Matthee (2019) which has calibrated EW_0 as a good empirical indicator of $f_{esc,Ly\alpha}$ (see Section 1.3.3). With a measurement of $f_{esc,Ly\alpha}$, $L_{Ly\alpha}$ can be converted to dust-corrected Hα luminosity assuming case-B recombination (Brocklehurst,
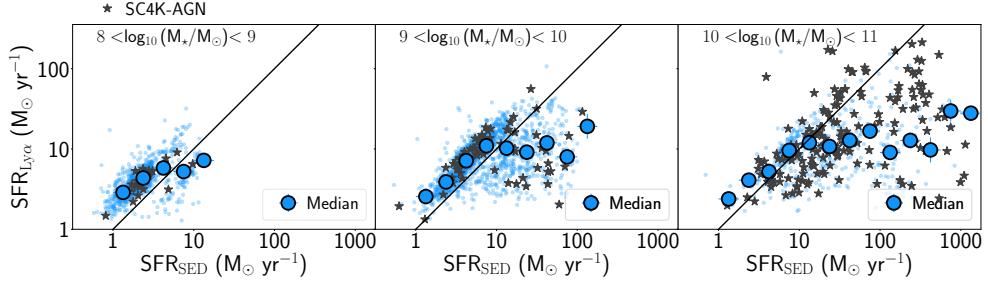
Figure 6.3: Emission line-based SFR vs SED-fitting SFR for the full sample of LAEs at different stellar masses. Blue circles are the median bin (excluding AGN) and individual points are plotted as scatter in the background. The black line is the 1-to-1 ratio. There is a small systematic offset at $M_\star = 10^{8-9}\,M_\odot\,yr^{-1}$ and for $SFR_{SED} = 1 - 10\,M_\odot\,yr^{-1}$ for all stellar mass ranges. For higher stellar masses and SFRs, there is a more significant difference between the two methods, with the emission line-based approach predicting lower SFRs. This is a likely consequence of Lyα not being sensitive to obscured regions in very massive galaxies, thus not being sensitive to their full contribution. Additionally, we plot AGN LAEs with black stars (purely stellar+dust SED-fitting with no AGN models) to show they are typically measured as having high stellar masses when blindly running SED codes with no AGN models in AGN samples.

1971) and transformed into SFR following Kennicutt (1998). For a Chabrier IMF $(0.1 - 100\,M_\odot)$ and assuming $f_{esc,LyC} = 0$, $L_{Ly\alpha}$ in erg s^{-1} and EW_0 in Å, the SFR thus becomes Sobral & Matthee (2019):

$$SFR_{Ly\alpha}\,[M_\odot\,yr^{-1}] = \frac{L_{Ly\alpha} \times 4.4 \times 10^{-42}}{0.042\,EW_0}, \qquad (6.4)$$

For $EW_0 > 210\,$Å, following Sobral & Matthee (2019), we set $f_{esc,Ly\alpha} = 1$ which corresponds to $SFR\,[M_\odot\,yr^{-1}] = 4.98 \times 10^{-43} \times L_{Ly\alpha}$, with $L_{Ly\alpha}$ in erg s^{-1}. This SFR is calibrated with dust-corrected Hα luminosities and thus should be interpreted as dust-corrected SFR. We show the SFR distribution in Figure 5.1 (right). As the SFR is derived from $L_{Ly\alpha}$, it is limited by the same detection limits, which causes a shift to higher SFR with increasing redshift. We measure SFRs in the range $\sim 1 - 300\,M_\odot\,yr^{-1}$, and measure a median $SFR_{Ly\alpha} = 5.9^{+6.3}_{-2.6}$ for SC4K LAEs (see Table 5.1).

6.1.6.2 SED-derived SFRs

As previously stated, MAGPHYS uses a bayesian approach to estimate the best likelihood SFR, comparing model SEDs (generated using some assumptions, see Section 5.6) with observed photometry. Due to our FIR measurements being mostly upper limits for $> 99\%$ of SC4K LAEs, it is not possible to directly measure the amount of SFR that is obscured by dust and the optical thickness of dust from IR-FIR. As such, the amount of dust and SFR is inferred from the UV-optical slope. We measure SFRs in the range $\sim 0.1 - 3000\,\mathrm{M_\odot\,yr^{-1}}$, and measure a median $\mathrm{SFR_{SED}} = 4.4^{+10.5}_{-2.4}\,\mathrm{M_\odot\,yr^{-1}}$ for SC4K LAEs (Table 5.1). For the small subsample of 46 FIR-detected LAEs, Calhau et al. (2020) obtained stacks in the FIR which reveal an average (median) SFR of $340^{+290}_{-260}\,\mathrm{M_\odot yr^{-1}}$ ($200^{+430}_{-110}\,\mathrm{M_\odot yr^{-1}}$).

6.1.6.3 $\mathrm{SFR_{Ly\alpha}}$ vs $\mathrm{SFR_{SED}}$

In this work, we estimate SFRs of individual LAEs using two approaches: emission line-based with Lyα ($\mathrm{SFR_{Ly\alpha}}$, Section 6.1.6.1) and from SED-fitting ($\mathrm{SFR_{SED}}$, Section 6.1.6.2). These two approaches are independent as $\mathrm{SFR_{Ly\alpha}}$ is derived directly from two properties of the Lyα emission-line (luminosity and $\mathrm{EW_0}$), while $\mathrm{SFR_{SED}}$ is obtained with MAGPHYS by removing the filter contaminated by Lyα and using up to ≈ 30 photometric data-points from the rest-frame UV to the rest-frame FIR.

In Figure 6.3 we show a comparison between $\mathrm{SFR_{Ly\alpha}}$ and $\mathrm{SFR_{SED}}$ at different mass ranges. We measure a small systematic offset at $\mathrm{M_\star} = 10^{8-9}\,\mathrm{M_\odot}$ and $\mathrm{SFR_{SED}} = 1 - 10\,\mathrm{M_\odot\,yr^{-1}}$ for all stellar mass ranges, with the emission line-based approach predicting slightly higher SFRs. As Lyα traces more recent star-formation than the UV-continuum, the higher predicted SFRs could be explained by on-going bursts of star-formation, which lead to slightly higher $\mathrm{SFR_{Ly\alpha}}$. Only for SFRs which are measured to be high from SED ($\mathrm{SFR_{SED}} > 10\,\mathrm{M_\odot\,yr^{-1}}$) there is a significant difference, with $\mathrm{SFR_{Ly\alpha}}$ being lower and its median maxing at $\approx 10\,\mathrm{M_\odot\,yr^{-1}}$. Such SFR ranges are typically only seen in more massive ranges ($\mathrm{M_\star} = 10^{9-11}\,\mathrm{M_\odot}$), which are thus more susceptible to have underestimated SFRs from Lyα. This is in line with what could be expected for very massive galaxies as Lyα will only be able to measure the contribution in regions of the galaxy

which are actively star-forming and unobscured, leading to underestimated SFRs in these regimes. Nevertheless, it is remarkable that two largely independent methods obtain such similar results. For the global populations of SC4K LAEs, these two methods also retrieve very similar SFRs of $5.9^{+6.3}_{-2.6}$ and $4.4^{+10.5}_{-2.4}\,\mathrm{M_\odot\,yr^{-1}}$ for the emission line-based and SED-based, respectively. Additionally, in the Appendix (Figure A.3), we show $\mathrm{SFR_{Ly\alpha}}$ vs $\mathrm{SFR_{SED}}$ at different redshift ranges. Both approaches predict very similar SFRs at all redshifts, outside the afore-mentioned ranges ($\mathrm{SFR_{SED}} \gtrsim 10\,\mathrm{M_\odot\,yr^{-1}}$) as the emission line-based approach becomes saturated at high SFRs.

Furthermore, in a recent study by Calhau et al. (2020), the SFR of the SC4K sample is derived through the stacking of radio imaging in the 3GHz band. For the stacking procedure, individual sources with direct detections are removed as these are likely AGN. They find median $\mathrm{SFR_{radio}} = 5.1^{+1.3}_{-1.2}\,\mathrm{M_\odot\,yr^{-1}}$ from the $z \sim 2-6$ stack, which is in very good agreement with emission line-based and SED-based SFR estimates of the sample.

6.1.7 Catalogue of SC4K LAE properties

We make public a catalogue with multiple measurements for individual LAEs in the SC4K sample. For each LAE we provide R.A., Dec, $\mathrm{L_{Ly\alpha}}$, $\mathrm{EW_0}$, X-ray and radio Flags (as given by Sobral et al. 2018a) and updated X-ray and radio Flags (as given by Calhau et al. 2020), $\mathrm{M_\star}$, β, $\mathrm{M_{UV}}$, $\mathrm{SFR_{Ly\alpha}}$ and $\mathrm{SFR_{SED}}$, with associated errors. We also provide our photometric measurements in Jansky for the 34 filters used in this work and a boolean SED flag which indicates unreliable SEDs. For LAEs in the unreliable SED subset, we do not derive SED-derived properties and thus set them to -99 in the public catalogues. We provide the catalogue of SC4K LAEs in electronic format in Appendix A.4.
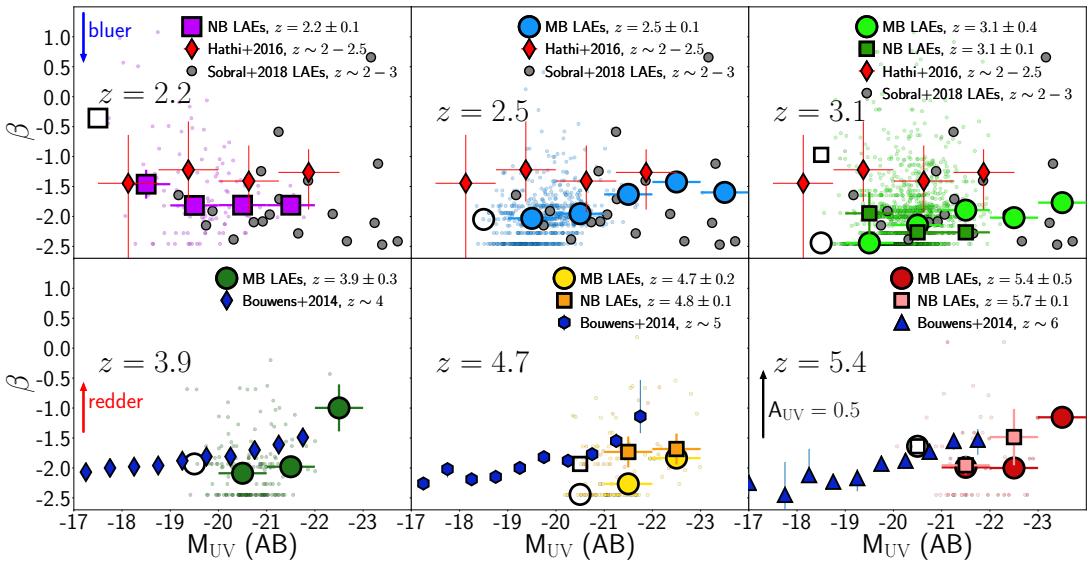
Figure 6.4: UV-continuum slope β (measured from SED fitting, see Section 6.1.4) vs UV luminosity M_{UV} (derived by integrating the SED fits at $\sim 1500\,\text{Å}$, see Section 6.1.3). Each panel contains LAEs from different redshift intervals (from left to right $z = 2.2, 2.5, 3.1, 3.9, 4.7, 5.4$). The median β of each M_{UV} bin of LAEs selected through medium (narrow) band filters is shown as filled coloured circles (squares) with the individual points being plotted as scatter in the background. Unfilled markers are likely biased bins, as discussed in Section 6.2.1. The clustering of points at $\beta = -2.44$ is a physically imposed model limitation as β can not become bluer without increasing the upper mass of the IMF to unreasonable values. For comparison we add measurements from LAEs at $z \sim 2 - 3$ (Sobral et al., 2018b) and UV-continuum selected samples at $z \sim 2 - 2.5$ (Hathi et al., 2016) and $z \sim 4$, $z \sim 5$ and $z \sim 6$ (Bouwens et al., 2014a). The black arrow is the size in β of $A_{UV} = 0.5$ ($A_{UV} = 4.43 + 1.99\beta$, Meurer et al. 1999). We find the median β in LAEs to be as blue or bluer than UV-selected samples at the same M_{UV} for all redshifts.
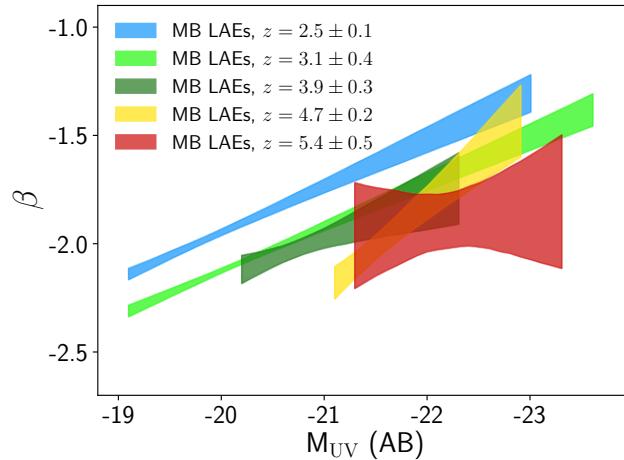
Figure 6.5: The evolution of the M_{UV}-β relation for LAEs. Shaded regions are the 1σ intervals obtained by bootstrapping the individual measurements for which we are not significantly biased (see Section 6.2.1). β increases with M_{UV} and this relation shifts down to smaller β as we move to higher redshifts. Most of this trend seems to be captured by a decrease in the normalisation of the relation.

6.2 Results and Discussion

6.2.1 $M_{UV} - \beta$ relation for LAEs and its evolution

The UV rest-frame luminosity (M_{UV}) and the UV β slope follow a tight correlation in UV-continuum selected samples (e.g. Bouwens et al., 2014a), with faint M_{UV} galaxies being typically bluer (more negative β). We measure how these two parameters are correlated for LAEs, whether they follow a similar M_{UV}-β relation as UV-continuum selected samples, and whether the relation evolves.

In Figure 6.4, we show the relation between M_{UV} (Section 6.1.3) and β (Section 6.1.4) for 6 redshift intervals ($z = 2.2, 2.5, 3.1, 3.9, 4.7, 5.4$). We note that at very faint M_{UV} we are biased towards redder sources. This is a consequence of redder sources being easier to detect in the optical filters, while sources with a very steep continuum slope will fall below our detection limits, particularly faint M_{UV} sources. As such, in Figure 6.4, we show the faintest M_{UV} bin as unfilled.

LAEs are found to be consistently bluer than UV-selected samples (Bouwens et al., 2014a; Hathi et al., 2016) at similar redshifts (up to ~ 1 dex bluer),

regardless of being NB or MB-selected, at all redshifts studied (see also Hashimoto et al., 2017). Our results are consistent with $z \sim 2 - 3$ LAEs measurements from Sobral et al. (2018b). Additionally, we measure an increase of β with M_{UV} (~ 0.5 dex per $\Delta M_{UV} = 2$), indicating that bright M_{UV} LAEs are redder than fainter LAEs at all redshift ranges, even though LAEs are typically bluer than LBGs. This tight correlation between M_{UV} and β is very similar to the one observed in LBG populations, implying an important overlap between the populations and also an important diversity within the LAE population.

In Figure 6.5, we show the 1σ contours for the M_{UV} vs β distribution. We compute the 1σ contours by bootstrapping our individual data points. We choose a random subset of 50% of the data points, determine the best fit, iterate the process 1000 times and define the 1σ contours as the 16th and 84th percentiles of all fits. As previously mentioned, faint M_{UV} bins will be biased towards redder sources, which are easier to detect in the continuum. As such, we apply a M_{UV} cut to our fits, equal to the lower limit of the faintest filled M_{UV} bin (Figure 6.4).

Overall we find a M_{UV}-β relation for LAEs, which is qualitatively very similar to the one observed in UV-selected samples. As can be seen in Figure 6.5, the normalisation of the M_{UV}-β relation slowly moves to bluer β with increasing redshift for LAEs. This might be explained due to a consistent average decrease in dust content and metallicity even within LAEs from low to high redshift.

6.2.2   Implications of $M_{UV} - \beta$ relation for LAEs

The UV continuum β slope can be an indicator of the dust attenuation of a galaxy as well as the age and metallicity of its stellar population, but because it is sensitive to all these effects, it can also be very complicated to interpret (see e.g. Popping et al., 2017). As shown by Bouwens et al. (2012) (see Figure 13 therein), a negative offset of $\sim 0.5 - 1$ in β should be dominated by a change in dust, albeit age and metallicity can also significantly steepen β, with a hotter population of stars. This suggests that LAEs are a subset of the SFG population which is very young and likely more metal-poor, with significant contribution from O and B stars which make the UV continuum steeper.
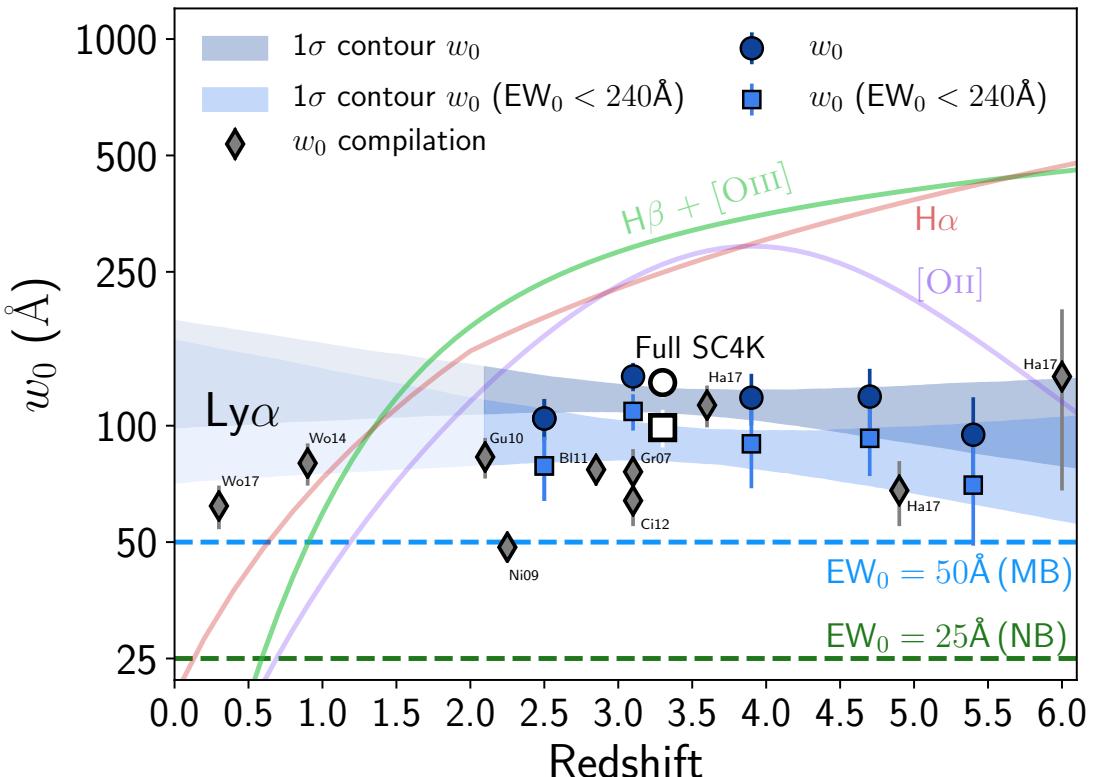
Figure 6.6: Global Lyα w_0 evolution with redshift. Best w_0 estimates are shown as blue circles (squares) for the full range of EW$_0$ (EW$_0$ < 240 Å). Blue contours are estimated by perturbing the w_0 bins within error bars (see Section 6.1.2.1 for details). We find evidence for little to no evolution of w_0. The white points show Lyα w_0 of the full SC4K sample. We present a compilation of Lyα w_0 from $z = 0.3$ to $z \sim 6$ (Blanc et al., 2011; Ciardullo et al., 2012; Gronwall et al., 2007; Guaita et al., 2010; Hashimoto et al., 2017; Nilsson et al., 2009; Wold et al., 2014, 2017). In addition, we show the [OII] (Hβ + [OIII]) rest-frame equivalent widths of emitters selected by these lines (Khostovan et al., 2016) as purple (green) fits and Hα EW$_0$ (Faisst et al., 2016; Matthee et al., 2017a) as red. Overall, the consensus of all data points is that there is no significant Lyα w_0 evolution with redshift despite the strong increase in the typical EW$_0$ of non-resonant lines for a wider population of SFGs.

In LBGs, β has been shown to depend on the UV luminosity, with a similar slope independent of redshift (e.g. Bouwens et al., 2012, 2014a). The normalisation of the relation is shifted to bluer β as we move to higher redshifts which can be explained by a lower dust content/lower dust extinction in galaxies at higher redshift (e.g. Finkelstein et al., 2012). As shown in Figure 6.5, LAEs have a very similar behaviour to LBG galaxies: β is tightly correlated with M_{UV}, with brighter M_{UV} galaxies being redder and the normalisation of this slope shifting to lower β with increasing redshift, which can be explained by a lower dust content at higher redshift even for LAEs. Similar observations of the M_{UV}-β trend and the β evolution with redshift have been shown by Hashimoto et al. (2017). The work presented by Hashimoto et al. (2017) reaches fainter M_{UV} than the work presented here and thus provides a consistent view of UV properties in LAEs from a complementary work using a different selection method (integral field spectroscopy with MUSE).

6.2.3 Lyα EW$_0$ and w_0: evolution for LAEs?

EW$_0$ is an indicator of the strength of an emission line relatively to the continuum. As such, it holds important information about a galaxy, with high EW$_0$ being associated with young stellar ages, low metallicities and top-heavy IMFs (Raiter et al., 2010; Schaerer, 2003). We use our sample of LAEs at well-defined redshift ranges to probe for redshift evolution of EW$_0$.

We find the median Lyα EW$_0$ of SC4K LAEs to remain constant at $\sim 140\,\text{Å}$ with redshift, both in MB and NB-selected samples (median EW$_0 = 138^{+284}_{-70}\,\text{Å}$). We show the little to no evolution of median EW$_0$ in Figure A.4. For individual filters, we detect a tentative higher than average EW$_0$ at $z \sim 5.7-5.8$, which could be caused by the small sample size or higher contamination fraction, although we highlight the large error bars.

The calculated median Lyα EW$_0$ can be very sensitive to selection effects, and it is possible that the non-evolution we measure is a consequence of the relatively high EW$_0 > 50\,\text{Å}$ cut applied in SC4K. In order to further tackle this, we also investigate the evolution of the scale parameter w_0 (Section 6.1.2.1). w_0 has been extensively probed in the literature (see e.g. Ciardullo et al., 2012; Hashimoto
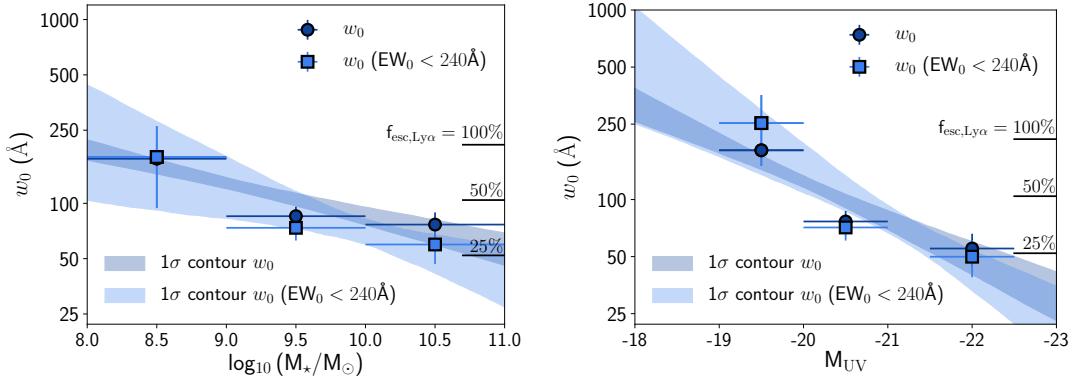
Figure 6.7: The Lyα w_0 dependence on $M_\star$ and M_{UV}. Best w_0 estimates are shown as blue circles (squares) for the full range of EW_0 ($EW_0 < 240\,\text{Å}$). A label with $f_{esc,Ly\alpha}$ ($= 0.048 w_0$; Sobral & Matthee 2019) is added for a potential physical interpretation of results. *Left:* Lyα w_0 is anti-correlated with stellar mass, such that the most massive LAEs have the lowest w_0 and likely the lowest $f_{esc,Ly\alpha}$. *Right:* Lyα w_0 is also anti-correlated with UV luminosity, with the faintest UV LAEs having the highest Lyα w_0.

et al., 2017), particularly because the exponential decay of the EW_0 distribution should be less affected by observational EW_0 cuts.

Our results are presented in Figure 6.6. We find no statistically significant evolution of the Lyα w_0 with redshift. Generally, w_0 is slightly higher when computed for the full sample of LAEs, and lower when we impose a restriction on the Lyα EW_0 ($< 240\,\text{Å}$), but no significant evolution is seen when using a single self-consistent method. We therefore conclude that both the observed median Lyα EW_0 and the distributions of Lyα w_0 for LAEs are not changing significantly from $z \sim 2$ to $z \sim 6$. A non-evolution of w_0 suggests there is no significant evolution in the typical or average properties of sources selected as LAEs across cosmic time. These include their typical metallicities and dust properties, but also perhaps more importantly their Lyα escape fraction, $f_{esc,Ly\alpha}$. As shown by Sobral & Matthee (2019), the observed Lyα EW_0 can be used to estimate $f_{esc,Ly\alpha}$. The non-evolution of Lyα EW_0 and w_0 across time implies non-evolving $f_{esc,Ly\alpha}$ for LAEs. For SC4K LAEs, we infer a constant $f_{esc,Ly\alpha}$ of $\approx 0.6 - 0.7$ across cosmic time ($\approx 0.5 - 0.6$ when applying the $EW_0 > 240\,\text{Å}$ cut). These median $f_{esc,Ly\alpha}$ values are consistent with those derived using radio

SFRs for SC4K Lyα emitters (0.7 ± 0.2, see Calhau et al., 2020).

However, it should be noted that different redshifts do not necessarily probe the same M_{UV} ranges (Figure 6.2, middle panel), which should be considered when discussing w_0 evolution with redshift, particularly as w_0 depends on M_{UV} (see Section 6.2.3.2). We attempt to explore potential bias effects by computing w_0 with a consistent M_{UV} cut. For the full SC4K sample, we compute w_0 for $-22 < M_{UV} < -19$ which is a M_{UV} range probed by all redshifts (see middle panel of Figure 6.2). For this cut, a flat relation (non-evolution) is still observed within $0.9\,\sigma$ for $EW_0 < 240\,\text{Å}$ and $1.8\,\sigma$ for the full EW_0 range. The different M_{UV} ranges probed by different selection filters/redshifts do not seem to be sufficient to explain the non-evolution of w_0 with redshift, which is likely a characteristic of the LAE population itself.

6.2.3.1   Comparison with other studies

In order to compare our results with other studies across different redshifts, in Figure 6.6 we show a compilation of Lyα w_0 in samples of LAEs, from $z \sim 0$ to $z \sim 6$ (Blanc et al., 2011; Ciardullo et al., 2012; Gronwall et al., 2007; Guaita et al., 2010; Hashimoto et al., 2017; Nilsson et al., 2009; Wold et al., 2014, 2017). Our results agree well with Hashimoto et al. (2017), Guaita et al. (2010) and Blanc et al. (2011). Furthermore, our extrapolation of w_0 to low redshift is consistent with the results from Wold et al. (2014, 2017).

Our measurements reveal higher values than those by Nilsson et al. (2009), Gronwall et al. (2007) and Ciardullo et al. (2012), all at intermediate redshifts ($z = 2.25 - 3.1$) and with selections that go to much lower EWs. We note however that the w_0 measured by Nilsson et al. (2009) is below our MB detection threshold and that our blind selection of LAEs is not sensitive to the lowest EW_0, as highlighted in Figure 6.6. Our LAE selection of high EW LAEs is much more similar to blind surveys done with MUSE (Hashimoto et al., 2017), but SC4K allows the selection and study of much higher luminosity LAEs. Furthermore, we note that our w_0 measurements shift to smaller values when the $EW_0 < 240\,\text{Å}$ cut is applied, becoming even more similar to the measurements reported in the literature.

While there are observed variations due to different sample selections which contribute to the scatter (Figure 6.6), overall we conclude that there is no clear evolution of the Lyα EW$_0$ and w_0 for LAEs when taking into account all measurements. Such parameters remaining constant for LAEs contrasts with measurements from other non-resonant emission lines for the general star-forming population, which are found to increase significantly with redshift. In order to provide a rough comparison, in Figure 6.6 we also show the redshift evolution of the rest-frame EW of line-emitters, including [OII] and Hβ + [OIII] emitters (Khostovan et al., 2016) and Hα EW$_0$ (Sobral et al., 2014). While at $z \sim 0$ those non-resonant rest-frame optical emission lines have typical EW$_0 < 25$ Å, by $z \sim 2$ they already exceed Lyα EW$_0$. This reveals a very significant evolution of the typical stellar populations of the general population of SFGs, while those selected to be LAEs have high Lyα EW$_0$ at all cosmic times. Since LAEs have typically high EWs in their rest-frame optical lines, it is very likely that we are seeing star-forming galaxies becoming, on average, LAEs, towards $z \sim 6$. Such possibility would easily explain the rise in the global Lyα/UV luminosity densities (see full discussion and implications in Sobral et al., 2018a).

6.2.3.2 The w_0 and f$_{\mathrm{esc,Ly\alpha}}$ dependence on M$_\star$ and M$_{\mathrm{UV}}$

LAEs seem to show no evolution in their typical Lyα w_0 across cosmic time. However, one could expect that LAEs with different physical properties may show different w_0, particularly as a consequence of different Lyα escape fractions (see e.g. Matthee et al., 2016; Oyarzún et al., 2017; Sobral & Matthee, 2019).

We start by investigating how Lyα w_0 may depend on the stellar mass of LAEs. The results are presented on the left panel of Figure 6.7, where we show the results when restricting the measurements to EW$_0 < 240$ Å and when using full samples. We find an anti-correlation between Lyα w_0 and stellar mass, with the least massive LAEs having $w_0 \approx 180$ Å and the most massive having $w_0 \approx 70$ Å. By using Sobral & Matthee (2019), this could be seen as a significant difference in the typical f$_{\mathrm{esc,Ly\alpha}}$ which would decline from $\approx 90\%$ for M $\sim 10^{8.5}$ M$_\odot$ LAEs to f$_{\mathrm{esc,Ly\alpha}} \approx 30\%$ for M $\sim 10^{10.5}$ M$_\odot$ LAEs. This trend is very similar to those

found by Matthee et al. (2016) for a general population of Hα emitters with much higher SFRs and lower $f_{esc,Ly\alpha}$ than our LAEs and by Oyarzún et al. (2017).

In Figure 6.7 (right panel) we also show how Lyα w_0 is clearly anti-correlated with M_{UV}. Our results show that UV luminous LAEs in our sample ($M_{UV} \approx -21.5$) have Lyα $w_0 \approx 50$ Å, which rises with declining UV luminosity to $w_0 \approx 180$ Å for $M_{UV} \approx -19.5$ LAEs. This implies that the UV faintest sources have the highest $f_{esc,Ly\alpha}$ (Sobral & Matthee, 2019) of around $\approx 85\%$, while the most UV luminous LAEs have $f_{esc,Ly\alpha} \approx 20 - 30\%$. Our results are in good agreement with Oyarzún et al. (2017) and reveal that even though LAEs have high Lyα w_0 across cosmic time, the population still shows important trends with stellar mass and rest-frame UV luminosity.

6.2.3.3 LAEs with extreme EW$_0$

The nature of LAEs with extremely high EW$_0$ and the processes behind the creation of such extreme lines are still a relatively unexplored topic despite a range of discoveries (e.g. Cantalupo et al., 2012; Hashimoto et al., 2017; Kashikawa et al., 2012; Maseda et al., 2018). Typical internal star formation processes should not be enough to power EW$_0 > 240$ Å in Lyα (Raiter et al., 2010; Schaerer, 2003), but studies like Cantalupo et al. (2012) suggest that such extreme objects which have been found could be explained by fluorescent "illumination" from e.g. a nearby quasar (see also Rosdahl & Blaizot, 2012; Yajima et al., 2012). Additionally, an extreme $z = 6.5$ LAE with EW$_0 = 436$ Å is reported in Kashikawa et al. (2012), with the authors arguing that such a high EW$_0$ requires a very young, massive and metal-poor stellar population, or even Population III stars.

The large volume covered by SC4K ($\sim 10^8$ Mpc^3) and the sensitivity to the highest EWs provides a unique opportunity to identify and quantify the number density of extremely high EW LAEs. In order to do so in a conservative way, rather than simply selecting sources with Lyα EW$_0$ higher than 240 Å, we take the photometric errors fully into account, and we use the 3σ errors. In practice, we look for LAEs within SC4K which satisfy EW$_0 > 240$ Å at a 3σ level [1] and for which we have no evidence of AGN activity. We find a total of 45

---

[1] EW$_0 - 3\Delta$EW$_0 > 240$ Å

"extreme" non-AGN LAEs in $\sim 61.5 \times 10^6\,\mathrm{Mpc}^3$ and we investigate how these are distributed across redshift. The results are shown in Table 6.1, where we use Poisson errors. Most of the extreme LAEs are found at $z \sim 2 - 3$. Furthermore, by taking into account the volumes surveyed, we find that the number density of extreme LAEs within SC4K rises, from $(0.12 \pm 0.08) \times 10^{-6}\,\mathrm{Mpc}^{-3}$ at $z \sim 5.4$ to $(1.50 \pm 0.61) \times 10^{-6}\,\mathrm{Mpc}^{-3}$ at $z \sim 2.5$, although such increase should be treated with caution, as the higher redshift sample does not reach very faint $\mathrm{M_{UV}}$ (> -20) ranges. Overall, we find a number density of $(0.73 \pm 0.11) \times 10^{-6}\,\mathrm{Mpc}^{-3}$ at $z \sim 2 - 6$, revealing that these sources are exceptionally rare. At $1\,\sigma$ confidence level, we find 318 LAEs with $\mathrm{EW_0} > 240\,\mathrm{\AA}$, resulting in a number density of $(5.17 \pm 0.29) \times 10^{-6}\,\mathrm{Mpc}^{-3}$. Spectroscopic follow-up observations are required to further understand their nature. We find our 45 "extreme" sources to be a diverse population, as they are found at all Lyα luminosities and stellar masses, but preferentially at faint UV luminosities which is a consequence of high EW + no dependence on Lyα luminosity. They typically have blue UV β slopes but some reach redder values ($\beta \sim -1.2$). We do not observe a spatial correlation between "extreme" LAEs and AGN, which we would expect if the high EWs in this sample of LAEs were generated by fluorescent "illumination".

Through a narrow band filter search, Cantalupo et al. (2012) targeted a field centred in a hyper luminous quasar and identified 18 LAEs at $z = 2.4$ in a comoving volume of $5500\,\mathrm{Mpc}^3$. Stacking of these sources results in Lyα $\mathrm{EW_0} > 800\,\mathrm{\AA}$ ($1\,\sigma$), which cannot be explained by typical star-formation processes. This implies a higher number density of extreme LAEs than the conservative number density we report in this work, although this can be easily explained by Cantalupo et al. (2012) specifically targeting a quasar field.

In a more comparable blank search, using deep MUSE data, Hashimoto et al. (2017) selected 6 LAEs with $\mathrm{EW_0} > 240\,\mathrm{\AA}$ at a $1\,\sigma$ level (zero at $3\,\sigma$) in $9.31 \times 10^4\,\mathrm{Mpc}^3$ (Drake et al., 2017a) at $z \sim 2 - 6$. This results in a number density of $\sim 6 \times 10^{-5}\,\mathrm{Mpc}^{-3}$, suggesting these "extreme" LAEs may be even more common at fainter luminosities than those in the SC4K sample.
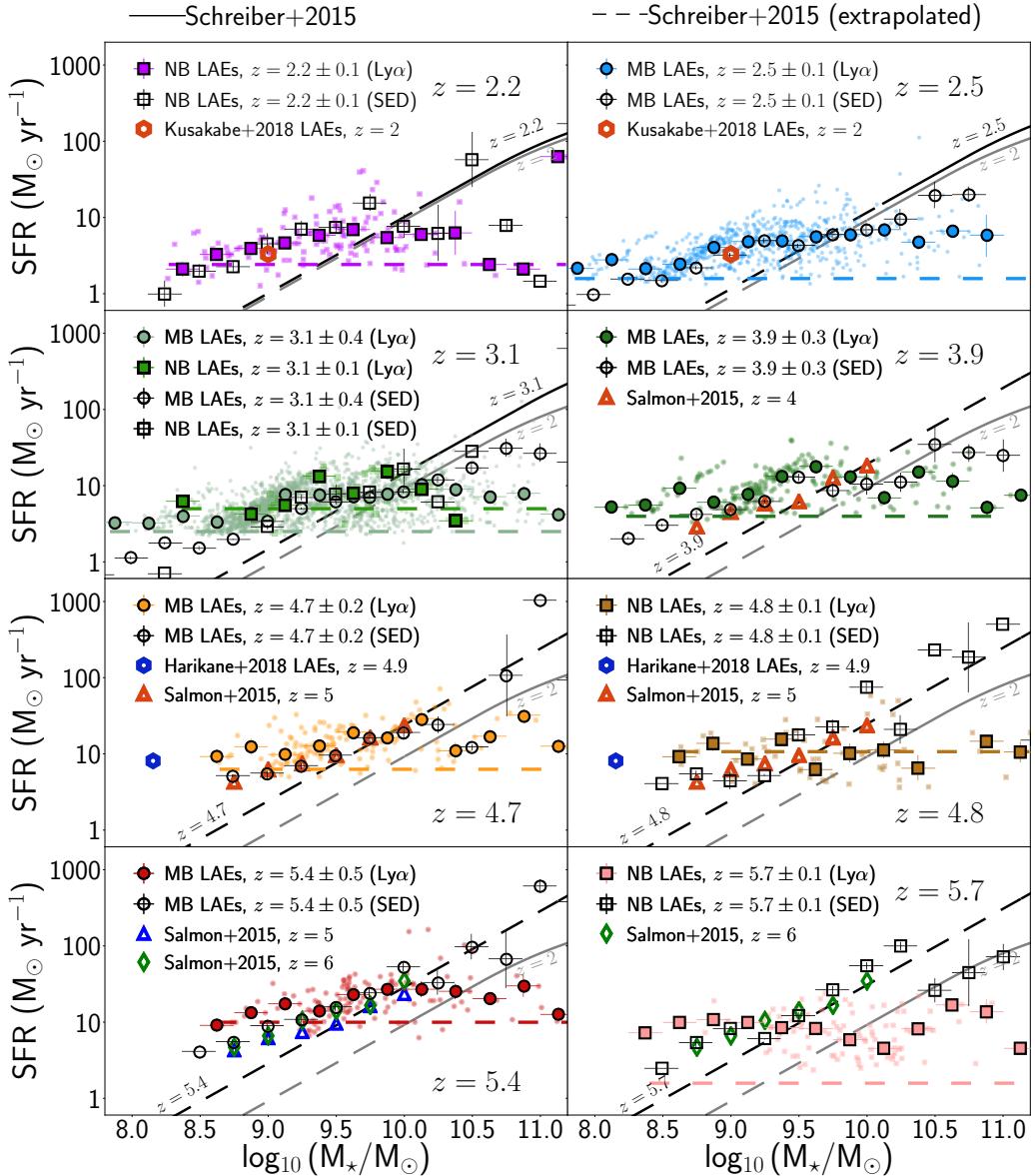
Figure 6.8: SFR (derived from Lyα and derived from SED fits, see Section 6.1.6) vs $M_\star$ (derived from SED fits, see Section 5.6). Each panel contains LAEs from different redshift intervals (from left to right $z = 2.2, 2.5, 3.1, 3.9, 4.7, 5.4$). The median $SFR_{Ly\alpha}$ of each $M_\star$ bin for LAE selected through medium (narrow) band filters is shown as filled coloured circles (squares) with the individual $SFR_{Ly\alpha}$ being plotted as scatter in the background. The median SFR_{SED} of each $M_\star$ bin for LAE selected through medium (narrow) band filters is shown as open circles (squares). The $M_\star$ bins of the two methods are defined with a 0.25 dex offset for better visibility. The dotted horizontal line is the average SFR depth, computed from the flux depth and average EW_0 of the sample. The continuous black lines are the best-fit relations from Schreiber et al. (2015) computed for the redshift of each panel and converted from Salpeter to Chabrier. These relations are shown as dashed lines for the mass ranges where they were extrapolated.

Table 6.1: Number count and number density of LAEs with $EW_0 > 240\,\text{Å}$ at a $3\,\sigma$ level, for different redshift intervals, using comoving volumes from Table 3.1. Errors are Poissonian. We find very low number densities of extreme LAEs, but these increase with decreasing redshift.

Redshift interval	N (# LAEs)	Φ $(10^{-6}\,\text{Mpc}^{-3})$
MB, $z = 2.5 \pm 0.1$	6 (± 2)	1.50 ± 0.61
MB, $z = 3.1 \pm 0.4$	15 (± 4)	0.82 ± 0.21
MB, $z = 3.9 \pm 0.3$	4 (± 2)	0.40 ± 0.20
MB, $z = 4.7 \pm 0.2$	2 (± 1)	0.17 ± 0.12
MB, $z = 5.4 \pm 0.5$	2 (± 1)	0.12 ± 0.08
Full sample	45 (± 7)	0.73 ± 0.11
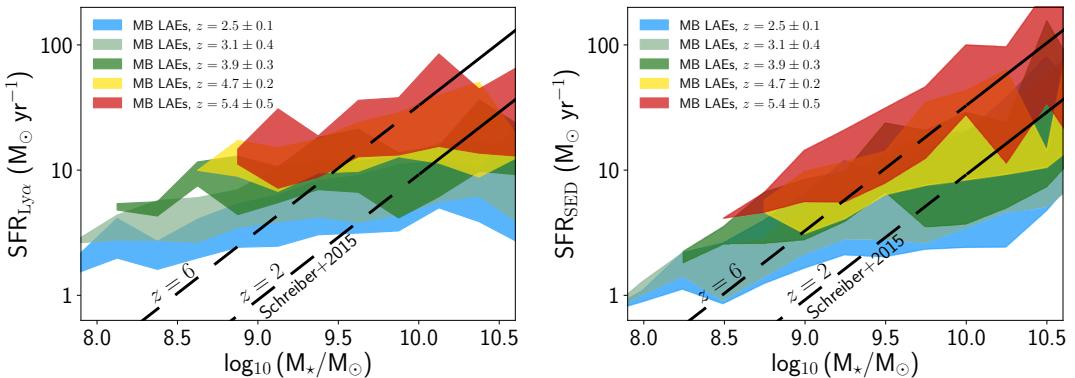
Figure 6.9: *Left:* Running average of SFR (derived from Lyα and EW_0, see Section 6.1.6) vs $M_\star$ (derived from SED fits, see Section 5.6). *Right:* Same but with SFR derived from MAGPHYS (see Section 6.1.6). The SFR-$M_\star$ slopes derived from the two methods are different, with the SED-derived slope being steeper. The difference is likely a consequence of $SFR_{Ly\alpha}$ not being able to reach very low ($< 1\,\text{M}_\odot\,\text{yr}^{-1}$) and very high SFRs ($> 20 - 30\,\text{M}_\odot\,\text{yr}^{-1}$), but we provide further discussion in Section 6.1.6.3. For comparison, we show the Main Sequence line for UV-continuum selected sources from Schreiber et al. (2015), where the dashed lines show the extrapolated values.

6.2.4 SFR-M$_\star$ relation and evolution

We test the dependence of SFR on M$_\star$ in our sample of LAEs and its potential evolution with redshift. In Figure 6.8 we show SFR derived from the SEDs and SFR derived from Lyα (see Section 6.1.6) vs M$_\star$ (derived from SEDs, Section 6.1.5) for our sample of LAEs and compare with SFRs derived from SED fitting. We compare our measurements with the Main Sequence relation as derived in Schreiber et al. (2015) (converted from Salpeter to Chabrier IMF, extrapolated to low mass ranges when required) and a few studies at different redshifts.

We find that in general there is a relation between SFR and M$_\star$ at all redshifts for LAEs. The relation is relatively shallow when using Lyα SFRs and steeper when using SED SFRs, as can be seen in Figure 6.8. The relation between SFR and M$_\star$ seems to steepen with increasing redshift for LAEs when using SED SFRs, as can also be seen in Figure 6.9 (right panel). This steepening with increasing redshift also seems to make the SFR-M$_\star$ relation much more in line with the extrapolated relations found for UV-continuum selected sources (e.g. Schreiber et al., 2015).

At $z < 4$, we find that LAEs are typically above the Main Sequence relation at their corresponding redshift. This is particularly evident for low stellar masses (M$_\star < 10^{9.5}$ M$_\odot$) although we find that more massive LAEs tend to be within the Main Sequence or even below it, a consequence of the slope of the relation being shallower. At higher redshifts, we find that even at low stellar masses ($10^{9.0-9.5}$ M$_\odot$) LAEs are closer to the Main Sequence or that the Main Sequence becomes closer to the relation valid for LAEs, as SFGs may become more LAE-like. Our results therefore suggest that at higher redshifts there is a wider overlap between LAEs and more "normal" populations of galaxies, as UV-continuum selected galaxies become LAE-like. This could explain the agreement between high-z LAEs and the results of Salmon et al. (2015). It is nonetheless important to point out (as shown in Figure 6.8) that the flux limit in Lyα corresponds to a rough cut in SFR and therefore a bias towards higher SFRs at the lowest masses. Similar flux cuts also affect continuum-selected samples, placing them well above the Main Sequence (see e.g. Tasca et al., 2015).

Our results are in good agreement with measurements from Lyα-selected samples from Kusakabe et al. (2018) at $z = 2$ within error bars. We also compare our results with those presented by Harikane et al. (2018). While we do not reach such low masses, our results are consistent with LAEs being above the Main Sequence at low stellar masses. With our SC4K sample of LAEs, we can now analyse the evolution of the SFR-$M_\star$ in wide mass ranges at different redshifts, no longer being constrained by single bins or having to stack sources to SED fit the stacked photometry, being able to probe the evolution of the relation within the same sample.

As previously discussed in Section 6.1.6.3, there are limitations to different SFR methods, which are important to highlight when comparing the SFR-$M_\star$ relation. $SFR_{Ly\alpha}$ consistently predicts higher SFR than SFR_{SED} for low stellar masses and lower SFR for very high stellar masses. In fact, individual measurements of $SFR_{Ly\alpha}$ seem to fully saturate at $\sim 100\,M_\odot\,yr^{-1}$ with the medians typically not going above $\sim 20 - 30\,M_\odot\,yr^{-1}$ (see Sobral et al., 2018b). $SFR_{Ly\alpha}$ also implies higher SFRs at lower masses, possibly due to tracing more recent star-formation which would be higher than the one measured from the continuum if LAEs are going through bursts of star-formation. This can be clearly seen with NB LAEs measurements at $z = 5.7$, where the low luminosity sample predicts SFRs $\approx 10\,M_\odot\,yr^{-1}$. SFR_{SED} may be better suited for such conditions, and as seen in Figure 6.8, it points towards a relation similar to the Schreiber et al. (2015) extrapolations for the entire mass range we can probe. Nevertheless, we find that the SFRs derived from the two approaches to be consistent, with the same trends being observed from both. In Figure 6.9 we show the running averages for $M_\star$ vs SFR. We find the normalisation of the relation to increase with redshift (left panel) but, as previously discussed, this is mostly driven by detection limits, as we are only capable of reaching down to SFR $< 5\,M_\odot\,yr^{-1}$ at $z \sim 2$.

6.2.5   LAEs: are they "Main Sequence" galaxies?

The stellar mass of a galaxy and its star formation rate are correlated in typical galaxies, creating a trend known as the "Main Sequence"[1] of SFGs (Brinchmann et al., 2004; Noeske et al., 2007). A priori, we can naively expect this correlation to occur as the stellar mass of a galaxy is the integral of SFR across time, so the total amount of stars produced will be proportional to the current SFR, assuming a continuous SFR. This dependence can lead to "tracks" in SED-fitting derived values which lead to a more stringent correlation between SFR and $M_\star$. Galaxies going through periods of intensive star formation, which may be a consequence of bursty star formation, will occupy a region above the Main Sequence. In typical galaxies, SFR and $M_\star$ are in tight correlation and the normalisation of the relation increases with redshift (e.g. Schreiber et al., 2015). Understanding whether the Main Sequence trend holds for LAEs provides important insight into how star formation occurs and how it is driven in this population of predominantly early, primeval galaxies. In principle, we do not expect a Lyα-selected sample to span uniformly around the Main Sequence, because we select on emission line strength which at fixed stellar mass always gives high $\mathrm{sSFR} \equiv \mathrm{SFR}/M_\star$. We therefore do not expect to use LAEs to measure the Main Sequence in an unbiased way, but we can use the comparison to the Main Sequence to determine how LAEs fit in the general galaxy population. Several measurements at $z > 2$ have measured the Main Sequence relation by probing $M_\star > 10^{10}\,M_\odot$, with the low mass limit typically rising to $M_\star > 10^{11}\,M_\odot$ at $z > 3.5$ (Schreiber et al., 2015), but some recent studies have measured the SFR-$M_\star$ slope and scatter down to $M_\star = 10^9\,M_\odot$ (Salmon et al., 2015). Our sample of high redshift, typically low $M_\star$ SFGs reaches a region still widely uncharted at these redshift ranges.

Our results point towards an intensive star formation nature for low mass LAEs at $z < 4$, which places them significantly above the extrapolation of the Main Sequence to the lowest masses. A more bursty star-forming nature could explain these SFRs above the Main Sequence. However, we cannot directly infer burstiness from our measurements. More massive LAEs seem to fall within the

---

[1] Note that galaxies do not evolve along the Main Sequence trend and it is therefore not an evolutionary sequence, see e.g. Matthee & Schaye (2019).

Main Sequence. At higher redshifts, SFR_{SED}-$M_\star$ measurements for LAEs start to resemble more the Main Sequence at all mass ranges. We also find $\text{SFR}_{\text{Ly}\alpha}$-$M_\star$ to follow a Main Sequence-like relation at $z > 4$, except for $M_\star \gtrsim 10^{10.5}$, when $\text{SFR}_{\text{Ly}\alpha}$ seems to saturate, likely due to dust, and is not able to reach SFRs as high as SFR_{SED}. This can easily be explained by more massive galaxies showing much higher dust extinction (see e.g. Garn & Best, 2010; Sobral et al., 2012; Whitaker et al., 2017), which at some point might completely absorb Lyα and UV photons in high SFR regions (Sobral et al., 2018b), making it impossible for them to be observed. In such cases, the FIR and some visible and NIR light can still escape, leading to a large discrepancy between SFR_{SED} and $\text{SFR}_{\text{Ly}\alpha}$. We note that $\text{SFR}_{\text{Ly}\alpha}$ contains an empirical correction for dust extinction (see Sobral & Matthee, 2019), but this was calibrated for typical LAEs where only moderate to low levels of dust extinction are present leading to Lyα and UV photons being attenuated, but not fully destroyed. At the highest masses, we are likely seeing LAEs with several star-forming regions that may be completely invisible in the UV and Lyα but where at least one region has a hole or a porous ISM (see also Popping et al., 2017).

Overall, we find that the SFR-$M_\star$ relation for LAEs steepens with redshift and that its normalisation also rises with look-back time (see Figure 6.9). As a consequence, by $z \sim 5-6$, LAEs and the general UV-continuum selected population essentially become indistinguishable. This increasing overlap of populations with increasing redshift is also observed in the morphologies and sizes of SFGs, which become LAE-like (compact, $r_e \sim 1\,\text{kpc}$) towards high redshift (Paulino-Afonso et al., 2018) and diverge towards lower redshift as LAEs remain compact at all redshifts. Our results are also fully consistent with the rapid rise of the cosmic average Lyα/UV luminosity density ratio with increasing redshift (Chapter 4) which imply that a higher fraction of star-forming galaxies share the properties associated with LAEs, leading to a rise of the cosmic averaged Lyα escape fraction and the cosmic averaged ionisation efficiency, ξ_{ion}. Such results are also in agreement with other studies showing a rise of the LAE fraction in UV-selected sources towards $z \sim 6$ (Curtis-Lake et al., 2012b; Schenker et al., 2014; Stark et al., 2017), and globally imply that by $z \sim 6$ LAEs become representative of the majority of the star-forming population.

6.3   Conclusions

In this Chapter, we determined and explored key properties of a large sample of LAEs from the publicly available SC4K survey ($\sim$4000 LAEs at $z \sim 2-6$ in the COSMOS field; Sobral et al. 2018a). We computed SFRs, M_{UV}, β and $M_\star$ for each individual LAE and we provide a full catalogue of SC4K LAEs with all the photometric measurements and derived properties. Our main results are:

- SC4K LAEs are typically low stellar mass sources (median $M_\star=10^{9.3^{+0.6}_{-0.5}}\,M_\odot$), very blue in the rest-frame UV (β=-2.1 $^{+0.5}_{-0.4}$) and have low SFRs (SFR$_{Ly\alpha}$: $5.9^{+6.3}_{-2.6}\,M_\odot\,yr^{-1}$; SFR$_{SED}$: $4.4^{+10.5}_{-2.4}\,M_\odot\,yr^{-1}$).

- We observe a tight correlation between β and M_{UV}, qualitatively similar to the one observed in UV-selected samples. The normalisation of this correlation shifts to more negative β (bluer) with increasing redshift, which is consistent with a decreasing dust content with increasing redshift in galaxies even for LAEs.

- Our LAEs are as blue or bluer than UV-selected Lyman Break Galaxies (LBGs) at similar redshifts (up to $\sim$1 in the redshift range $z \sim 2-6$), suggesting they always constitute the youngest, most metal-poor and/or most dust-poor subset of the UV-selected sources.

- We find evidence for little to no evolution in the typical Lyα EW$_0$ and the scale parameter w_0 with redshift, suggesting the median $f_{esc,Ly\alpha}$ in LAEs is always high and not evolving strongly with redshift.

- We find that the Lyα w_0 (and thus $f_{esc,Ly\alpha}$) for LAEs declines with increasing stellar mass, implying that $f_{esc,Ly\alpha}$ is highest for the lowest stellar mass LAEs and lowest for the most massive LAEs. A similar trend is found with rest-frame UV luminosity, where the faintest LAEs have the highest typical EWs and the highest $f_{esc,Ly\alpha}$.

- We explore extreme EW$_0$ measurements in our large sample of LAEs and find 45 non-AGN LAEs with EW$_0$ > 240 Å at a 3σ level, resulting in a number density $(7\pm1)\times10^{-7}\,Mpc^{-3}$. These extreme emitters are incredibly

rare but can provide insight into extreme Lyα emission that is neither purely from typical star-formation or AGN.

- By using Lyα EW$_0$ to infer f$_{esc,Ly\alpha}$ (Sobral & Matthee, 2019) we compute Lyα SFRs which are independent of SED fitting measurements and we compare both. Lyα and SED-fitting based SFRs show a remarkable agreement for M$_\star = 10^{9-10}$ M$_\odot$ and SFR$_{SED} = 1 - 10$ M$_\odot$ yr^{-1}. SFR$_{Ly\alpha}$ predicts lower SFRs at more massive regimes, likely due to not being sensitive to heavily obscured parts of very massive galaxies.

- LAEs show a relation between stellar mass and SFR at all redshifts, but this is typically shallower than the relation found for the general star-forming population. We also find that the relation steepens and rises with increasing redshift for LAEs.

- LAEs are typically above the "Main Sequence" at $z < 4$ and M$_\star < 10^{9.5}$ M$_\odot$, indicating LAEs are experiencing more intense star formation than the general population of galaxies of similar mass at similar redshifts, with one possible explanation being a bursty star-formation nature of LAEs. For higher masses and redshifts, this offset decreases, implying a larger overlap between LAEs and more "normal" SFGs.

Overall, we find that LAEs are typically very young, low mass galaxies, albeit they still span an important range of properties, and within the LAE population there are important trends with stellar mass and UV luminosity. Typical properties of LAEs seem to have little evolution between $z = 2$ and $z = 6$, although they still become bluer and the relation between SFR and stellar mass steepens and rises slightly. By $z \gtrsim 4$, the overlap between LAEs and the more general UV-selected population becomes significant and by $z \sim 6$ they seem to become indistinguishable, as typical SFGs essentially become LAE-like. Our results reveal how galaxies selected as LAEs constitute mostly the youngest, most primeval galaxies at any redshift, and also that LAEs are ideal sources to study the dominant population of SFGs towards $z \gtrsim 6$ and therefore also likely a population that significantly contributed to reionising the Universe (see also discussions in Sections 1.1.2 and 4.6).

Chapter 7

The UV luminosity function and galaxy stellar mass function of $z \sim 2 - 6$ LAEs with SC4K

Abstract

We measure the evolution of the rest-frame UV luminosity function (LF) and the stellar mass function (SMF) of Lyman-α (Lyα) emitters (LAEs) from $z \sim 2$ to $z \sim 6$ by exploring ~ 4000 LAEs from the SC4K sample (see Chapter 2). We find a correlation between Lyα luminosity ($L_{Ly\alpha}$) and rest-frame UV (M_{UV}), with best-fit $M_{UV} = -1.6^{+0.2}_{-0.3} \times \log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) + 47^{+12}_{-11}$ and a shallower relation between $L_{Ly\alpha}$ and stellar mass ($M_\star$), with best-fit $\log_{10}(M_\star/M_\odot) = 0.9^{+0.1}_{-0.1} \times \log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) - 28^{+4.0}_{-3.8}$. As a consequence of these dependences, an increasing $L_{Ly\alpha}$ cut predominantly lowers the number density of faint M_{UV} LAEs, but the decrease is more uniform over the full $M_\star$ range. We estimate a proxy for the full UV LFs and SMFs of LAEs with simple assumptions of the faint end slope. To reduce selection bias, we analyse the same luminosity range ($\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$) at all redshifts. For the UV LF of LAEs, we find a characteristic number density (Φ^*) increase from $\log_{10}(\Phi^*/\mathrm{Mpc^{-3}}) \sim -5.6$ at $z = 2.5$ to ~ -4.6 at $z \sim 3$, remaining constant up to $z \sim 5 - 6$ and no clear evolution of the characteristic UV luminosity (M_{UV}^*). We find no significant evolution of the SMF of LAEs with redshift, with $\log_{10}(\Phi^*/\mathrm{Mpc^{-3}})$ staying constant at ~ -5.5 from $z \sim 2.5$ to $z \sim 6$ and the characteristic stellar mass staying constant at $\log_{10}(M_\star^*/M_\odot) \sim 11$ for the same redshift range. For $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ LAEs, ρ_{UV} increases from $10^{24.3}$ to $10^{25.0}$ erg s^{-1} Hz^{-1} Mpc^{-3} and ρ_M remains constant at $\sim 10^{5.5}\,M_\odot\,\mathrm{Mpc^{-3}}$, being always lower than the total luminosity and stellar mass densities of more typical galaxies but approaching it with increasing redshift. Both ρ_{UV} and ρ_M of LAEs are extrapolated to converge to the measurements of

continuum-selected galaxies at $z > 6$, which suggests a key role of LAEs in the epoch of reionisation.

7.1 Introduction

Multiple studies have used the Lyman-α (Lyα, $\lambda_{0,\mathrm{vacuum}} = 1215.67\,\text{Å}$) emission line to successfully select large samples of galaxies at $z > 2$ (e.g. Cowie & Hu, 1998; Drake et al., 2017a; Konno et al., 2018; Malhotra & Rhoads, 2004; Matthee et al., 2015; Ouchi et al., 2008; Rauch et al., 2008; Santos et al., 2016; Sobral et al., 2018a; Taylor et al., 2020; van Breukelen et al., 2005). Lyα emission is typically associated with young star-forming galaxies (SFGs, e.g. Partridge & Peebles, 1967) but can also be emitted from active galaxy nuclei (AGN; e.g. Calhau et al., 2020; Miley & De Breuck, 2008; Sobral et al., 2018b).

LAEs are typically young/primeval, low mass, low dust extinction sources (e.g. Gawiser et al., 2006, 2007; Lai et al., 2008; Pentericci et al., 2007), but a significant diversity of properties within the Lyα population has been reported in the literature (e.g. Acquaviva et al., 2012; Finkelstein et al., 2009; Hagen et al., 2016; Lai et al., 2008; Matthee et al., 2016; Oyarzún et al., 2017; Santos et al., 2020). Sources with high Lyα equivalent width (EW) typically have young stellar ages, low metallicities and top-heavy initial mass functions (e.g. Hashimoto et al., 2017; Raiter et al., 2010; Schaerer, 2003). LAEs have been shown to be very compact in the UV (e.g. Malhotra et al., 2012; Paulino-Afonso et al., 2018), with the compact morphology possibly being favourable to the escape of Lyα photons. Additionally, studies have shown that high-redshift LAEs may be progenitors of a wide range of galaxies, from present-day galaxies (e.g. Gawiser et al., 2007; Guaita et al., 2010; Yajima et al., 2012) to bright cluster galaxies (BCGs; e.g. Khostovan et al., 2019), highlighting the significance of LAEs in galaxy evolution studies.

Studies of UV-continuum selected galaxies have found that the Lyα fraction ($\chi_{\mathrm{Ly}\alpha}$, percentage of galaxies with Lyα emission) increases with redshift up to $z \sim 6$ (e.g. Caruana et al., 2018; Cassata et al., 2015; De Barros et al., 2017; Kusakabe et al., 2020; Pentericci et al., 2011; Stark et al., 2010). This might be explained by an average lower dust content in higher redshift galaxies (e.g. Bouwens et al., 2006; Stanway et al., 2005), increasing the Lyα escape fraction ($\mathrm{f}_{\mathrm{esc,Ly}\alpha}$, ratio between observed and intrinsic Lyα photons in a galaxy; e.g. Hayes et al., 2011) and/or increasing the ionising efficiency (ξ_{ion}, number of produced

ionising photons per unit UV luminosity; e.g. Matthee et al., 2017a). $\chi_{\mathrm{Ly}\alpha}$ is typically computed with large spectroscopic samples, with $\chi_{\mathrm{Ly}\alpha}$ being the ratio between the number of galaxies with Lyα emission detected above some Lyα EW threshold and the total number of probed galaxies (see e.g. Stark et al., 2010). $\chi_{\mathrm{Ly}\alpha}$ is found to be higher for galaxies fainter in the rest-frame UV (M_{UV}, e.g. Pentericci et al., 2011), implying such galaxies have higher escape fraction of Lyα photons and/or have a higher ξ_{ion} (e.g Maseda et al., 2020). This can also be linked with faint M_{UV} galaxies having higher Lyα EW (see e.g. Kusakabe et al., 2018; Shimizu et al., 2011) and thus being more susceptible to being picked as LAEs, although such trend could also be a consequence of selection effects or survey limits (see e.g. Ando et al., 2006; Hashimoto et al., 2017; Nilsson et al., 2009; Zheng et al., 2014). Some studies report no strong correlation between $\chi_{\mathrm{Ly}\alpha}$ and M_{UV} (Kusakabe et al., 2020) and attribute the typical high $\chi_{\mathrm{Ly}\alpha}$ of faint M_{UV} galaxies to selection biases in Lyman break galaxy (LBG) samples, which are biased towards selecting sources with high Lyα EW, as strong Lyα emission will boost the photometry and enhance the Lyman break, making such sources easier to detect.

Alternatively, $\chi_{\mathrm{Ly}\alpha}$ could in principle be inferred from the ratio between luminosity functions (LF, number density per luminosity bin vs luminosity) of Lyα-selected and UV continuum-selected samples. The UV LF of continuum-selected galaxies has been extensively constrained in multiple studies up to $z \sim 10$ (e.g. Alavi et al., 2016; Arnouts et al., 2005; Bouwens et al., 2015; Finkelstein et al., 2015; Mehta et al., 2017; Ono et al., 2018; Sawicki & Thompson, 2006; Steidel et al., 1999). The characteristic number density (Φ^*) is found to decrease with an increasing redshift, from $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3}) \sim -2.5$ at $z \sim 2$ (Reddy & Steidel, 2009) to ~ -3.5 at $z \sim 6$ (Bouwens et al., 2015; Ono et al., 2018), with no characteristic luminosity evolution (L^*), although other studies have discussed a possible L^* instead of the Φ^* evolution. The UV LF of LAEs has also been probed by multiple studies (see e.g. Hu et al., 2004; Ouchi et al., 2008; Shimasaku et al., 2006), targeting volumes of up to $\sim 10^6 \, \mathrm{Mpc}^3$. Ouchi et al. (2008) found no evolution of the UV LF of LAEs at $z \sim 3-4$, but an increase of UV bright LAEs at $z = 5.7$. It is important to establish whether such evolutionary trends hold for much larger volumes ($\sim 10^8 \, \mathrm{Mpc}^3$) and larger samples of LAEs.

Furthermore, it is important to establish how LAEs contribute to the total mass budget of galaxies. LAEs are typically low stellar mass galaxies, but can span a wide range of stellar masses, with some LAEs being very massive ($> 10^{10}\,\mathrm{M_\odot}$, e.g. Finkelstein et al., 2009). The stellar mass function (SMF) of continuum-selected galaxies has been well studied up to $z \sim 4$ (see e.g. Ilbert et al., 2013; Mortlock et al., 2011; Muzzin et al., 2013; Pozzetti et al., 2010; Santini et al., 2012). The SMF of continuum-selected galaxies is found to shift to lower number densities (Φ^* decrease) with increasing redshift, from $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3}) \sim -3.3$ at $z = 2 - 2.5$ to ~ -5.0 at $z = 3 - 4$ (Muzzin et al., 2013). For LAEs, our understanding of the SMF is very limited as most studies are only able to determine stellar masses of stacks of the population (e.g. Kusakabe et al., 2018). Estimating stellar masses of high-redshift LAEs is challenging due to near-infrared (NIR) coverage typically not being deep enough. Recent programs such as UltraVISTA (McCracken et al., 2012) DR4 provide ultra-deep NIR imaging which can be used to better constrain the spectral energy distribution of high-redshift galaxies. Measurements of the stellar mass of individual galaxies in large samples spanning wide redshift ranges can significantly improve our view on the evolution of LAEs and how they compare with more typical galaxy samples.

In this work, we use a uniformly selected sample of ~ 4000 LAEs (SC4K, see Chapter 2) to compute UV LFs and SMFs in the wide redshift range $z \sim 2 - 6$. We use the publicly available catalogues from Calhau et al. (2020) which identify AGN candidates in the SC4K sample using X-ray and radio measurements and the publicly available catalogues from Chapters 2 and 6 which have measurements of the UV luminosity and stellar mass of LAEs in the SC4K sample. By comparing the luminosity and stellar mass density of LAEs with measurements of continuum-selected galaxies, we can infer how representative LAEs are of the overall population of galaxies at different redshifts.

This Chapter is structured as follows: in Section 7.2, we give a brief summary of the SC4K sample of LAEs, together with some galaxy properties derived in previous Chapters. We present our methodology to derive UV LFs and SMFs in Section 7.3. We present and discuss our results in Section 7.4, probing the evolution of the UV LF and SMF parameters across time, as well as estimating

the evolution of $\Phi_{\mathrm{LAE}}/\Phi_{\mathrm{LBG}}$ (proxy of $\chi_{\mathrm{Ly\alpha}}$) and the luminosity and stellar mass densities. We present our conclusions in Section 7.5.

7.2 Sample and properties

7.2.1 SC4K sample of LAEs

The public SC4K sample of LAEs (Slicing COSMOS with 4k LAEs, Chapter 2) consists of 3908 LAEs selected with 12+4 medium+narrow band (MB+NB) filters (see Table 7.1 for an overview) over the 2 deg^2 of the COSMOS field (Capak et al., 2007; Scoville et al., 2007; Taniguchi et al., 2015a). For full details on the selection criteria applied, we refer the reader to Section 2.3. Briefly, LAEs are selected based on 1) Lyα EW$_0$ > 50 Å (25 Å for NBs and 5 Å for NB392); 2) significant excess emission (Σ > 3; see Bunker et al., 1995); 3) colour break blueward of the Lyα emission; 4) exclusion of sources with strong red colours (prevents lower redshift interlopers with strong Balmer breaks); 5) full visual inspection to remove spurious detections.

Multiple studies have used the SC4K sample to derive properties of LAEs. Paulino-Afonso et al. (2018) and Shibuya et al. (2019) find small UV sizes with little evolution from $z \sim 2$ to $z \sim 6$. Clustering analysis reveals dark matter halo masses strongly depend on the Lyα luminosity (L$_{\mathrm{Ly\alpha}}$, Khostovan et al., 2019). Calhau et al. (2020) analysed X-ray and radio data on the COSMOS field and measured a low (< 10%) overall AGN fraction, dependent on L$_{\mathrm{Ly\alpha}}$, significantly increasing with increasing luminosity and approaching 100% at L$_{\mathrm{Ly\alpha}}$ > 10^{44} erg s^{-1}. SED fitting from Chapter 5 shows that SC4K LAEs are typically very blue ($\beta = -2.1$), low mass (M$_\star = 10^{9.3}$ M$_\odot$), and above the star-forming Main Sequence at $z < 4$ and M$_\star < 10^{9.5}$ M$_\odot$. SC4K sources are also the prime focus of follow-up spectroscopic observations focusing on studying primeval galaxies (Amorín et al., 2017).

Table 7.1: Overview of the SC4K sample of LAEs used in this study (summary table of Tables 2.3 and 5.1). Given values are the median of all measurements for each galaxy property, with the errors being the 16th and 84th percentile of the distribution. (1) LAE selection filter (see Table 2.3); (2) Redshift range the filter is sensitive to Lyα emission, based on the filter FWHM; (3) Number of LAEs; (4) Comoving volume probed by each filter; (5) Median-likelihood stellar mass parameter from SED fitting, see Section 7.2.2.2; (6) UV magnitude computed by integrating the SED at $\lambda_0 = 1500\,\text{Å}$, see Section 7.2.2.1.

(1) Filter	(2) Lyα z	(3) # LAEs	(4) Volume ($\times 10^6$ Mpc^3)	(5) M$_\star$ ($\log_{10}$ (M$_\star$/M$_\odot$))	(6) M$_{\mathrm{UV}}$ (AB)
NB392	$2.20 - 2.24$	159	0.6	$9.5^{+0.5}_{-0.6}$	$-19.6^{+1.0}_{-0.6}$
IA427	$2.42 - 2.59$	741	4.0	$9.2^{+0.5}_{-0.5}$	$-19.7^{+0.6}_{-0.6}$
IA464	$2.72 - 2.90$	311	4.2	$9.1^{+0.6}_{-0.3}$	$-20.2^{+0.5}_{-0.5}$
IA484	$2.89 - 3.08$	711	4.3	$9.0^{+0.7}_{-0.3}$	$-20.0^{+0.6}_{-0.7}$
NB501	$3.08 - 3.16$	45	0.9	$9.6^{+0.4}_{-0.5}$	$-20.4^{+1.1}_{-0.8}$
IA505	$3.07 - 3.26$	483	4.3	$9.4^{+0.5}_{-0.5}$	$-20.2^{+0.6}_{-0.6}$
IA527	$3.23 - 3.43$	641	4.5	$9.4^{+0.6}_{-0.6}$	$-20.2^{+0.5}_{-0.6}$
IA574	$3.63 - 3.85$	98	4.9	$9.3^{+0.7}_{-0.2}$	$-20.8^{+0.5}_{-0.4}$
IA624	$4.00 - 4.25$	142	5.2	$9.2^{+0.5}_{-0.5}$	$-20.5^{+0.5}_{-0.6}$
IA679	$4.44 - 4.72$	79	5.5	$9.5^{+0.8}_{-0.3}$	$-21.2^{+0.6}_{-0.5}$
IA709	$4.69 - 4.95$	81	5.1	$9.4^{+0.5}_{-0.3}$	$-21.1^{+0.5}_{-0.4}$
NB711	$4.83 - 4.89$	78	1.2	$9.7^{+0.6}_{-0.6}$	$-20.9^{+0.5}_{-0.8}$
IA738	$4.92 - 5.19$	79	5.1	$9.6^{+0.7}_{-0.3}$	$-21.3^{+0.4}_{-0.7}$
IA767	$5.17 - 5.47$	33	5.5	$9.7^{+0.3}_{-0.4}$	$-21.6^{+0.4}_{-0.5}$
NB816	$5.65 - 5.75$	192	1.8	$9.9^{+0.4}_{-0.5}$	$-21.4^{+0.6}_{-0.6}$
IA827	$5.64 - 5.92$	35	4.9	$9.9^{+0.6}_{-0.4}$	$-22.0^{+0.8}_{-1.0}$
Full SC4K	2.20-5.92	3908	62.0	$9.3^{+0.6}_{-0.5}$	$-20.2^{+0.7}_{-0.8}$
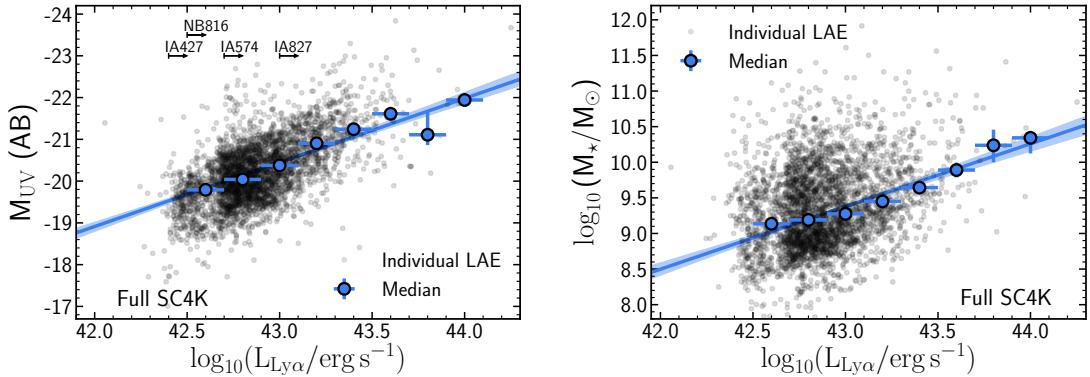
Figure 7.1: *Left:*
M_{UV} dependence on $L_{Ly\alpha}$ within our sample of LAEs. Individual measurements are plotted as scatter in the background. We calculate the median M_{UV} per $L_{Ly\alpha}$ bin (blue circles), with the error bars being the 16th and 84th percentile of the M_{UV} distribution divided by $\sqrt{N}$, with N being the number of sources inside the bin. Bins are defined with 0.2 bin width, starting at $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) = 42.5$, which corresponds to the 3σ $L_{Ly\alpha}$ limit for the MB at $z = 2.5$. The blue shaded contour is the 16th and 84th percentiles of 1000 iterations of fits, obtained by perturbing the median bins within their asymmetric error bars. We find M_{UV} and $L_{Ly\alpha}$ to be well correlated (best-fit $M_{UV} = -1.6^{+0.2}_{-0.3} \log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) + 47^{+12}_{-11}$) in our sample of LAEs, with bright M_{UV} typically corresponding to bright $L_{Ly\alpha}$, but with an important scatter. There is a clear and gradual median brightening at $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) = 42.5 - 43.5$, from -19.8 to -21.4. The higher number of sources above $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \approx 42.7$ (also observed in the right panel) is a consequence of flux limit differences between narrow and medium bands. For reference, we show the 3σ $L_{Ly\alpha}$ limits for the IA427 (MB, $z = 2.5$), IA574 (MB, $z = 3.7$), IA827 (MB, $z = 5.7$) and NB816 (NB, $z = 5.7$) samples. *Right:* Same but for stellar mass ($M_\star$). We also find a correlation between $M_\star$ and $L_{Ly\alpha}$ (best-fit $\log_{10}(M_\star/M_\odot) = 0.9^{+0.1}_{-0.1} \log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) - 28^{+4.0}_{-3.8}$), which is shallower than the correlation found for M_{UV}, revealing how a stellar mass selection and a $Ly\alpha$ selection can differ. The median evolution is less evident than in the left panel, with the median $\log_{10}(M_\star/M_\odot)$ only increasing by 0.2 in the luminosity range $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) = 42.5 - 43.0$.

7.2.1.1 X-ray and radio AGN in SC4K

The SC4K sample includes 254 LAEs detected in X-ray and 120 detected in radio (56 in both), resulting in 318 AGN candidates (Calhau et al., 2020) out of 3705 SC4K LAEs with X-ray or radio coverage. Following the same methodology as in Section 5.2.1, we classify these sources as AGNs since pure star-forming processes would require extremely high SFRs ($\gtrsim 1000\,M_\odot\,\mathrm{yr}^{-1}$) to be detectable at such wavelengths and redshifts. Throughout this work, SC4K AGNs are removed from any fitting/binning and median values in tables, except in Figures 7.3 and 7.5 (left panel), where we show the number densities of AGN LAEs.

7.2.1.2 Redshift binning

In addition to analysing the properties of LAEs from specific selection filters, we group filters with similar central wavelengths to analyse specific redshift bins in a more statistically robust way. We use a grouping scheme similar to Sections 3.1.5 and 5.2.2:

- $z = 2.5 \pm 0.1$ (IA427);

- $z = 3.1 \pm 0.4$ (IA464, IA484, NB501, IA505, IA527);

- $z = 3.9 \pm 0.3$ (IA574, IA624);

- $z = 4.7 \pm 0.2$ (IA679, IA709, NB711);

- $z = 5.4 \pm 0.5$ (IA738, IA767, NB816, IA827).

Here, we include NBs in the redshift bins (even though they typically reach fainter Lyα luminosities) as we perform Lyα luminosity cuts to ensure the samples are directly comparable (see Section 7.4.3 and Section 7.4.6).

7.2.2 Spectral energy distribution and properties of SC4K LAEs

Spectral energy distribution (SED) fitting of the full SC4K sample is presented in Section 5.6. Briefly, SED-fitting is done using the publicly available SED-fitting
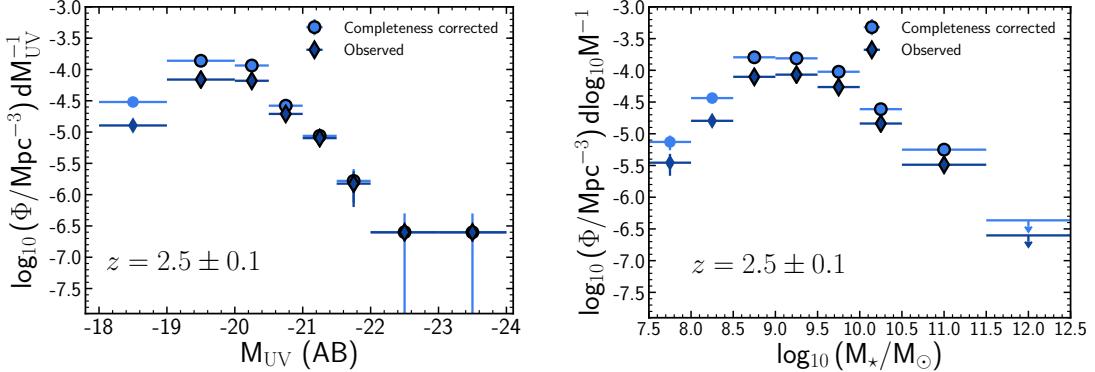
Figure 7.2: *Left:* The rest-frame UV LF of the $z = 2.5$ (IA427) sample of LAEs. We show the luminosity values before (dark blue diamonds) and after (blue circles) applying the completeness correction. The completeness correction is based on Lyα flux (selection criteria). Since M_{UV} and $L_{Ly\alpha}$ typically correlate (see Figure 7.1, left panel), the completeness corrections are larger for the faintest M_{UV} bins. *Right:* Same but for the SMF. As the correlation between $L_{Ly\alpha}$ and $M_\star$ is shallower (see Figure 7.1, right panel), completeness corrections are not a strong function of stellar mass for a specific Lyα cut.

code MAGPHYS[1] (da Cunha et al., 2008, 2012) with the high-redshift extension (see da Cunha et al., 2015), which models stellar and dust emission from galaxies. We obtain photometric measurements from publicly available imaging, taken with 34 rest-frame UV-FIR filters in the COSMOS field (Capak et al., 2007; Lutz et al., 2011; McCracken et al., 2012; Oliver et al., 2012; Sanders et al., 2007; Steinhardt et al., 2014). As the SED-fitting code does not include nebular emission, we exclude the NB or MB with observed Lyα emission from the SED-fitting. Derived parameters are the median-likelihood parameters obtained by comparing modelled SEDs with libraries of galaxies at similar redshift. MAGPHYS uses dust attenuation models from Charlot & Fall (2000) and the stellar population synthesis model from Bruzual & Charlot (2003) with a Chabrier (2003) IMF (range 0.1-100 $M_\odot$). The prescription of Madau (1995) is used to model the intergalactic medium (IGM).

In this work, we will focus on two SED-derived properties: rest-frame UV luminosity (M_{UV}) and stellar mass ($M_\star$). We use the public catalogues provided by

---

[1]http://www.iap.fr/magphys/

Santos et al. (2020) (see Appendix A.4), which contain coordinates, photometry and derived galaxy properties for the full SC4K sample of LAEs.

7.2.2.1   Rest-frame UV luminosity (M_{UV})

The UV luminosity of a galaxy (M_{UV}) can be used as a tracer of recent star-formation on $\sim 100\,\mathrm{Myr}$ timescales (e.g. Boselli et al., 2001; Salim et al., 2009). M_{UV} is computed in Section 6.1.3 by integrating the best-fit SEDs at rest-frame $\lambda_0 = 1400 - 1600\,\text{\AA}$. The median of the SC4K sample is $M_{UV} = -20.2^{+0.7}_{-0.8}$ (Table 7.1), which corresponds to $0.09 \times L^*_{z=3}$ (Steidel et al., 1999).

Similarly but for shorter timescales, Lyα emission also traces recent star-formation, due to being a tracer of Lyman-Continuum (e.g. Sobral & Matthee, 2019) like Hα (Kennicutt, 1998). As the massive young stars responsible for producing the UV continuum also produce the ionising photons that lead to Lyα emission, we can expect these two properties to be related. For our sample of LAEs, we observe that these two properties are typically correlated (see Figure 7.1, left panel), with the median M_{UV} significantly brightening from -19.8 to -21.4 in the luminosity range $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) = 42.5 - 43.5$. We compute a best-fit of $M_{UV} = -1.6^{+0.2}_{-0.3} \log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) + 47^{+12}_{-11}$ from the median distribution. However, Lyα luminosity ($L_{Ly\alpha}$) does not necessarily translate to M_{UV} and vice-versa (see the scatter around $M_{UV} = -20$ and see discussion in Matthee et al. 2017c and Sobral et al. 2018b). This is also made evident from LBG samples, where there are bright M_{UV} sources with no significant Lyα detection, as shown by the Lyα fraction (e.g. Arrabal Haro et al., 2018; Kusakabe et al., 2020; Pentericci et al., 2011; Stark et al., 2010). Furthermore, earlier works (e.g. Shapley et al., 2003) show that LBGs only have Lyα emission $\sim 50\%$ of the time, with a significant number of LBGs even showing Lyα absorption (e.g. DLAs).

7.2.2.2   Stellar Mass ($M_\star$)

The shape and normalisation of an SED is a reflection of the content of stars in a galaxy, thus its total mass of stars (stellar mass, $M_\star$) can be derived by fitting and modelling the SED. LAEs typically have low $M_\star$ but there is an important diversity within the population. The median of the SC4K sample

of LAEs is computed in Section 6.1.5 using MAGPHYS: $\log_{10}(M_\star/M_\odot) = 9.3^{+0.6}_{-0.5}$ (Table 7.1), which corresponds to $0.006 \times M^*_{\star,z=3-4}$ (Muzzin et al., 2013). We find the median $M_\star$ and $L_{Ly\alpha}$ to be correlated (see Figure 7.1, right panel), with best-fit $\log_{10}(M_\star/M_\odot) = 0.9^{+0.1}_{-0.1} \log_{10}(L_{Ly\alpha}/erg\,s^{-1}) - 28^{+4.0}_{-3.8}$, but with a significant scatter of individual detections. This relation is shallower than the one measured between M_{UV} and $L_{Ly\alpha}$, with a more modest increase: the median $\log_{10}(M_\star/M_\odot)$ only increases by 0.2 in the luminosity range $\log_{10}(L_{Ly\alpha}/erg\,s^{-1}) = 42.5 - 43.0$. We note that the SED-fitting used to derive $M_\star$ does not include nebular emission, so the two properties are independently derived. Additionally, there is an anti-correlation between $M_\star$ and $Ly\alpha$ EW_0, and thus $Ly\alpha$ escape fraction of LAEs (Section 6.2.3.2).

7.3 Luminosity and stellar mass functions

In this section, we present our methodology and computations to derive UV LFs and SMFs for our sample of ~ 4000 LAEs at well-defined redshift intervals between $z \sim 2$ and $z \sim 6$.

7.3.1 Determining the luminosity/mass functions

We measure the number densities of well-defined M_{UV} and $M_\star$ bins which we use to construct the UV LF and SMF. We choose bin widths depending on M_{UV} and $M_\star$, as the most (and least) luminous and massive bins have fewer sources. We define bins with width 0.5 dex in the range $-22 < M_{UV} < -20$ ($M_{UV} > -22.5$ for the deeper NB816) and $7 < \log_{10}(M_\star/M_\odot) < 10.5$ and 1 dex outside these ranges, where the number densities are the lowest. We use Poissonian errors for any individual LF realisation.

We define the number density of a luminosity bin as:

$$\log_{10}(\phi_j) = \log_{10}\left(\frac{1}{d\log_{10} L}\frac{N_j}{V}\right), \tag{7.1}$$

where ϕ_j is the number density of a bin j, N_j is the number of sources within dlog$_{10}$L of j, and V is the volume probed by the NBs or MBs for that specific bin (see Table 7.1), which is computed from the redshift range that each filter is sensitive to Lyα emission.

7.3.2 Completeness correction

Faint sources and those with low Lyα EW may be missed by our selection criteria, leading to an underestimation of number densities. We estimate completeness corrections based on Lyα line flux (same corrections used for the Lyα LFs in Section 3.1.2.1) and apply them to the UV LFs and the SMFs of our sample of LAEs. Briefly, for each NB or MB, we obtain a sample of high-redshift non-line-emitters by applying the same colour break we used to target the Lyman break in our LAEs and by selecting sources with photometric redshifts (obtained from Laigle et al., 2016) ± 0.2 the redshift range given in Table 7.1. For each non-line-emitter sample, flux is incrementally added to the NB or MB and BB (see Table 2.3). By reapplying our selection criteria after each step, we determine the fraction of galaxies which are picked as emitters per Lyα luminosity value. We only consider sources with $> 30\%$ completeness.

We apply completeness corrections to each LAE individually, based on their observed Lyα flux, and not their M_{UV} or $M_\star$. In Figure 7.2, we show M_{UV} and $M_\star$ number densities for $z = 2.5$ (IA427) LAEs, before and after completeness corrections. We note that the completeness correction is based on Lyα flux and thus larger for fainter LAEs but not necessarily correlated with other properties. Since M_{UV} and $L_{Ly\alpha}$ typically correlate (see Figure 7.1, left panel), the corrections will typically be smaller for LAEs which are brighter in M_{UV} (see Figure 7.2, left panel). Since the correlation between $M_\star$ and $L_{Ly\alpha}$ is weak (see Figure 7.1, right panel), the corrections will be similar for the entire mass range (see Figure 7.2, right panel).

Including the completeness correction, applied to each source, Equation 7.1 becomes:

$$\log_{10}(\phi_j) = \log\left(\frac{1}{d\log_{10}L}\sum_i^{N_j}\frac{c_i}{V}\right), \qquad (7.2)$$

where c_i is the completeness correction for a source i.

For the luminosity or stellar mass bins with zero counts, we compute the upper limit as one source at the volume probed by the NB or MB, with the completeness correction equal to the total completeness correction applied to the previous luminosity or stellar mass bin.

7.3.3 Fitting the UV luminosity function

In order to compare our results with previous studies, we adopt the common parameterisation of Schechter (1976) function, which consists of a power-law with a slope α for faint luminosities and a declining exponential for brighter luminosities. The transition between the two regimes is given by a characteristic luminosity (L*) and a characteristic number density (Φ^*). The Schechter equation has the following form:

$$\Phi(\mathrm{L}) = \frac{\Phi^*}{\mathrm{L}^*}\left(\frac{\mathrm{L}}{\mathrm{L}^*}\right)^{\alpha}\exp\left(-\frac{\mathrm{L}}{\mathrm{L}^*}\right), \qquad (7.3)$$

Equation 7.3.3 can be rewritten for absolute magnitudes by using the substitution $\Phi(L)\mathrm{d}L = \Phi(\mathrm{M_{UV}})\mathrm{dM_{UV}}$:

$$\Phi(\mathrm{M_{UV}}) = \frac{\ln 10}{2.5}\Phi^* 10^{-0.4(\alpha+1)\Delta \mathrm{M_{UV}}}\exp\left(-10^{-0.4\Delta \mathrm{M_{UV}}}\right), \qquad (7.4)$$

where $\Delta \mathrm{M_{UV}} = \mathrm{M_{UV}} - \mathrm{M_{UV}^*}$.

The observed UV luminosity distribution of LAEs shows the same behaviour at all redshifts: there is a peak number density at an intermediate UV luminosity,
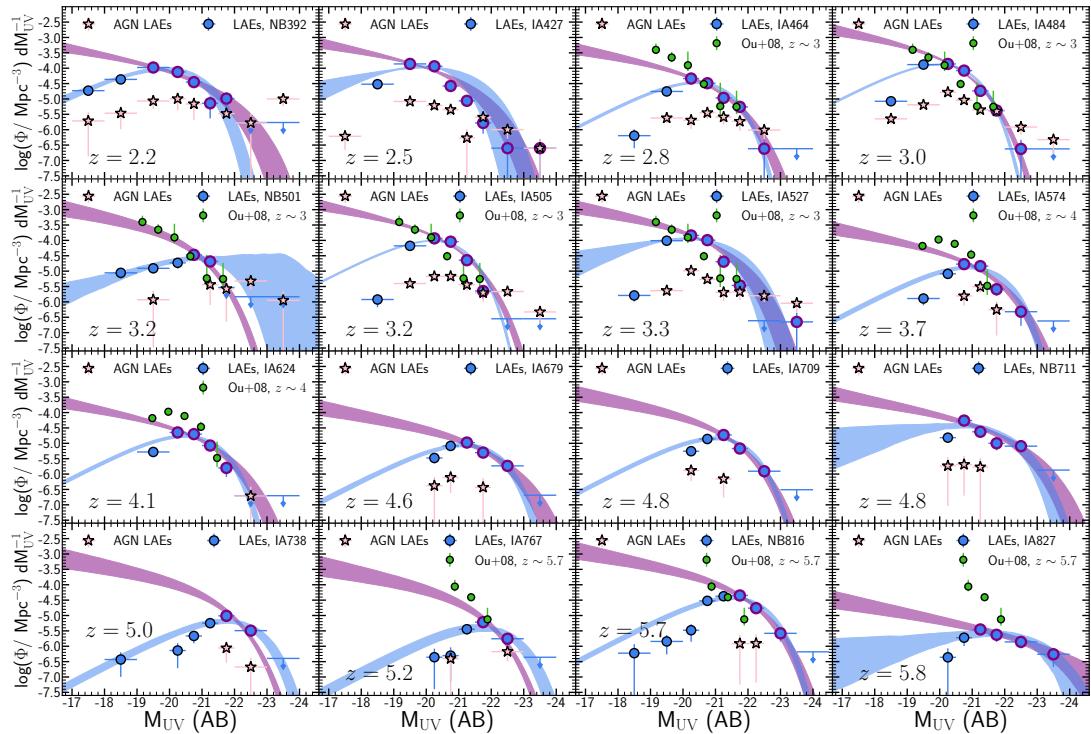
Figure 7.3: The rest-frame UV LF (blue circles) for each of the 16 individual selection filters in this study, without any Lyα flux cut and excluding AGNs (see Section 7.2.1.1). Luminosity bins brighter than the peak of number densities are marked with a purple edge colour. The blue contours are the 16th and 84th percentile of multiple iterations of fits to the luminosity bins, obtained by perturbing the luminosity bins within their error bars (see Section 7.3.5) for the full UV luminosity range. The purple contours represent the same but only fitting the points above the number density peak. We compare our results with those of Ouchi et al. (2008) at $z \sim 3, 4, 5.7$, finding a good agreement, with the offset at $z = 5.8$ being easily explained by differences in Lyα flux limits (see Section 7.4.3). We show the number densities of AGNs (pink stars), which predominantly dominate the bright M_{UV} regime (< -22) at $z < 4$, often having higher number densities than non-AGN LAEs. At $z > 4$ there are significantly less AGN LAEs, composing only a small fraction of LAEs at all M_{UV} ranges.

with a subsequent decline in number density for both brighter and fainter UV luminosities (see Figure 7.3). While such a distribution does not resemble the Schechter function with a steep faint end which is typically measured in LBG samples (e.g. Bouwens et al., 2015; Finkelstein et al., 2015), we argue that such observed distribution of UV luminosities can be expected for a sample which is selected by Lyα line flux above some threshold (corresponding to a vertical cut in Figure 7.1), causing an incomplete sampling of M_{UV}. This incomplete sampling is most significant at the faint UV luminosities, which is shown in Figure 7.5 (right panel) where an increasing $L_{Ly\alpha}$ limit will cause a preferential decline of number densities at faint UV luminosities and hence the observed turn-over. Thus, in order to conduct a detailed analysis of the UV luminosity distribution of LAEs, we explore two separate scenarios:

- fit to the full UV luminosity range (blue in Figure 7.3): the entire observable UV luminosity range is considered, including the turn-over at faint UV luminosities. While the low number densities at faint UV luminosities may be driven by our $L_{Ly\alpha}$ limits, this approach provides the best-fit to the directly observed number densities.

- fit to the UV luminosity range brighter than the number density peak (purple in Figure 7.3): the bins fainter than the number density peak (dominated by an incomplete sampling) are thus not included in the fitting, and the faint UV luminosity regime becomes unconstrained. The peak in number density is different for different filters (see Figure 7.3) and different $L_{Ly\alpha}$ limits (see Figure 7.5, right panel). With the simple assumptions of a steep faint end slope (as measured in UV luminosity-selected samples) and by not including the bins below of the turn-over (which are heavily dominated by our $L_{Ly\alpha}$ limits), we obtain a proxy of the full distribution of LAEs.

We provide the Schechter parameters of the best-fits to both cases in Table 7.2. For the fit to the full UV luminosity range, we find the set of parameters (α, M^*_{UV}, Φ^*) which minimises the reduced χ^2 (χ^2_{red}) in log-space. Alternatively, fitting can also be performed in linear-space, where χ^2 is less sensitive to bins with low number densities. A fit in log-space thus tends to favour slightly brighter

Table 7.2: Best-fit Schechter parameters for the UV LF of LAEs from $z = 2$ to $z = 6$, for each of the individual selection filters and for different redshift bins (see Section 7.2.1.2). The number of sources provided here is the number of sources included in the luminosity functions, i.e. non-AGN LAEs with available SEDs and with completeness $> 30\%$. We provide best fits for the two cases considered in this study: fit to the full UV luminosity range (blue in Figure 7.3) and fit to the bins brighter than the number density peak (purple in Figure 7.3). We provide the best set of parameters (α, M^*_{UV} and Φ^*) which minimise χ^2_{red}, with α being fixed for the latter case as it cannot be directly constrained. When χ^2_{red} is very large, the errors should be interpreted with caution as the best parameters found still do not provide a good model. Additionally, M^*_{UV} is also fixed for the individual filters with less than three luminosity bins (although we perturb these parameters when exploring the uncertainties of the bins/fits, see Section 7.3.5). For the redshift bins we also show the Schechter parameters when applying a $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s}^{-1}) \geq 43.0$ cut.

Redshift	# Filters	# Sources	Full UV range				UV brighter than the peak ($\alpha_{fix} = -1.5$)		
			$\log_{10}\Phi^*$ (Mpc^{-3})	M^*_{UV} (AB)	α	χ^2_{red}	$\log_{10}\Phi^*$ (Mpc^{-3})	M^*_{UV} (AB)	χ^2_{red}
2.2 ± 0.1	1	129	$-3.57^{+0.03}_{-0.04}$	$-19.78^{+0.16}_{-0.17}$	$-0.7^{+0.1}_{-0.0}$	2.5	$-4.16^{+0.11}_{-0.15}$	$-21.15^{+0.26}_{-0.51}$	0.9
2.5 ± 0.1	1	519	$-3.48^{+0.01}_{-0.01}$	$-19.46^{+0.03}_{-0.04}$	$-0.9^{+0.0}_{-0.0}$	33.9	$-3.86^{+0.05}_{-0.06}$	$-20.65^{+0.11}_{-0.15}$	14.0
2.8 ± 0.1	1	139	$-4.04^{+0.03}_{-0.02}$	$-20.02^{+0.06}_{-0.06}$	$-1.5^{+0.0}_{-0.0}$	9.4	$-4.19^{+0.09}_{-0.09}$	$-20.95^{+0.15}_{-0.18}$	1.0
3.0 ± 0.1	1	565	$-3.43^{+0.01}_{-0.02}$	$-19.61^{+0.03}_{-0.03}$	$-1.4^{+0.0}_{-0.0}$	36.8	$-3.43^{+0.06}_{-0.06}$	$-20.37^{+0.07}_{-0.09}$	3.6
3.2 ± 0.1	1	31	$-4.11^{+0.61}_{-0.09}$	$-21.09^{+0.49}_{-3.20}$	$-0.3^{+0.1}_{-1.1}$	1.2	$-3.65^{+0.04}_{-0.04}$	-20.37 (fix)	1.7
3.2 ± 0.1	1	413	$-3.59^{+0.01}_{-0.01}$	$-19.82^{+0.03}_{-0.03}$	$-0.7^{+0.0}_{-0.0}$	29.2	$-3.58^{+0.05}_{-0.05}$	$-20.58^{+0.08}_{-0.10}$	13.6
3.3 ± 0.1	1	565	$-3.47^{+0.01}_{-0.01}$	$-19.77^{+0.03}_{-0.03}$	$-0.9^{+0.0}_{-0.0}$	36.4	$-3.49^{+0.06}_{-0.05}$	$-20.54^{+0.08}_{-0.09}$	12.3
3.7 ± 0.1	1	53	$-4.45^{+0.03}_{-0.03}$	$-20.51^{+0.09}_{-0.12}$	$-1.3^{+0.1}_{-0.1}$	8.5	$-4.45^{+0.11}_{-0.12}$	$-21.17^{+0.16}_{-0.23}$	3.2
4.1 ± 0.1	1	116	$-4.33^{+0.03}_{-0.02}$	$-20.10^{+0.07}_{-0.12}$	$-0.8^{+0.0}_{-0.1}$	13.4	$-4.50^{+0.10}_{-0.12}$	$-21.07^{+0.18}_{-0.27}$	2.9
4.6 ± 0.1	1	69	$-4.64^{+0.03}_{-0.03}$	$-20.84^{+0.06}_{-0.15}$	$-1.3^{+0.0}_{-0.1}$	2.6	$-4.85^{+0.09}_{-0.11}$	$-21.90^{+0.14}_{-0.21}$	0.4
4.8 ± 0.1	1	50	$-4.42^{+0.03}_{-0.03}$	$-20.63^{+0.06}_{-0.08}$	$-0.9^{+0.1}_{-0.1}$	6.3	$-4.26^{+0.08}_{-0.08}$	$-21.22^{+0.09}_{-0.12}$	0.5
4.8 ± 0.1	1	41	$-4.04^{+0.07}_{-0.35}$	$-21.30^{+0.31}_{-0.72}$	$-0.7^{+0.1}_{-0.1}$	8.4	$-4.44^{+0.10}_{-0.10}$	$-22.10^{+0.20}_{-0.33}$	2.2
5.0 ± 0.1	1	29	$-4.74^{+0.04}_{-0.04}$	$-21.33^{+0.12}_{-0.20}$	$-0.4^{+0.1}_{-0.1}$	3.0	$-4.02^{+0.03}_{-0.03}$	-21.22 (fix)	7.0
5.2 ± 0.1	1	17	$-4.91^{+0.06}_{-0.05}$	$-21.19^{+0.13}_{-0.24}$	$0.0^{+0.1}_{-0.1}$	2.9	$-4.25^{+0.04}_{-0.03}$	-21.22 (fix)	3.0
5.7 ± 0.1	1	107	$-3.98^{+0.03}_{-0.04}$	$-21.09^{+0.09}_{-0.12}$	$-1.0^{+0.1}_{-0.1}$	3.9	$-3.84^{+0.08}_{-0.10}$	$-21.67^{+0.10}_{-0.13}$	0.1
5.8 ± 0.1	1	14	$-5.27^{+0.14}_{-0.82}$	$-22.43^{+0.46}_{-2.53}$	$-0.5^{+0.1}_{-0.1}$	0.7	$-5.85^{+0.18}_{-0.25}$	$-23.56^{+0.45}_{-0.72}$	0.0
2.5 ± 0.1	1	519	$-3.48^{+0.01}_{-0.01}$	$-19.46^{+0.03}_{-0.04}$	$-0.9^{+0.0}_{-0.0}$	33.9	$-3.86^{+0.05}_{-0.06}$	$-20.65^{+0.11}_{-0.15}$	14.0
3.1 ± 0.4	5	1713	$-3.60^{+0.01}_{-0.00}$	$-19.77^{+0.02}_{-0.02}$	$-1.1^{+0.0}_{-0.0}$	86.8	$-3.59^{+0.03}_{-0.04}$	$-20.50^{+0.04}_{-0.06}$	17.6
3.9 ± 0.3	2	169	$-4.32^{+0.03}_{-0.02}$	$-20.10^{+0.04}_{-0.11}$	$-0.8^{+0.0}_{-0.0}$	3.3	$-4.30^{+0.09}_{-0.08}$	$-20.87^{+0.11}_{-0.13}$	2.4
4.7 ± 0.2	3	160	$-4.44^{+0.02}_{-0.02}$	$-20.73^{+0.03}_{-0.06}$	$-0.9^{+0.1}_{-0.0}$	11.6	$-4.62^{+0.05}_{-0.06}$	$-21.73^{+0.08}_{-0.09}$	2.5
5.4 ± 0.5	4	167	$-4.69^{+0.03}_{-0.03}$	$-21.26^{+0.07}_{-0.07}$	$-0.9^{+0.1}_{-0.0}$	8.5	$-4.62^{+0.05}_{-0.06}$	$-21.92^{+0.07}_{-0.09}$	0.0
Full SC4K	16	2857	$-3.98^{+0.00}_{-0.01}$	$-20.45^{+0.01}_{-0.04}$	$-1.1^{+0.0}_{-0.0}$	191.3	$-4.39^{+0.02}_{-0.02}$	$-21.36^{+0.03}_{-0.04}$	14.6
$\log_{10}(L_{Ly\alpha}) \geq 43.0\,\mathrm{erg\,s}^{-1}$									
2.5 ± 0.1	1	47	$-4.79^{+0.09}_{-0.20}$	$-21.03^{+0.36}_{-1.98}$	$-0.1^{+0.0}_{-0.1}$	1.2	$-5.60^{+0.30}_{-0.31}$	$-22.94^{+0.83}_{-1.03}$	0.8
3.1 ± 0.4	5	411	$-4.43^{+0.02}_{-0.02}$	$-20.32^{+0.05}_{-0.06}$	$0.4^{+0.0}_{-0.1}$	13.5	$-4.84^{+0.08}_{-0.09}$	$-21.70^{+0.18}_{-0.25}$	5.5
3.9 ± 0.3	2	107	$-4.70^{+0.03}_{-0.03}$	$-20.40^{+0.07}_{-0.09}$	$0.4^{+0.0}_{-0.1}$	5.1	$-4.71^{+0.10}_{-0.11}$	$-21.12^{+0.14}_{-0.20}$	0.5
4.7 ± 0.2	3	132	$-4.54^{+0.02}_{-0.02}$	$-20.75^{+0.04}_{-0.05}$	$0.4^{+0.0}_{-0.1}$	8.8	$-4.61^{+0.05}_{-0.06}$	$-21.60^{+0.08}_{-0.10}$	1.4
5.4 ± 0.5	4	91	$-4.88^{+0.04}_{-0.03}$	$-21.33^{+0.08}_{-0.11}$	$0.4^{+0.0}_{-0.1}$	5.3	$-4.85^{+0.06}_{-0.07}$	$-22.03^{+0.09}_{-0.11}$	0.1
Full SC4K	16	789	$-4.65^{+0.02}_{-0.01}$	$-20.79^{+0.03}_{-0.04}$	$0.4^{+0.0}_{-0.1}$	27.8	$-5.03^{+0.03}_{-0.04}$	$-21.95^{+0.07}_{-0.08}$	7.7

characteristic luminosities which provide a better fit to the very bright luminosities. We find that the observed distribution is best fit by shallow faint end slopes ($\alpha > -1$), which are able to represent the turn-over at the faintest luminosities, with α even being positive at some redshift ranges.

When constraining only the UV luminosities brighter than the number density peak, we are not able to directly constrain the α slope of the power-law, and thus fix α to -1.5 (similar to the UV LF of LAEs from e.g. Ouchi et al., 2008), but we still perturb this parameter to quantify uncertainties (see Section 7.3.5). Here, we make the assumption that α does not evolve with redshift, which is a necessary caveat due to not being able to directly constrain it. We measure the UV LF of LAEs selected in each MB or NB by determining the pair (M^*_{UV}, Φ^*) which minimises χ^2_{red} in log-space of the M_{UV} luminosity bins with associated Poissonian error bars. In Figure 7.3, we show the luminosity bins and luminosity functions of LAEs from the 16 selection filters. For the filters with only two luminosity bins brighter than the number density peak, we can only fit one free parameter, so we fix M^*_{UV} to a similar nearby filter (NB501 uses $M^*_{UV,IA505} = -20.37$ and IA738+IA767 use $M^*_{UV,IA709} = -21.22$). We provide the Schechter parameters of the best fits in Table 7.2.

7.3.4 Fitting the stellar mass function

Following a similar logic to what was done in Section 7.3.3, Equation 7.3.3 can be rewritten in $\log M$ space:

$$\Phi(M_\star) = \ln 10 \, \Phi^* \, 10^{(\alpha+1)\Delta M} \exp\left(-10^{\Delta M}\right), \qquad (7.5)$$

where $\Delta M = \log_{10} M_\star - \log_{10} M^*_\star$. At $z < 1$, a double Schechter function has been commonly used (see e.g. Ilbert et al., 2013; Pozzetti et al., 2010), with two α and two Φ^*, which are capable of reproducing a bimodal population, which includes quiescent galaxies. In this work, we restrain ourselves to a single Schechter as the quiescent population should not contribute to our Lyα-selected sample, particularly at the redshift range that we probe.

Similarly to the observed UV LF, the observed number density distribution of the stellar mass peaks at an intermediate stellar mass, and declines for both lower and higher stellar masses (see Figure 7.4). While a Schechter distribution with a steep slope could be expected for a mass-selected sample, as our LAEs are selected by being above some Lyα line flux (corresponding to a vertical cut in Figure 7.1) determined by observational constrains, there is a turn-over at low stellar masses. The preferential decline of low stellar masses with increasing Lyα line flux is shown in Figure 7.8 (right panel), and we further discuss how to interpret the shape of the SMF in Section 7.4.1.

Following the same reasons listed for the UV LF, we conduct our fitting procedure in two stellar mass ranges: full stellar mass range (blue in Figure 7.4) and stellar mass range above the turn-over, with an assumption of the α slope (blue in Figure 7.4). The former provides a fit to the directly observed number densities and the later provides a proxy SMF of the full distribution of LAEs. We provide the best Schechter fits to both cases in Table 7.3. For the fit of the full stellar mass range, we find the set of parameters (α, $M_\star^*$, Φ^*) which minimises $\chi^2_{\rm red}$ in log-space. The observed distribution with a turn-over for the smallest stellar masses, results in shallow faint end slopes ($\alpha > -1$).

When constraining only the stellar masses bigger than the number density peak, we are not able to directly constrain the α slope of the power-law. We fix α to -1.3, but we vary all parameters, including α in Section 7.3.5. Similarly to the UV LF, we introduce the caveat that α does not evolve with redshift, which is a necessary assumption due to us not being able to directly constrain it. In Figure 7.4, we show the stellar mass bins and SMFs of LAEs from the 16 selection filters. For the filters with only two stellar mass bins, we can only fit one free parameter, so we fix $M_\star^*$ to a similar nearby filter (NB711+IA767 use $M_{\star,{\rm IA738}}^* = 10^{10.68}\,{\rm M_\odot}$). We provide the Schechter parameters of in best fits in Table 7.3.

7.3.5 Perturbing the luminosity and mass functions

We explore the uncertainties in our UV LFs and SMFs by perturbing the luminosity or mass bins within their Poissonian error bars. For each iteration, we
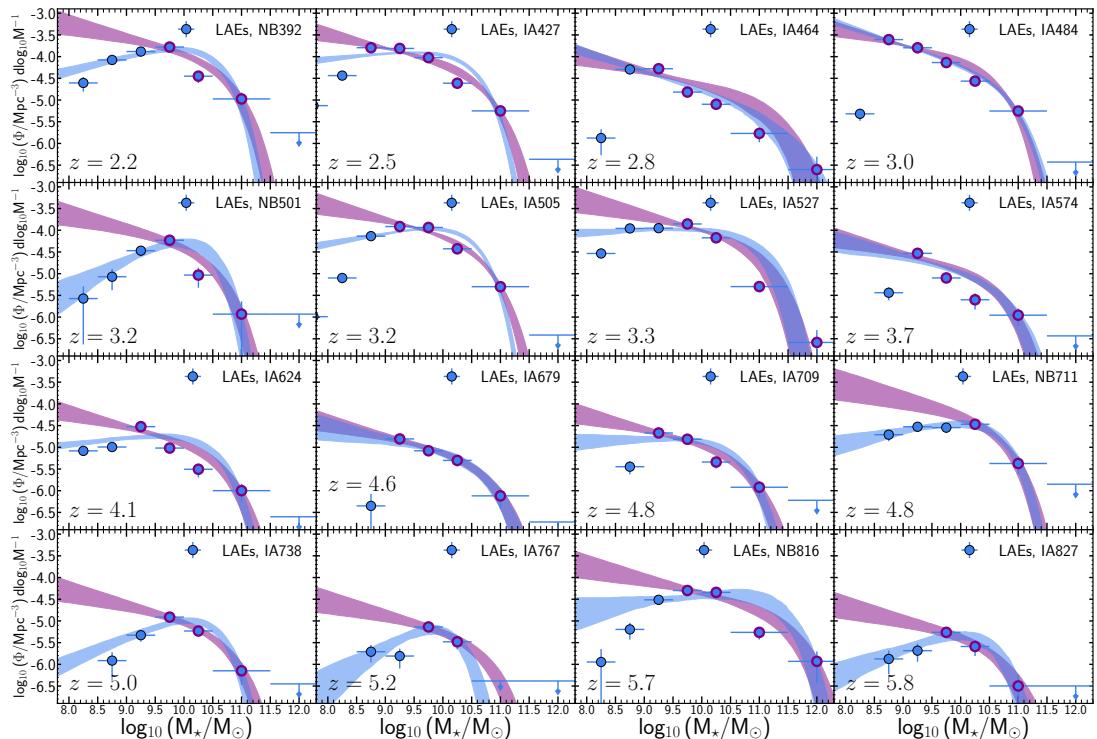
Figure 7.4: The SMF for each of the 16 individual selection filters. Stellar mass bins are shown as blue circles. Stellar mass bins more massive than the peak of number densities are marked with a purple edge colour. The blue contours are the 16th and 84th percentile of multiple iterations of fits to the stellar mass bins, obtained by perturbing the stellar mass bins within their error bars (see Section 7.3.5) for the full stellar mass range. The purple contours represent the same but only fitting the points above the number density peak. Candidate AGN are removed from the analysis here (see Section 7.2.1.1)

Table 7.3: Best-fit Schechter parameters for the SMF of LAEs from $z = 2$ to $z = 6$, for each of the individual selection filters and for different redshift bins (see Section 7.2.1.2). The number of sources provided here is the number of sources included in the stellar mass functions, i.e. non-AGN LAEs with available SEDs and with completeness corrections $> 30\%$. We provide best fits for the two cases considered in this study: fit to the full stellar mass range (blue in Figure 7.4) and fit to the bins more massive than the number density peak (purple in Figure 7.4). We provide the best set of parameters (α, $M_\star^*$ and Φ^*) which minimise $\chi^2_{\rm red}$, with α being fixed for the latter case as it cannot be directly constrained. When $\chi^2_{\rm red}$ is very large, the errors should be interpreted with caution as the best parameters found still do not provide a good model. Additionally, $M_\star^*$ is also fixed for the individual filters with less than three luminosity bins (although we perturb these parameters when exploring the uncertainties of the bins/fits, see Section 7.3.5). For the redshift bins we also show the Schechter parameters when applying a $\log_{10}(L_{\rm Ly\alpha}/{\rm erg\,s^{-1}}) \geq 43.0$ cut.

Redshift	# Filters	# Sources	$\log_{10}\Phi^*$ (Mpc^{-3})	$M_\star^*$ (AB)	α^*	$\chi^2_{\rm red}$	$\log_{10}\Phi^*$ (Mpc^{-3})	$M_\star^*$ (AB)	$\chi^2_{\rm red}$
				Full $M_\star$ range			**$M_\star$ above the peak** ($\alpha_{\rm fix} = -1.3$)		
2.2 ± 0.1	1	129	$-3.94^{+0.08}_{-0.07}$	$10.41^{+0.06}_{-0.06}$	$-0.7^{+0.1}_{-0.0}$	7.8	$-4.44^{+0.04}_{-0.03}$	$10.69^{+0.05}_{-0.05}$	7.7
2.5 ± 0.1	1	519	$-4.14^{+0.01}_{-0.01}$	$10.40^{+0.02}_{-0.02}$	$-0.9^{+0.0}_{-0.0}$	76.2	$-4.66^{+0.02}_{-0.02}$	$10.67^{+0.04}_{-0.03}$	13.4
2.8 ± 0.1	1	139	$-5.90^{+0.20}_{-0.40}$	$11.41^{+0.79}_{-0.19}$	$-1.5^{+0.0}_{-0.0}$	17.6	$-5.33^{+0.04}_{-0.05}$	$11.10^{+0.12}_{-0.08}$	14.1
3.0 ± 0.1	1	565	$-4.79^{+0.02}_{-0.03}$	$10.73^{+0.05}_{-0.03}$	$-1.4^{+0.0}_{-0.0}$	67.9	$-4.57^{+0.02}_{-0.01}$	$10.61^{+0.02}_{-0.02}$	15.4
3.2 ± 0.1	1	31	$-4.25^{+0.09}_{-0.11}$	$10.05^{+0.15}_{-0.09}$	$-0.3^{+0.1}_{-0.1}$	3.6	$-4.68^{+0.07}_{-0.07}$	$10.31^{+0.12}_{-0.07}$	3.1
3.2 ± 0.1	1	413	$-4.03^{+0.01}_{-0.05}$	$10.33^{+0.03}_{-0.02}$	$-0.7^{+0.0}_{-0.0}$	39.4	$-4.62^{+0.02}_{-0.02}$	$10.66^{+0.02}_{-0.03}$	12.8
3.3 ± 0.1	1	565	$-4.19^{+0.01}_{-0.01}$	$10.76^{+0.03}_{-0.03}$	$-0.9^{+0.0}_{-0.0}$	58.0	$-4.60^{+0.02}_{-0.02}$	$10.93^{+0.04}_{-0.03}$	29.7
3.7 ± 0.1	1	53	$-5.50^{+0.16}_{-0.24}$	$10.72^{+0.19}_{-0.10}$	$-1.3^{+0.1}_{-0.1}$	28.6	$-5.40^{+0.04}_{-0.04}$	$10.66^{+0.07}_{-0.06}$	9.9
4.1 ± 0.1	1	116	$-4.86^{+0.03}_{-0.09}$	$10.37^{+0.09}_{-0.04}$	$-0.8^{+0.0}_{-0.0}$	19.0	$-5.38^{+0.04}_{-0.04}$	$10.63^{+0.07}_{-0.05}$	8.4
4.6 ± 0.1	1	69	$-5.65^{+0.11}_{-0.34}$	$10.74^{+0.40}_{-0.10}$	$-1.3^{+0.0}_{-0.0}$	3.7	$-5.64^{+0.05}_{-0.06}$	$10.74^{+0.11}_{-0.08}$	0.4
4.8 ± 0.1	1	50	$-5.04^{+0.07}_{-0.13}$	$10.49^{+0.10}_{-0.05}$	$-0.9^{+0.1}_{-0.1}$	15.8	$-5.47^{+0.04}_{-0.05}$	$10.73^{+0.10}_{-0.06}$	1.9
4.8 ± 0.1	1	41	$-4.53^{+0.07}_{-0.08}$	$10.51^{+0.10}_{-0.06}$	$-0.7^{+0.1}_{-0.1}$	0.3	$-4.79^{+0.05}_{-0.04}$	$10.68\,({\rm fix})$	0.1
5.0 ± 0.1	1	29	$-4.96^{+0.07}_{-0.08}$	$10.32^{+0.07}_{-0.05}$	$-0.4^{+0.1}_{-0.1}$	4.6	$-5.52^{+0.05}_{-0.05}$	$10.68^{+0.08}_{-0.06}$	0.2
5.2 ± 0.1	1	17	$-5.10^{+0.04}_{-0.06}$	$9.83^{+0.11}_{-0.06}$	$0.0^{+0.1}_{-0.1}$	3.4	$-5.75^{+0.04}_{-0.04}$	$10.68\,({\rm fix})$	0.2
5.7 ± 0.1	1	107	$-4.80^{+0.16}_{-0.10}$	$11.38^{+0.11}_{-0.09}$	$-1.0^{+0.1}_{-0.1}$	11.8	$-5.18^{+0.04}_{-0.04}$	$11.54^{+0.12}_{-0.07}$	7.3
5.8 ± 0.1	1	14	$-5.35^{+0.11}_{-0.13}$	$10.34^{+0.14}_{-0.09}$	$-0.5^{+0.1}_{-0.1}$	1.4	$-5.86^{+0.08}_{-0.10}$	$10.66^{+0.17}_{-0.11}$	0.1
2.5 ± 0.1	1	519	$-4.14^{+0.01}_{-0.01}$	$10.40^{+0.02}_{-0.02}$	$-0.9^{+0.0}_{-0.0}$	76.2	$-4.66^{+0.02}_{-0.02}$	$10.67^{+0.04}_{-0.03}$	13.4
3.1 ± 0.4	5	1713	$-4.57^{+0.01}_{-0.01}$	$10.86^{+0.02}_{-0.02}$	$-1.1^{+0.0}_{-0.0}$	193.2	$-4.90^{+0.01}_{-0.01}$	$11.02^{+0.03}_{-0.03}$	55.2
3.9 ± 0.3	2	169	$-4.89^{+0.07}_{-0.02}$	$10.40^{+0.03}_{-0.05}$	$-0.8^{+0.0}_{-0.0}$	47.8	$-5.39^{+0.02}_{-0.03}$	$10.64^{+0.05}_{-0.03}$	18.9
4.7 ± 0.2	3	160	$-5.07^{+0.06}_{-0.02}$	$10.53^{+0.03}_{-0.05}$	$-0.9^{+0.1}_{-0.1}$	25.7	$-5.52^{+0.02}_{-0.04}$	$10.81^{+0.07}_{-0.05}$	0.0
5.4 ± 0.5	4	167	$-5.31^{+0.07}_{-0.09}$	$11.19^{+0.07}_{-0.05}$	$-0.9^{+0.1}_{-0.1}$	26.9	$-5.78^{+0.03}_{-0.04}$	$11.37^{+0.07}_{-0.04}$	14.2
Full SC4K	16	2857	$-4.93^{+0.00}_{-0.01}$	$10.94^{+0.02}_{-0.01}$	$-1.1^{+0.0}_{-0.0}$	372.4	$-5.19^{+0.01}_{-0.00}$	$11.05^{+0.01}_{-0.01}$	95.2
			$\log_{10}(L_{\rm Ly\alpha}) \geq 43.0\,{\rm erg\,s^{-1}}$						
2.5 ± 0.1	1	47	$-4.79^{+0.09}_{-0.20}$	$-21.03^{+0.36}_{-1.98}$	$-0.6^{+0.0}_{-0.0}$	1.2	$-5.60^{+0.30}_{-0.31}$	$-22.94^{+0.83}_{-1.03}$	0.8
3.1 ± 0.4	5	411	$-4.43^{+0.02}_{-0.02}$	$-20.32^{+0.05}_{-0.06}$	$-1.2^{+0.0}_{-0.0}$	13.5	$-4.84^{+0.08}_{-0.09}$	$-21.70^{+0.18}_{-0.25}$	5.5
3.9 ± 0.3	2	107	$-4.70^{+0.03}_{-0.03}$	$-20.40^{+0.07}_{-0.09}$	$-0.8^{+0.0}_{-0.0}$	5.1	$-4.71^{+0.10}_{-0.11}$	$-21.12^{+0.14}_{-0.20}$	0.5
4.7 ± 0.2	3	132	$-4.54^{+0.02}_{-0.02}$	$-20.75^{+0.04}_{-0.05}$	$-1.0^{+0.0}_{-0.0}$	8.8	$-4.61^{+0.05}_{-0.06}$	$-21.60^{+0.08}_{-0.10}$	1.4
5.4 ± 0.5	4	91	$-4.88^{+0.04}_{-0.03}$	$-21.33^{+0.08}_{-0.11}$	$-0.9^{+0.0}_{-0.0}$	5.3	$-4.85^{+0.06}_{-0.07}$	$-22.03^{+0.09}_{-0.11}$	0.1
Full SC4K	16	789	$-4.65^{+0.02}_{-0.01}$	$-20.79^{+0.03}_{-0.04}$	$-1.1^{+0.0}_{-0.0}$	27.8	$-5.03^{+0.03}_{-0.04}$	$-21.95^{+0.07}_{-0.08}$	7.7

perturb each bin within their error bars (assuming a normal probability distribution function centred at each bin and with FWHM equal to the error) and determine the value for the current realisation. We compute the best Schechter fit to the bins of the current realisation and iterate the process 1000 times. We obtain the 16th and 84th percentile of all fits, which we plot as contours in all figures. For each iteration, we also perturb the fixed Schechter parameters (α for all redshifts and $M_\star^*$ or M_{UV}^* for the filters with only two bins) by picking a random value in a ±0.2 dex range centred in the fixed values (same method used in Chapter 3).

7.3.6 Obtaining UV and stellar mass densities

We integrate UV LFs and SMFs to obtain the luminosity density (ρ_{UV}) and the stellar mass density (ρ_M), respectively. In order to fully take into account the uncertainties in our luminosities/stellar masses, we perturb our measurements within their errors and fit and integrate each of the 1000 realisations (see Section 7.3.5). The computed ρ_{UV} and ρ_M are the median of all integrals, with the errors being the 16th and 84th percentile of the distribution of all realisations. To obtain ρ_{UV}, we compute the integral of the UV LFs in the range $-23 < M_{UV} < -17$ (similar to e.g. Bouwens et al., 2015; Finkelstein et al., 2015). To obtain ρ_M, we compute the integral of the SMFs in the range $10^{8-13}\,M_\odot$ (similar to e.g. Davidzon et al., 2017). All ρ_M measurements in this study assume a Chabrier IMF, and values from the literature are converted to a Chabrier IMF if another IMF was used.

7.4 Results and Discussion

7.4.1 Interpreting the observed UV LF and SMF

As detailed in the previous sections, the observed distribution of both the UV LF and SMF of LAEs has a turn-over at the faintest UV luminosities and smallest stellar masses, respectively (see Figure 7.3 and 7.4). While such a turn-over has

not been observed in UV-selected or mass-selected samples, it is an expected distribution of a Lyα-selected sample, where Lyα correlates with both M_{UV} and $M_\star$ but with significant scatter (see Figure 7.1), as there will be incomplete sampling of M_{UV} and $M_\star$, particularly at the faint UV luminosity and low stellar mass regimes. As shown in Figure 7.5 (right panel), an increasing $L_{Ly\alpha}$ cut will preferentially decreases the number densities of the faintest UV luminosities, creating the turn-over which is a consequence of selection and not an intrinsic property of the UV LF of LAEs. A similar dependence is measured for the SMF in Figure 7.8 (right panel), albeit the dependence is not as strong. We make the assumption that the incomplete sampling will introduce only small contributions above the turn-over, which is supported by our measurements (Figure 7.5, right panel): when extending the luminosity cut from $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) = 42.5$ to 42.7 (and even further into 42.9 and 43.1), the number densities always have a very significant drop below the turn-over but remain roughly constant above it. By only fitting the regime above the turn-over and by fixing α as a steep slope, we are able to measure a distribution which is not dominated by incomplete sampling, and compute a proxy for the full UV LFs and SMFs.

We provide in Table 7.2 and 7.3 the best Schechter parameters of the distribution of 1) the full UV luminosity (or stellar mass) ranges (see the blue contours in Figure 7.3 and 7.4) and 2) the UV luminosity (or stellar mass) range above the turn-over, with a fixed steep α slope (see the purple contours in Figure 7.3 and 7.4). As we aim to understand the full LAE population, in the analysis conducted in the following sections we use the second fitting procedure, which gives a proxy of the full distribution of LAEs. We note nonetheless that the $L_{Ly\alpha}$ limits can have some influence on the number densities even above the turn-over, so when probing redshift evolution we extend the analysis to always use the same $L_{Ly\alpha}$ cut and ensure the samples are comparable (see discussion in Section 7.4.3).

7.4.2 The global UV LF of LAEs at $z \sim 2 - 6$

We start by measuring the UV LF of the full sample of SC4K LAEs, exploring a large volume of $\sim 10^8$ Mpc^3 at $z \sim 2-6$. With our large sample of ~ 4000 LAEs, we are capable of probing extremely bright UV luminosities, down to $M_{UV} =$
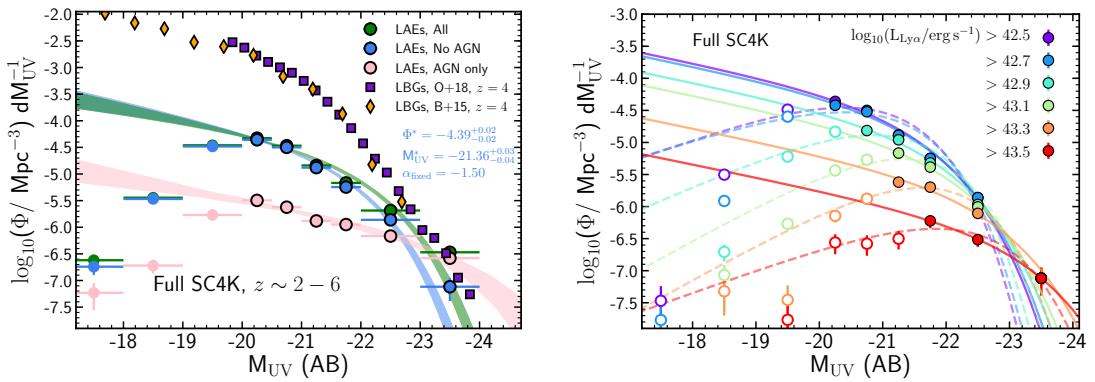
Figure 7.5: *Left:* UV LF of the full SC4K sample of LAEs: including AGN (green), no AGN (blue; what we use throughout this work) and AGN only (pink). AGN LAEs dominate the bright end ($-24 < M_{UV} < -23$) of the UV LF of LAEs. The contours are the 16th and 84th percentiles, which we obtain by perturbing the bins within their Poissonian error bars and iterating the fitting 1000 times (see Section 7.3.5). For reference, we show UV LFs of LBGs at $z \sim 4$, from Bouwens et al. (2015) (orange diamonds) and Ono et al. (2018) (purple squares). The number density of $M_{UV} = -20$ LAEs is ~ 1.5 dex lower than LBGs, but they converge to the same number densities at $M_{UV} < -23$. *Right:* UV LF of the full SC4K sample at different $L_{Ly\alpha}$ cuts. We show the best Schechter fits to the full UV luminosity range as dashed lines, and to the number densities above the turn-over as filled lines (see Section 7.3.3). The increasing $L_{Ly\alpha}$ cuts reduce the number densities, predominantly for fainter M_{UV}, which can be linked with $L_{Ly\alpha}$ and M_{UV} being typically correlated (see Figure 7.1, left panel). However, note that the UV LF of more luminous LAEs yields a declining Φ^* but a brightening in M_{UV}^*.

-24, which even in UV-continuum searches has typically only been reached in very wide area ground-based surveys (e.g. Bowler et al., 2017; Ono et al., 2018). Additionally, we have a statistically robust sample up to $M_{UV} = -20$, providing a robust probe in a range of 4 magnitudes in M_{UV}, with individual LAEs as faint as $M_{UV} = -17$.

We show in Figure 7.5 (left panel) the UV LF for three subsets of SC4K LAEs: 1) All LAEs; 2) All LAEs after removing AGN (this is the subset we use throughout this work; see Section 7.2.1.1); 3) AGN LAEs only. We show the best Schechter fits to each case as 1σ contours, which we obtain by perturbing the luminosity bins within their Poissonian errors and fitting 1000 realisations of the perturbed bins (see Section 7.3.5). We find that the UV LF of all LAEs resembles a Schechter distribution, although there is an excess at $M_{UV} < -23$, where the UV LF starts deviating from a Schechter function. A single power-law with best-fit $\log_{10}(\Phi/\mathrm{Mpc}^{-3}) = 0.71^{+0.005}_{-0.005}M_{UV} + 10.26^{+0.11}_{-0.11}$ is also a very good fit ($\chi^2_{\mathrm{reduced}} = 4.01$). When excluding AGNs, the number density significantly drops by 0.7 dex at the bright end ($-24 < M_{UV} < -23$), and the LF becomes steeper, with the single power law, with best-fit $\log_{10}(\Phi/\mathrm{Mpc}^{-3}) = 0.91^{+0.005}_{-0.005}M_{UV} + 14.49^{+0.12}_{-0.11}$, becoming less preferable ($\chi^2_{\mathrm{reduced}} = 60.63$). We observe that AGN LAEs clearly dominate the bright end ($-24 < M_{UV} < -23$) of the UV LF, with only minor contributions to the faint end ($-22 < M_{UV} < -20$). This trend is qualitatively similar to the one found in Chapter 3 for the Lyα LF of LAEs. Such a similar behaviour between the UV LF and Lyα LF is a consequence of $L_{Ly\alpha}$ and M_{UV} being typically correlated (see Figure 7.1, left panel), although the complicated radiative transfer physics behind Lyα emission should be noted (see Section 1.3.3).

7.4.3 UV LF with varying $L_{Ly\alpha}$ cuts

Due to an increasing luminosity distance with redshift, we are only capable of reaching the faintest Lyα luminosities (down to $10^{42.5}\,\mathrm{erg\,s^{-1}}$) at $z \sim 2.5$, or at higher redshifts with NBs. We aim to ensure that when comparing UV LFs at different redshifts, results are not driven by differences in depth. As such, we need to estimate how different Lyα luminosity limits affect the UV LF of LAEs. We show in Figure 7.5 (right panel) the UV LF of the full SC4K sample with

varying $L_{Ly\alpha}$ cuts, from $10^{42.5}$ to $10^{43.5}\,\mathrm{erg\,s^{-1}}$. As expected from the dependence of M_{UV} and $L_{Ly\alpha}$, an increasing $L_{Ly\alpha}$ cut predominantly decreases the number densities of fainter M_{UV} LAEs. For the full SC4K sample, between $10^{42.5}$ and $10^{43.3}\,\mathrm{erg\,s^{-1}}$, $\log_{10}\Phi$ decreases by 2.0 dex at $M_{UV} = -20.25$ but only decreases by 0.3 dex at $M_{UV} = -22.5$. This trend is qualitatively the same at all redshifts.

It is thus clear that a varying Lyα flux limit will significantly affect the UV LF as a whole, both in shape and characteristic parameters, with number densities being significantly more affected for fainter M_{UV}. To compare UV LFs at different redshifts and interpret any evolution, it is therefore necessary to ensure we use the same luminosity ranges, otherwise a potential evolution in the UV LF of LAEs may not be intrinsic but instead could be a consequence of the different Lyα luminosity limits. As such, when comparing LFs, we not only compare the full samples, but also compare a homogeneous subset, defined by a single Lyα luminosity cut of $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$, which we will apply to all redshifts. We choose this value as it excludes the lower $L_{Ly\alpha}$ regime which can only be reached at lower redshift or by the deep NBs, and covers a luminosity regime which is probed at all redshifts, ensuring we are comparing similar samples of LAEs. While this cut will only remove a small fraction of LAEs from MBs at $z > 3.5$, it will significantly reduce the number of sources at the lower redshifts, with only 10% of non-AGN LAEs at $z = 2.5$ being above this Lyα cut.

7.4.3.1 The $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ population of LAEs

In order to probe evolution in the same luminosity ranges, we have defined a sub-sample of the SC4K sample of LAEs, with $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ at all red-shifts. In comparison, the characteristic $L_{Ly\alpha}$ is measured to be $\log_{10}(L_{Ly\alpha}^{*}/\mathrm{erg\,s^{-1}}) = 42.93_{-0.11}^{+0.15}$ (Chapter 3), so these sources are extremely bright LAEs, rare dust-free starbursts. Amorín et al. (2017) has shown that such sources (galaxies in that study are also selected as LAEs in the SC4K sample) are analogues of high-z primeval galaxies.

7.4.4 Redshift evolution of the UV LF from $z \sim 2$ to $z \sim 6$

We will now use our sample of LAEs, selected with 16 unique NBs and MBs in 16 well defined redshift slices, to probe the evolution of the UV LF of LAEs from $z \sim 2$ to $z \sim 6$. We have shown in Figure 7.3 the UV LF for LAEs selected from each of the 16 individual NB and MB filters, together with best-fit Schechter and 1σ countours. We provide all the Schechter parameter estimates in Table 7.2. All samples are well represented by Schechter distributions. Our measurements agree well with Ouchi et al. (2008) at $z \sim 3$, $z \sim 4$ and $z \sim 5.7$, but we report lower number densities at $z = 5.8$, particularly for fainter M_{UV}. This discrepancy can be explained by differences in Lyα flux limits, as the MB that we use is only sensitive to $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$. We also note that our M_{UV} measurements are estimated from SED fitting with 30+ bands, including the recent ultra-deep NIR data from UltraVISTA DR4, instead of directly from adjacent photometric bands.

For a statistically robust study of the evolution of UV LFs of LAEs with redshift, we group LAEs from multiple filters that probe similar redshifts to explore five different bins of redshift ($z = 2.5$, $z = 3.1$, $z = 3.9$, $z = 4.7$ and $z = 5.4$; see Section 7.2.1.2), as well as the full SC4K sample. The completeness corrections are applied to LAEs individually, based on their Lyα luminosity (see 7.3.2) and the volume per redshift bin is the sum of the volume of individual redshift slices included in the redshift bin (see Table 7.1).

We show in Figure 7.6 (left panel) the UV LF at different redshifts ($z = 2.5$, $z = 3.1$, $z = 3.9$, $z = 4.7$ and $z = 5.4$), without any $L_{Ly\alpha}$ cut. We also show in Figure 7.6 (right panel) the 1σ, 2σ and 3σ contours of $\Phi^* - M_{UV}^*$. We observe a brightening (M_{UV}^* becomes more negative) of the UV LF with increasing redshift, from ~ -20.5 at $z = 2.5$ to ~ -22 at $z = 5.4$, and a $\log_{10}(\Phi^*/\mathrm{Mpc^{-3}})$ decrease from ~ -3.5 to ~ -4.5 for the same redshifts. While in UV-continuum studies (e.g. Bouwens et al., 2015; Finkelstein et al., 2015) Φ^* of the UV LF is also measured to decrease with increasing redshift, M_{UV}^* is found to become fainter (increase), which is the opposite of what we measure in our sample of LAEs (before applying any luminosity cut).
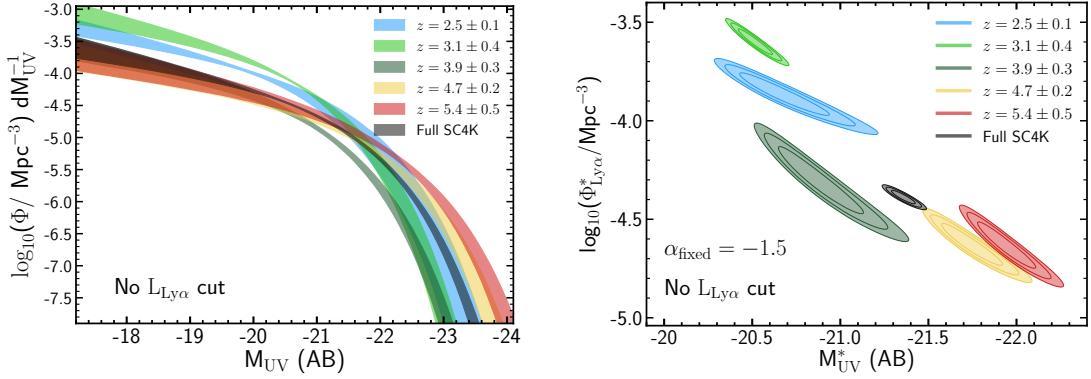
Figure 7.6: *Left:* Evolution of the UV LF with redshift, with no $L_{Ly\alpha}$ cut. The shaded contours are the 16th and 84th percentiles of all iterations obtained by perturbing the luminosity bins (see Section 7.3.5) *Right:* $\Phi^* - M^*_{UV}$ 1σ, 2σ and 3σ contours. We observe an M^*_{UV} increase from ~ -20.5 at $z \sim 2.5$ to ~ -22 at $z \sim 5 - 6$, and a $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3})$ decrease from ~ -3.5 to ~ -4.5 for the same redshifts.
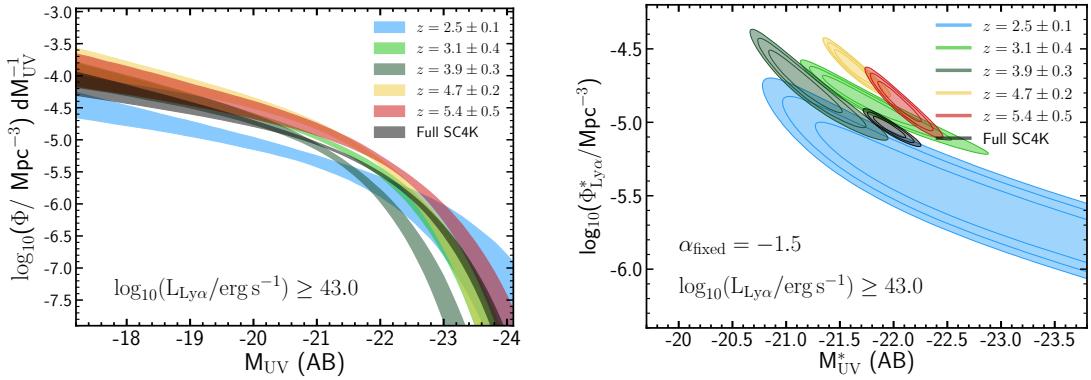
Figure 7.7: *Left:* Evolution of the UV LF with redshift, with a luminosity cut of $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s}^{-1}) \geq 43.0$. *Right:* $\Phi^* - M^*_{UV}$ 1σ, 2σ and 3σ contours. With a uniform cut for the entire sample, we note no clear evolutionary trend in M^*_{UV}, while $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3})$ remains roughly constant at -4.7 at $z \sim 3-6$. The constraints at $z = 2.5$ are worse likely due to only $\sim 10\%$ of the original sample remaining above the luminosity cut.

However, as previously discussed (Section 7.4.3), different Lyα luminosity limits play a very significant role on the shape and characteristic parameters of the UV LF. We thus conduct the same analysis for a subset of our sample of LAEs, obtained by applying the luminosity cut of $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$. By using a uniform cut at all redshifts (see Figure 7.7), we are able to probe evolution in comparable Lyα luminosity regimes, and reduce the effects of the Lyα flux limit bias. We now observe an increase of Φ^* with increasing redshift, from $\mathrm{pmblog}_{10}(\Phi^*/\mathrm{Mpc}^{-3}) \sim -5.5$ at $z = 2.5$ to -4.5 at $z \sim 3 - 6$,, which contrasts with the decrease (becomes fainter) observed in UV-continuum selected samples. We do not observe a clear evolution in M_{UV}^*, which also contrasts the increase in M_{UV}^* observed in UV-continuum selected samples.

7.4.5 The global SMF of LAEs at $z \sim 2 - 6$

Following the same methodology that we use for the UV LF, we now analyse the global SMF of ~ 4000 LAEs at $z \sim 2 - 6$. The study of the SMF of such a large sample of LAEs over such a wide volume is unprecedented at these redshift ranges. We have a robust sample of LAEs at $10^{9.0} - 10^{12.5}\,M_\odot$, with individual measurements down to $\sim 10^{7.5}\,M_\odot$. Studies that have estimated stellar masses of $z > 2$ galaxies, typically only probe $> 10^{10}\,M_\odot$ galaxies (e.g. Schreiber et al., 2015) but with our population of LAEs, we are capable of reaching galaxies with very low stellar masses, while still having detections of very massive systems ($> 10^{11}\,M_\odot$).

We show in Figure 7.8 (left panel) the SMF of the full SC4K sample of $z \sim 2-6$ LAEs after removing AGN (which is what we use throughout this work, see Section 7.2.1.1). Unlike the UV LF, we do not explore how AGNs influence the SMF since we are not able to accurately estimate the stellar mass of AGNs with our stellar+dust SED-fitting code which does not use AGN models. We show the Schechter fit to the SMF and the 1σ contour which we estimate by perturbing the stellar mass bins within their Poissonian errors and fitting 1000 realisations of the perturbed bins (see Section 7.3.5). The SMF resembles a Schechter distribution, but with an excess in number densities at $10^{12}\,M_\odot$.
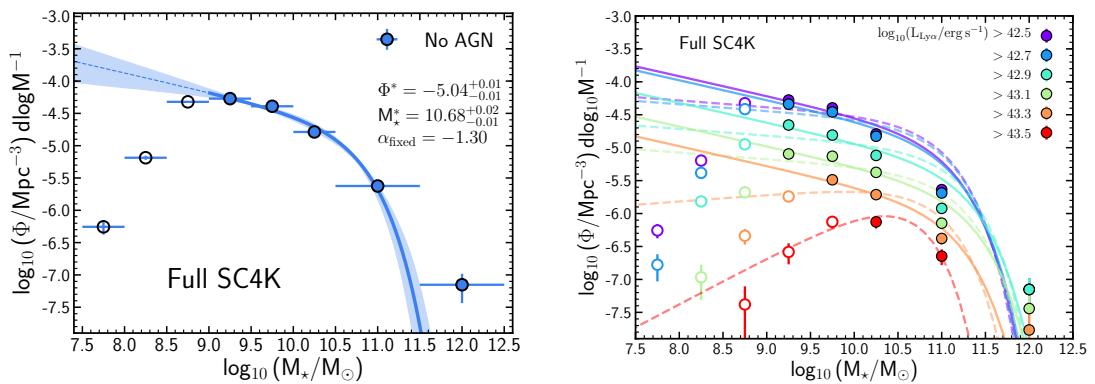
Figure 7.8: *Left:* Stellar mass function of the full SC4K sample after removing AGN (blue points, what we use throughout this study). The AGN sample is not shown here as we cannot obtain accurate mass estimations for AGNs using the stellar+dust SED-fitting we use in this study. The contours are the 16th and 84th percentiles, which we obtain by perturbing the bins within their Poissonian error bars and iterating the fitting 1000 times (see Section 7.3.5). *Right:* SMF of the full SC4K sample at different $L_{Ly\alpha}$ cuts. We show the best Schechter fits to the full stellar mass range as dashed lines, and to the number densities above the turn-over as filled lines (see §7.3.4). The increasing $L_{Ly\alpha}$ cuts reduce the number densities at all mass ranges. The decay of the number density is much more uniform across the entire mass range compared to the UV LF (Figure 7.5, right panel), which can be explained by $L_{Ly\alpha}$ and $M_\star$ having a shallower correlation with significant scatter (see Figure 7.1, right panel).

7.4.6 SMF with varying $L_{Ly\alpha}$ cuts

Here, we explore how different Lyα luminosity limits affect the SMF. For the UV LF of LAEs, we have observed that an increasing $L_{Ly\alpha}$ cut significantly affects the shape and characteristic parameters of the distribution, with a more significant effect on the number density of fainter UV luminosities, which are typically linked with lower Lyα luminosities. Such a trend is not necessarily expected for the SMF, as the relation between $M_\star$ and $L_{Ly\alpha}$ is very shallow, if even present (see Figure 7.1, right panel).

We show in Figure 7.8 (right panel) the SMF of the full SC4K sample with varying $L_{Ly\alpha}$ cuts, from $10^{42.5}$ to $10^{43.5}\,\mathrm{erg\,s^{-1}}$. As the stellar mass and $L_{Ly\alpha}$ have a shallow relation, an increasing $L_{Ly\alpha}$ limit produces a much more uniform decay of the number densities over the entire stellar mass range. Between $10^{42.5}$ and $10^{43.3}\,\mathrm{erg\,s^{-1}}$, $\log_{10}(\Phi^*/\mathrm{Mpc^{-3}})$ decreases by 1.6 dex at $\log_{10}(M_\star/M_\odot) = 9.25$ and by 1.0 dex at $\log_{10}(M_\star/M_\odot) = 11.0$, which is much more modest than the large difference observed for the UV LF.

As such, when comparing SMFs at different redshifts, we will not only look at the full samples, but we will also make use of a luminosity cut $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq$ 43.0, for the same reasons that we do for the UV LF (Section 7.4.6). This produces a luminosity range which all filters can target and is consistent with our approach to compare UV LFs.

7.4.7 Redshift evolution of the SMF of LAEs from $z \sim 2$ to $z \sim 6$

We probe the evolution of the SMF with redshift, using ~ 4000 LAEs selected in 16 well defined redshift slices from $z \sim 2$ to $z \sim 6$. We showed the SMF of LAEs selected from individual filters in Figure 7.4, together with 1σ Schechter contours. All redshift slices resemble a Schechter distribution and we provide the best-fit parameters in Table 7.3.

In order to obtain statistically robust comparisons of the evolution of the SMF of LAEs with redshift, we follow the same grouping scheme that we use for the UV LFs. We define five redshift intervals ($z = 2.5$, $z = 3.1$, $z = 3.9$, $z = 4.7$ and
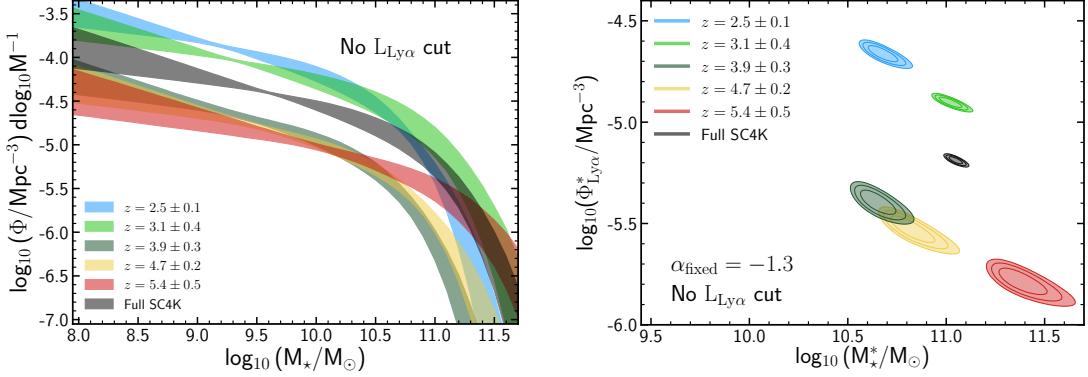
Figure 7.9: *Left:* Evolution of the SMF with redshift, with no $L_{Ly\alpha}$ cut. The shaded contours are the 16th and 84th percentiles of all iterations obtained by perturbing the luminosity bins (see Section 7.3.5) *Right:* $\Phi^*-M_\star^*$ 1σ, 2σ and 3σ contours. We observe a $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3})$ decrease from -4.5 at $z = 2.5$ to -5.5 at $z = 5 - 6$ and $\log_{10}(M_\star^*/M_\odot)$ stays constant at ~ 10.7, although we measure a small increase at $z = 5.4$.
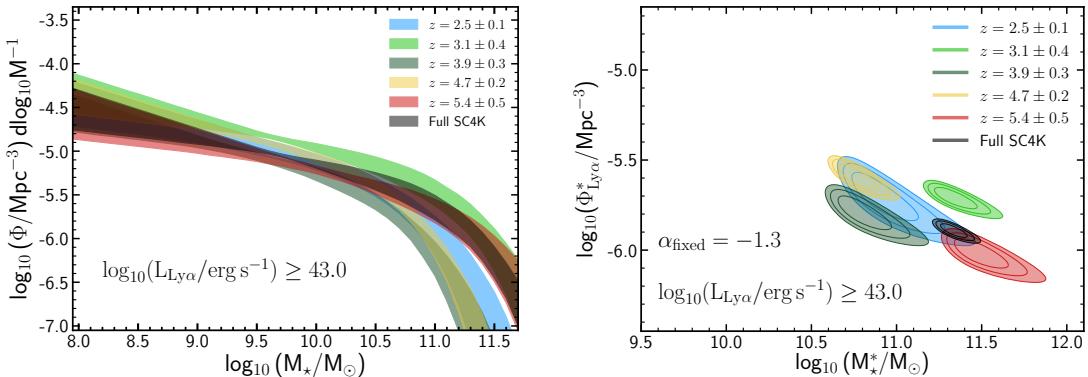
Figure 7.10: *Left:* Evolution of the SMF with redshift, with a luminosity cut of $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s}^{-1}) \geq 43.0$. *Right:* $\Phi^*-M_\star^*$ 1σ, 2σ and 3σ contours. With a uniform cut for the entire sample, we do not observe clear evidence of evolution with redshift of the SMF of LAEs. We find little $M_\star^*$ and Φ^* evolution with redshift, remaining constant at $\log_{10}(M_\star^*/M_\odot) \sim 11$ and $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3}) \sim 5.8$.

$z = 5.4$; see Section 7.2.1.2) and also use the global SMF of the full $z \sim 2-6$ sample. The completeness corrections are applied to LAEs individually, based on their Lyα luminosity (see 7.3.2) and the volume per redshift bin is the sum of the volume of individual redshift slices included in the redshift bin (see Table 7.1). We show in Figure 7.9 (left panel) the SMF at different redshifts ($z = 2.5$, $z = 3.1$, $z = 3.9$, $z = 4.7$ and $z = 5.4$), without any L$_{\mathrm{Ly}\alpha}$ cut. We also show in Figure 7.9 (right panel) the 1σ, 2σ and 3σ contours of Φ^*-M$^*_\star$. We observe a clear evolution of the SMF with redshift (before applying any Lyα luminosity restriction), with the low mass end shifting down by 1 dex from $z = 2.5$ to $z = 5.4$. This is reflected as a gradual $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3})$ decrease with redshift from -4.6 at $z = 2.5$ to -5.8 at $z = 5.4$. The shift down to lower Φ^* with increasing redshift is also observed in the SMF of more typical galaxies (e.g. Muzzin et al., 2013), which suggests the observed trends are qualitatively the same, however, an analysis using the same luminosity regime is still required.

As previously discussed in Section 7.4.6, different Lyα luminosity limits play a very significant role on the shape and characteristic parameters of the SMF. We thus conduct the same analysis for a subset of our sample of LAEs, obtained by applying the luminosity cut of $\log_{10}(\mathrm{L}_{\mathrm{Ly}\alpha}/\mathrm{erg\,s}^{-1}) \geq 43.0$. By using a uniform cut at all redshifts (see Figure 7.10), we are able to probe evolution in comparable Lyα luminosity regimes, and reduce the effects of the Lyα flux limit bias. While there is a clear evolution in the observed Schechter fits of the full samples, we find no evidence of such evolution when comparing samples of LAEs within the same Lyα regime. We find little M$^*_\star$ and Φ^* evolution with redshift, remaining constant at $\log_{10}(\mathrm{M}^*_\star/\mathrm{M}_\odot) \sim 11$ and $\log_{10}(\Phi^*/\mathrm{Mpc}^{-3}) \sim -5.8$. The evolution that we find when looking at the same luminosity regimes is thus not qualitatively the same that is observed in more typical galaxies. Analysis of the evolution of the stellar mass density, will provide more insight into this.

7.4.8 Evolution of the Lyα fraction

We attempt to infer the Lyα fraction ($\chi_{\mathrm{Ly}\alpha}$) dependence on redshift and M$_{\mathrm{UV}}$. We compute the ratio between the observed UV number densities in our sample of LAEs and the UV number densities of LBGs from the literature: $\Phi_{\mathrm{LAE}}/\Phi_{\mathrm{LBG}}$,
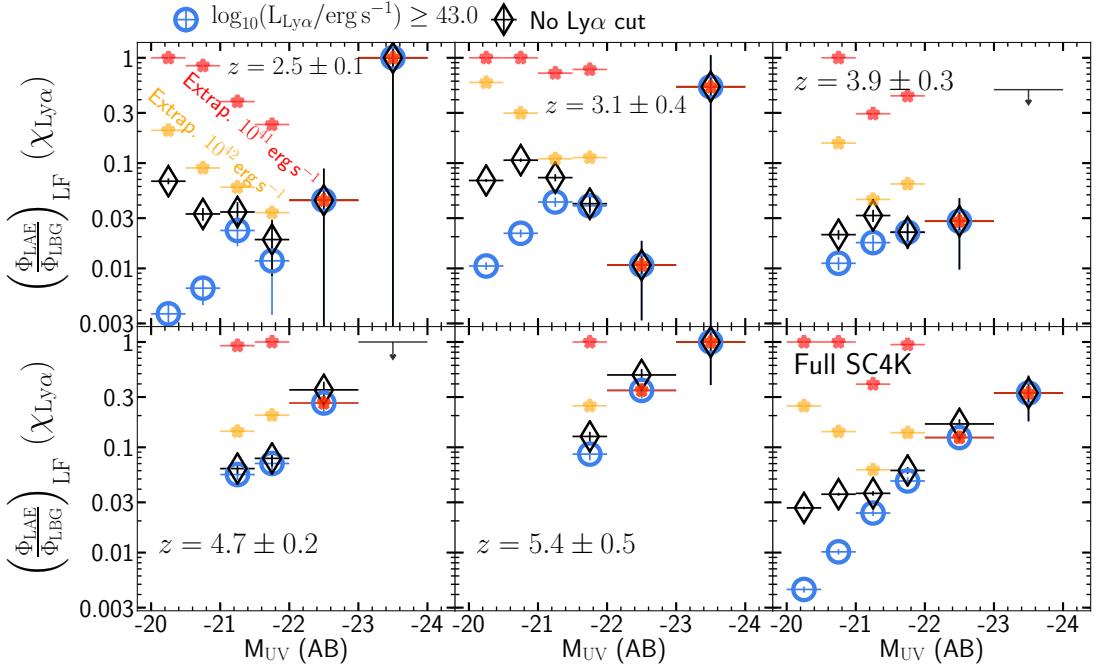
Figure 7.11: $\Phi_{\text{LAE}}/\Phi_{\text{LBG}}$ ratio (interpreted as $\chi_{\text{Ly}\alpha}$) dependence on M_{UV} for different redshifts. $\Phi_{\text{LAE}}/\Phi_{\text{LBG}}$ measurements are shown when applying a uniform $\log_{10}(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) \geq 43.0$ cut (blue circles), and when applying no cut (unfilled diamonds). The ratio is computed from a compilation of UV LFs from UV-selected galaxies at $z = 2.3$, $z = 3.05$ (Reddy & Steidel, 2009), $z = 4$, $z = 5$ and $z = 6$ (Ono et al., 2018). We show simple extrapolations to $\log_{10}(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) \geq 42.0$ (orange stars) and $\log_{10}(L_{\text{Ly}\alpha}/\text{erg s}^{-1}) \geq 41.0$ (red stars), computed from $z = 2.5$ and applied to all redshift intervals. For better visualisation, the ratio is collapsed to $\Phi_{\text{LAE}}/\Phi_{\text{LBG}} = 1$ when it surpasses that value.

which can be interpreted as the fraction of LBGs that are LAEs (above some Lyα detection limit), or $\chi_{\text{Ly}\alpha}$. To compute this fraction, we use a UV LF compilation consisting of: $z = 2.3$, $z = 3.05$ (Reddy & Steidel, 2009), $z = 4$, $z = 5$ and $z = 6$ (Ono et al., 2018) (which we use for the redshifts $z = 2.5$, $z = 3.1$, $z = 3.9$, $z = 4.7$ and $z = 5.4$, respectively). For the full SC4K sample (median $z = 4.1$) we use the $z = 4$ literature measurements from Ono et al. (2018), which being a very wide area LBG survey, provides a fair comparison with our wide area LAE survey. To prevent any biases from fitting, the ratio is computed directly from the luminosity bins in this study and the literature, with the latter being interpolated to the M_{UV} values used in this study.

As clearly seen for the full sample in Figure 7.5 (left panel), the number density of faint M_{UV} LBGs is multiple times higher than the number density of faint M_{UV} LAEs. The number densities of $M_{UV} = -20$ LAEs is ~ 1.5 dex lower than LBGs, but they converge to the same number densities for M_{UV} brighter than -23. We note, however, that this should not be interpreted as all M_{UV} bright LBGs being strong LAEs, which is evidently not the case (e.g. Shapley et al., 2003). We show the ratio of the two number densities in Figure 7.11 for five redshift intervals (and the full SC4K sample), before and after applying a $L_{Ly\alpha}$ cut. The $z = 2.5$ panel shows that as we probe fainter $Ly\alpha$ luminosities, we get closer to unity in the Φ_{LAE}/Φ_{LBG} fraction, and that the effect of the $Ly\alpha$ cut depends on M_{UV}, as shown in Section 7.4.3. For very bright M_{UV} (< -23), we are always able to retrieve most galaxies, as M_{UV} bright are typically also $Ly\alpha$ bright (see Figure 7.1, left panel) and the Φ_{LAE}/Φ_{LBG} will always be close to unity. This holds true for all redshifts, with the ratio always tending to unity at the brightest UV luminosities. When comparing Φ_{LAE}/Φ_{LBG} at different redshifts, for the comparable $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ subsample, we observe that Φ_{LAE}/Φ_{LBG} is typically higher at $z > 4$ than for the lower redshift samples. This may imply that LAEs become a bigger subset of LBGs with increasing redshift (same trend found in e.g. Arrabal Haro et al., 2020), but we explore this further by measuring the UV luminosity density., but we explore this further by measuring the UV luminosity density.

We make a direct extrapolation of the measurements of $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ and ≥ 42.5 $z = 2.5$ LAEs to lower $L_{Ly\alpha}$ cuts by scaling the increment in Φ_{LAE}. The extrapolated values for $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 42.0$ and ≥ 41.0 are shown in Figure 7.11. We find that for $M_{UV} = -20$ at $z = 2.5$, we would approach unity if we could reach $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) = 41.0$. We make the simple assumption that the extrapolation we predict for $z = 2.5$ is valid for all redshifts, as the higher flux limits of the other redshifts are not capable of reaching $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 42.5$ and thus do not allow a direct extrapolation. We find that for $z \gtrsim 3$ the ratio approaches unity even for $M_{UV} = -21$ to -22. We note that for $z > 4$ and for the full SC4K sample, the extrapolation at $M_{UV} = -22.5$ can be below the measurement without applying any $Ly\alpha$ cut, which is a consequence of applying

the $z = 2.5$ extrapolation estimation, which has a null increment for that M_{UV} value.

Furthermore, the Φ_{LAE}/Φ_{LBG} ratio can be interpreted as a way to constrain the duty cycle – the fractional time that a galaxy spends above some SFR threshold, making it observable (e.g. Jaacks et al., 2012). When the ratio is high, it implies that galaxies at that specific UV luminosity are typically going through a Lyα-bright phase. It should also be noted that while all galaxies going through a UV-bright phase can also potentially emit strongly in Lyα, the morphological analysis of LAEs shows that LAEs are typically compact (Paulino-Afonso et al., 2018). This suggests that more compact galaxies have better conditions for the Lyα photons to escape, and thus be observable (strong Lyα emission).

7.4.9 Redshift evolution of the UV luminosity density of $z \sim 2 - 6$ LAEs

We measure the UV luminosity density (ρ_{UV}) at the aforementioned redshift intervals in our sample of LAEs and explore its evolution. We detail how the integration is conducted in Section 7.3.6, with α being fixed to -1.5 but perturbed within ± 0.2 dex. We show our ρ_{UV} measurements in Figure 7.12 and compare them with measurements from LBG samples from the literature from $z \sim 2$ to $z \sim 8$ (Bouwens et al., 2015; Finkelstein et al., 2015; Reddy & Steidel, 2009).

When applying no luminosity restriction, we measure that $\log_{10}(\rho_{UV}/\text{erg s}^{-1} \text{Hz}^{-1} \text{Mpc}^{-3})$ is anti-correlated with redshift, decreasing from 25.3 at $z = 2.5$ to 25.0 at $z \sim 5-6$. When applying the luminosity cut of $\log_{10}(L_{Ly\alpha}/\text{erg s}^{-1}) \geq 43.0$, $\log_{10}(\rho_{UV}/\text{erg s}^{-1} \text{Hz}^{-1} \text{Mpc}^{-3})$ of LAEs changes from 24.3 to 25.0. In comparison, $\log_{10}(\rho_{UV}/\text{erg s}^{-1} \text{Hz}^{-1} \text{Mpc}^{-3})$ of LBGs is always higher and decreases with redshift, from 26.5 at $z = 2.5$ to 26.0 at $z = 6$. We extrapolate the ratio between the luminosity densities of $\log_{10}(L_{Ly\alpha}/\text{erg s}^{-1}) \geq 43.0$ LAEs and LBGs and determine it tends to unity at $z = 9$. Overall, our measurements of ρ_{UV} suggest that at $z \sim 2$ LAEs constitute a much smaller subset of LBGs and that with increasing redshift, both populations converge to the same values of ρ_{UV}. This is qualitatively similar to the trends found in Chapter 4 by integrating Lyα LFs.
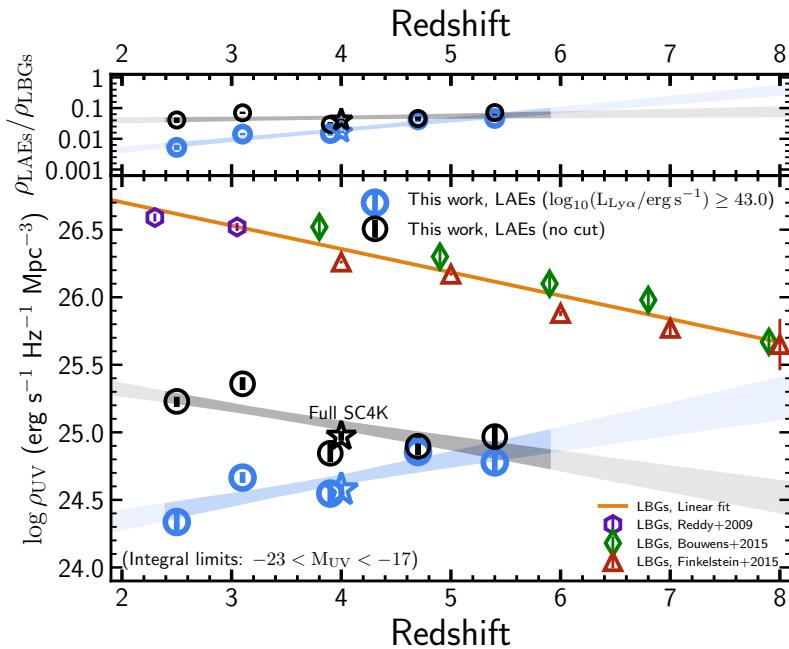
Figure 7.12: Evolution of the UV luminosity density (ρ_{UV}) with redshift. We show the ρ_{UV} measurements of our LAEs when applying a consistent $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ cut (blue circles) and with no $L_{Ly\alpha}$ cut (black circles). ρ_{UV} of the full SC4K sample is shown as stars, using the same colour scheme. The shaded contours are the 16th and 84th percentiles of the fits, obtained by perturbing the M_{UV} bins at each redshift (see Section 7.3.5). We find no evidence for ρ_M evolution with redshift when applying a consistent $L_{Ly\alpha}$ cut. We compare our results with measurements from the literature, from continuum-selected LBG populations: $z = 2.3$, $z = 3.05$ (Reddy & Steidel, 2009), $z = 3.8$, $z = 4.9$, $z = 5.9$, $z = 6.8$, $z = 7.9$ (Bouwens et al., 2015), $z = 4$, $z = 5$, $z = 6$, $z = 7$, $z = 8$ (Finkelstein et al., 2015).

7.4.10 Redshift evolution of the stellar mass density of $z \sim 2 - 6$ LAEs

Using our best-derived fits (Table 7.3), we estimate the stellar mass density (ρ_{M}) of our LAEs at different redshifts, by integrating the SMFs in the range. We obtain ρ_{M} using the procedure described in Section 7.3.6. In Figure 7.13, we show our ρ_{M} measurements and compare them with measurements from the literature. The observed ρ_{M} (without applying any luminosity cuts) changes from $\log_{10}(\rho_{\mathrm{M}}/\mathrm{M}_{\odot}\,\mathrm{Mpc}^{-3}) \sim 6.3$ at $z \sim 2.5$ to ~ 5.5 at $z \sim 5 - 6$. By applying the consistent $\log_{10}(\mathrm{L}_{\mathrm{Ly}\alpha}/\mathrm{erg\,s}^{-1}) \geq 43.0$ cut, the estimated ρ_{M} of our LAE sample remains roughly constant with redshift at $\log_{10}(\rho_{\mathrm{M}}/\mathrm{M}_{\odot}\,\mathrm{Mpc}^{-3}) \sim 5.5$.

We compare our results with measurements from the literature, from continuum-selected populations: Davidzon et al. (2017), Caputi et al. (2011), Caputi et al. (2015), Duncan et al. 2014, González et al. (2011), Grazian et al. (2015), Ilbert et al. (2013), Mortlock et al. (2011), Mortlock et al. (2015), Muzzin et al. (2013), Reddy et al. (2012), Santini et al. (2012), Song et al. (2016), and Tomczak et al. (2014). The ρ_{M} measurements of typical populations of galaxies from the literature indicate a decrease from $\log_{10}(\rho_{\mathrm{M}}/\mathrm{M}_{\odot}\,\mathrm{Mpc}^{-3}) \sim 7.5$ at $z \sim 2.5$ to ~ 6.5 at $z \sim 5 - 6$. This implies that galaxies selected as LAEs always have low stellar mass densities, and as we move to higher redshifts, their properties become similar to the ones derived from more typical populations of galaxies, suggesting that with an increasing redshift more galaxies become LAE-like. The ratio between the stellar mass densities for the $\log_{10}(\mathrm{L}_{\mathrm{Ly}\alpha}/\mathrm{erg\,s}^{-1}) \geq 43.0$ population and the values from the literature decreases from ~ 0.005 at $z \sim 2.5$ to ~ 0.05 at $z \sim 5 - 6$. We extrapolate the ratio between the stellar mass densities of $\log_{10}(\mathrm{L}_{\mathrm{Ly}\alpha}/\mathrm{erg\,s}^{-1}) \geq 43.0$ LAEs and LBGs and determine it tends to unity at $z = 10$. This implies that these bright LAEs, contribute very significantly to the total stellar mass density during the epoch of reionisation, highlighting the importance of LAEs to the evolution of primeval galaxies in the early Universe.
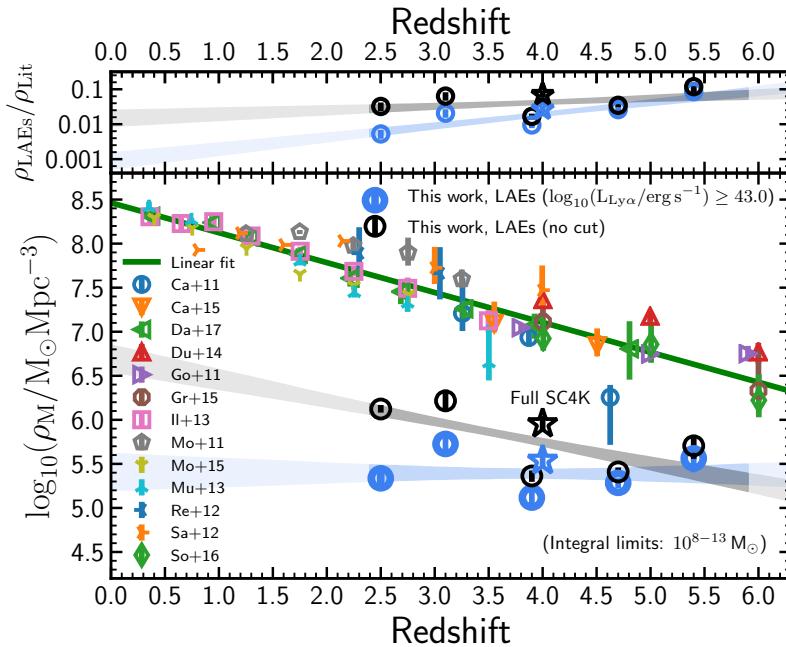
Figure 7.13: Evolution of the stellar mass density (ρ_M) with redshift. We show the ρ_M measurements of our LAEs when applying a consistent $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ cut (blue unfilled circles) and with no $L_{Ly\alpha}$ cut (black unfilled circles). ρ_M of the full SC4K sample is shown as stars, using the same colour scheme. The shaded contours are the 16th and 84th percentiles of the fits, obtained by perturbing the stellar mass bins at each redshift (see Section 7.3.5). We find no evidence for ρ_M evolution with redshift when applying a consistent $L_{Ly\alpha}$ cut. We compare our results with measurements from the literature, from continuum-selected populations: Davidzon et al. 2017 (Da+17), Caputi et al. 2011 (Ca+11), Caputi et al. 2015 (Ca+15), Duncan et al. 2014 (Du+14), González et al. 2011 (Go+11), Grazian et al. 2015 (Gr+15), Ilbert et al. 2013 (Il+13), Mortlock et al. 2011 (Mo+11), Mortlock et al. 2015 (Mo+15), Muzzin et al. 2013 (Mu+13), Reddy et al. 2012 (Re+12), Santini et al. 2012 (Sa+12), Song et al. 2016 (So+16), and Tomczak et al. 2014 (To+14). All ρ_M were converted to Chabrier, when another IMF was used. We show the best fit to this compilation as a green line. The ratio between the ρ_M from the literature and ρ_M from $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}}) \geq 43.0$ LAEs (top panel) decreases from ~ 300 at $z \sim 2.5$ to ~ 30 at $z \sim 5-6$, suggesting an increasing overlap between populations with increasing redshift.

7.5 Conclusions

In this Chapter, we determine the UV luminosity functions (LFs) and stellar mass functions (SMFs) of ~ 4000 LAEs from the SC4K sample at $z \sim 2-6$. Our main results are:

- M_{UV} and $L_{Ly\alpha}$ are typically correlated ($M_{UV} = -1.6^{+0.2}_{-0.3} \log_{10}(L_{Ly\alpha}/\text{erg s}^{-1}) + 47^{+12}_{-11}$) in our sample of LAEs. The relation between $M_\star$ and $L_{Ly\alpha}$ is shallower ($\log_{10}(M_\star/M_\odot) = 0.9^{+0.1}_{-0.1} \log_{10}(L_{Ly\alpha}/\text{erg s}^{-1}) - 28^{+4.0}_{-3.8}$).

- Different $L_{Ly\alpha}$ limits significantly affect the shape and normalisation of the UV LF and SMF of LAEs. An increasing $L_{Ly\alpha}$ cut predominantly reduces the number density of lower stellar masses and faint UV luminosities, more significantly for the UV LF. We estimate a proxy for the full UV LF and SMF of LAEs, making simple assumptions of fitting range and faint end slope. We further address the issues of incomplete sampling by applying a uniform luminosity cut of $\log_{10}(L_{Ly\alpha}/\text{erg s}^{-1}) \geq 43.0$ to our entire sample, producing a subsample of rare bright primeval galaxies.

- For the UV LF of LAEs, we find a characteristic number density (Φ^*) increase from $\log_{10}(\Phi^*/\text{Mpc}^{-3}) \sim -5.6$ at $z = 2.5$ to ~ -4.8 at $z \sim 3$, with no evolution up to $z \sim 5-6$, and no clear evolution of the characteristic UV luminosity (M^*_{UV}).

- For the SMF of LAEs, we find no significant evolution with redshift, with $\log_{10}(\Phi^*/\text{Mpc}^{-3})$ staying constant at ~ -5.5 from $z \sim 2.5$ to $z \sim 6$ and the characteristic stellar mass staying constant at $\log_{10}(M^*_\star/M_\odot) \sim 11$ for the same redshift range.

- We compute Φ_{LAE}/Φ_{LBG} (proxy of $\chi_{Ly\alpha}$) which tends to unity with increasing M_{UV} at all redshifts, as bright LAEs are typically also bright in M_{UV}. For fainter LAEs, the ratio tends to one as we reach fainter Lyα fluxes, with a simple extrapolation implying that by reaching $\log_{10}(L_{Ly\alpha}/\text{erg s}^{-1}) = 41.0$ we would approach unit for $M_{UV} = -20$ galaxies at $z = 2.5$.

- From $z = 2.5$ to $z \sim 5-6$, the luminosity density (ρ_{UV}) increases from $10^{24.3}$ to $10^{25.0}$ erg s^{-1} Hz^{-1} Mpc^{-3}. For the same redshift range, the stellar mass density (ρ_{M}) shows little evolution, remaining constant at $\sim 10^{5.5}$ M$_\odot$ Mpc^3, and being always lower than the total luminosity and stellar densities of continuum-selected galaxies but approaching it with increasing redshift. Overall, we find that the stellar and luminosity density of bright LAEs are extrapolated to converge to the measurements of continuum-selected galaxies at $z \sim 9-10$, pointing to the very significant role of LAEs in the epoch of reionisation.

Chapter 8

Conclusions and Future Work

In this thesis, we contribute to the understanding of the nature, typical properties and evolution of galaxies with Lyα emission, between the end of the epoch of reionisation at $z \sim 6$ and the peak of cosmic SFH at $z \sim 2$. Here, we summarise the main results of the thesis, give some final remarks and discuss future lines of research.

8.1   Conclusions

We present a new large sample of Lyα emitters (LAEs) - SC4K (Slicing COSMOS with 4k LAEs) - and study the evolution with redshift of physical properties of galaxies in this sample. The catalogues with the SC4K sample and derived properties are made fully public (see Appendices A.1 and A.4) so the community can fully benefit from the work presented in this thesis. Multiple studies have already been conducted using such catalogues (e.g. Calhau et al., 2020; Khostovan et al., 2019; Marques-Chaves et al., 2020; Paulino-Afonso et al., 2018; Shibuya et al., 2019).

SC4K: sample and luminosity functions
(Chapters 2, 3 and 4)

Most studies of Lyα emitting galaxies focus on a single redshift slice and/or are limited by very small areas/volumes. Inferring potential evolution of galaxy

properties by comparing multiple different surveys is challenging with hard to quantify biases. With this study, we aimed to develop a single self-consistent survey to construct a large sample of LAEs, obtained using the same data reduction methods, same (wide) areas, similar filters and similar selection criteria. This is done in an effort to build a sample where comparison of redshifts can be done with as little bias as possible. The SC4K survey fully shows the capabilities of narrow/medium band surveys in efficiently probing a large area of the sky to efficiently build large samples of young star-forming galaxies (and AGNs). Our sample consists of 3908 LAEs, 318 of which have AGN signatures (Calhau et al., 2020). The Lyα LFs reveal an AGN component which dominates the brightest luminosities at $z < 4$, but disappears at higher redshifts.

Furthermore, by combining our measurements with results from the literature, including IFU and blind spectroscopy surveys, we are able to obtain a much more complete view of the Lyα emitting population, which would not be possible from any single selection method. Our results indicate a brightening with redshift of the characteristic Lyα luminosity. We find an increase by a factor of ≈ 2 of the Lyα luminosity density from $z = 2$ to $z = 3$ but no evolution at $z \sim 3 - 6$, which results in an increase of $\mathrm{SFRD_{UV}/SFRD_{Ly\alpha}}$ from 4% at $z = 2.2$ to 30% at $z = 6$. This can be translated into a factor of ≈ 2 increase in the ionising efficiency and a ≈ 4 increase in the Lyα escape fraction, from $z \sim 2$ to $z \sim 6$.

SC4K: properties and evolution
(Chapters 5 and 6)

In these chapters, we conduct PSF photometry for each individual source in the SC4K sample of LAEs, using 34 bands from rest-frame UV to FIR, which we use to measure the SEDs with the stellar+dust SED fitting code MAGPHYS. The large number and wide wavelength coverage of filters used allow a very accurate constraints of SEDs and parameters. This is further enhanced by very recent programs such as UltraVISTA DR4 which improves the accuracy of the SEDs of our high redshift galaxies with ultra-deep NIR imaging. Overall, we find that while (as expected) LAEs are typically very young, very blue and low mass, there is a significant diversity in the properties of LAEs. We also find that the typical properties of LAEs show little evolution with redshift, implying that galaxies

191

that are selected as LAEs always have the same properties. When compared with continuum-selected galaxies, we find that they become LAE-like at $z > 5$.

SC4K: UV luminosity and stellar mass functions
(Chapter 7)

We estimate the UV LF and the SMF of $z \sim 2 - 6$ LAEs, both still widely unconstrained up to this point. For the UV LF, we find an increase in the characteristic number density, with increasing redshift. Such an evolutionary trend is qualitatively very different from the UV LF of LBGs (e.g. Bouwens et al., 2015; Finkelstein et al., 2015), which has a decreasing characteristic number density with increasing redshift. Additionally, our measurements of the SMF are unprecedented, as no other survey has been able to derive stellar masses for such a large sample of LAEs over such a wide redshift range. Our measurements of the UV luminosity and stellar mass densities of LAEs are always below the measurements of continuum-selected galaxies but are extrapolated to converge to the same values at $z \sim 8 - 10$.

Overall the analysis of multiple properties of SC4K: the evolution of UV sizes, Lyα escape fraction, SFR-M$_\star$ relation, $\rho_{\rm UV}$ and $\rho_{\rm M}$, all build a picture where as we move to higher redshifts, the properties of continuum-selected galaxies and LAEs become increasingly similar, implying that by $z > 5$ star-forming galaxies are LAE-like. Furthermore, this also suggests the important role that LAEs play in the epoch of reionisation and their importance in galaxy evolution.

8.2   Future work

While we now have a better understanding of the properties of LAEs and their evolution with redshift, there are still many open questions. Here we discuss possible research paths and the questions they may answer:

- **spectroscopic follow-up of SC4K LAEs:** extensive spectroscopic follow up will fully reveal the interloper fraction of our sample of LAEs (estimated 10-15%, see §2.3.5). Public COSMOS data spectroscopically confirms 119

SC4K LAEs (see §2.3.5), in addition to the 10 "primeval galaxies" also detected in the VUDS program (Amorín et al., 2017). We have already conducted observations of 514 SC4K LAEs using the multi-fibre spectrograph AF2 at the William Herschel Telescope. I have developed a PYTHON pipeline to reduce and analyse this AF2 data, with the results to be published at a later date. X-SHOOTER observations are being planned, as the deep and high-resolution spectra will show the Lyα line profiles and reveal valuable information about the escape of Lyα photons and the HI column densities.

- **multi-wavelength follow-up of SC4K LAEs:** with e.g. ALMA would reveal the dust content properties of typical LAEs (see e.g. Matthee et al., 2017d, 2019) and their evolution from $z \sim 2$ to $z \sim 6$.

- **extend the redshift range probed by SC4K:** we have used the SC4K sample to get a consistent view of the evolution of LAEs from $z \sim 2$ to $z \sim 6$. For higher redshifts, we have conducted observations with the HAWK-i instrument at the VLT using NB1061 to select LAEs at $z = 8.8$ (to be presented in Wade et al. in prep). For $z < 2$, Lyα cannot be detected from the ground, so a space telescope with an imaging instrument capable of observing with narrow bands in the UV would be required (but see Östlin et al., 2014, who constructed the Lyα reference sample at $z = 0.028 - 0.19$, LARS, using a combination of GALEX, SDSS and HST observations).

- **extend the volumes probed by SC4K:** larger volumes would allow us to find more extremely luminous $L_{Ly\alpha} > 10^{44} \, \mathrm{erg \, s^{-1}}$ sources, which are ideal candidates for spectroscopic and multi-wavelength follow-up. Such as study would improve the constrains on the number densities and properties of the brightest populations, which require us to probe extremely low number densities. This would allow us to determine if the excess in number densities of bright LAEs, driven by AGN, which we observe at $z \sim 2 - 3$ but not at $z > 4$, intrinsically disappears or if it just falls below our current detection limits.

- **conduct the same $z \sim 2-6$ search with a non-resonant line:** while we have shown how useful Lyα emission can be in probing galaxy evolution, the fact remains that Lyα's resonant nature makes it a challenging line to interpret. A $z > 2.5$ study with e.g. Hα would offer a more direct view of star-formation, although it cannot be conducted from the ground and no current space telescope can conduct such survey until the launch of the *James Webb Space Telescope* in the nearby future.

Appendix A

Appendices

A.1 Catalogue of Lyman-α emitters (SC4K)

We publicly release the full SC4K catalogue of 3,908 LAEs at $z \sim 2 - 6$ derived and used in this work, based on data obtained with 16 different medium- and narrow-band filters over the full COSMOS field. We show 5 example entries of the catalogue in Table A.1. The full electronic version of the catalogue is available with the refereed paper in a FITS table[1]. Table A.2 presents the colour terms used to correct medium-band magnitudes and to compute emission line fluxes.

A.2 [OIII]+Hβ excess in the K_s band at $z \approx 3$

A diagnostic that provides information on the validity of the sample of LAEs, and simultaneously provides insight into their nature is the evolution of the $H - K_s$ colours with redshift. The flux in these filters may be boosted by strong Hα and [OIII]+Hβ emission lines (e.g. Faisst et al., 2016), depending on the redshift, affecting the $H - K_s$ colours (see also Forrest et al. 2017). If our sample had significant number of redshift interlopers, such effects on the colours would not be seen, as interlopers will not show them.

---

[1]https://academic.oup.com/mnras/article/476/4/4725/4858393#supplementary-data
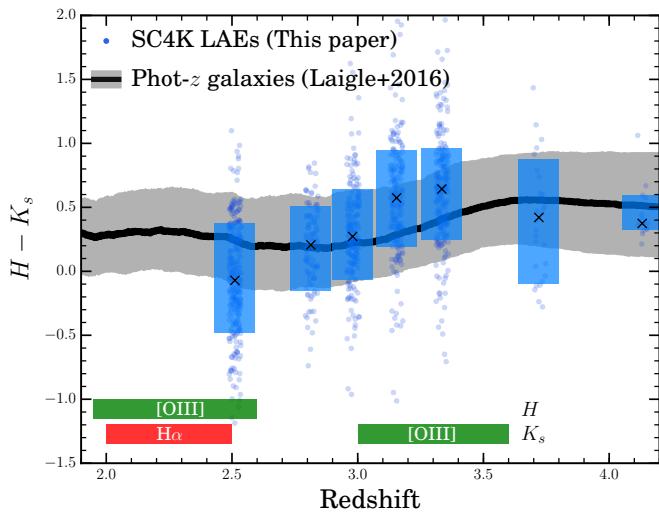
Figure A.1: $H - K_s$ colours as a function of redshift for our sample of LAEs at $z \sim 2.5 - 4$. The grey region shows the 16-84 percentile range of the colours of the general galaxy population (Laigle et al., 2016), while the blue boxes show the percentiles for the SC4K LAEs. We use photometric-redshifts for the general galaxy sample, but assign the redshift where Lyα falls in the MB for the SC4K emitters (points are randomly shifted for visualisation purposes). The green and red boxes indicate the redshifts where the strong Hα and [OIII] lines fall in the H and K_s filters and can affect the colours. The LAEs at $z \approx 3.1 - 3.4$ have systematically redder $H - K_s$ colours compared to the general galaxy population, indicative of strong [OIII] emission in the K_s filter.

Table A.1: Our full SC4K catalogue of candidate LAEs which we release with this work. The SC4K catalogue contains the samples obtained with the 12 COSMOS medium-bands, together with 4 narrow-band samples from Santos et al. (2016), Sobral et al. (2017), Matthee et al. (2017b) and Perez et al. (in prep.). We provide five example entries. The full catalogue is available in electronic format (FITS table). Errors on EW_0, Flux and $L_{Ly\alpha}$ are computed by independently perturbing the MB and BB magnitudes along their Gaussian uncertainties 10,000 times per source and computing the 16th and 84th percentiles of each computed quantity. Note that for faint sources EWs are affected by large uncertainties; see e.g. IA427-141. The AGN flag in the catalogue provides information on the matches with public X-ray (including coverage) and radio catalogues (see Section 3.1.7): 0 – no match/no coverage; 1 – X-ray detected; 2 – radio detected.

ID (SC4K-)	R.A. (J2000)	Dec. (J2000)	MB or NB (AB)	BB (AB)	EW_0 (Å)	Flux/10^{-17} (erg s^{-1} cm^{-2})	$\log_{10} L_{Ly\alpha}$ (erg s^{-1})	AGN flag (X-ray or radio)
IA427-141	10 03 20.01	+02 13 38.8	24.83 ± 0.06	26.47 ± 0.28	2000^{+2000}_{-1400}	$13.4^{+0.7}_{-0.7}$	$42.83^{+0.02}_{-0.02}$	0
IA427-446	10 02 38.96	+02 14 16.3	24.87 ± 0.07	25.97 ± 0.18	289^{+270}_{-124}	$10.6^{+0.9}_{-1.0}$	$42.73^{+0.03}_{-0.04}$	0
IA427-865	10 02 17.97	+02 15 03.2	24.84 ± 0.07	25.82 ± 0.15	205^{+136}_{-98}	$10.2^{+1.0}_{-1.8}$	$42.71^{+0.04}_{-0.08}$	0
IA427-1169	10 03 10.85	+02 15 37.6	24.28 ± 0.04	25.25 ± 0.09	193^{+126}_{-38}	$16.8^{+2.4}_{-1.0}$	$42.92^{+0.06}_{-0.03}$	0
IA427-1559	10 02 13.65	+02 16 28.9	24.90 ± 0.07	25.46 ± 0.11	65^{+7}_{-21}	$6.5^{+0.3}_{-1.6}$	$42.51^{+0.02}_{-0.13}$	0

Figure A.1 shows the median $H - K_s$ colours of the general galaxy population in the COSMOS field (Laigle et al., 2016) and of the SC4K sample of LAEs from $z \sim 2$ to $z \sim 4$. Several interesting trends can be seen. The sample of LAEs at $z \approx 2.5$ has systematically bluer $H - K_s$ colours than the general galaxy sample, which indicates that the H band is significantly boosted by strong [OIII]+Hβ emission, while the majority of the sample does not have Hα falling in the K_s filter. The LAEs at $z = 3.1 - 3.4$ have systematically redder $H - K_s$ colours than typical galaxies. This indicates LAEs have relatively strong [OIII]+Hβ emission, which is similar to the spectroscopic results from Nakajima et al. (2016). As no strong lines affect the $H - K_s$ colours at $z \approx 2.7 - 3.0$ and $z > 3.6$, the colours of LAEs at these redshifts are similar to the colours of the general population.

Table A.2: The colour coefficients for each medium-band, used to correct the observed medium-band magnitudes (MB_0) into MB, as defined in Equation 2.2: $MB = MB_0 - (m \times (BB - BB_{adjacent}) + b)$. For sources without a colour determination ($BB - BB_{adjacent}$) we add the median correction listed in the table. Note that we use the MB magnitudes (and not MB_0) for our SC4K catalogue (see Table A.1) and all derived quantities.

| MB	$BB - BB_{adjacent}$	m	b	Median correction
IA427	$B - U$	0.33	-0.11	0.01
IA464	$B - V$	0.0	0.0	0.0
IA484	$B - V$	0.0	0.0	0.0
IA505	$V - B$	0.0	0.0	0.0
IA527	$V - B$	0.0	0.0	0.0
IA574	$r^+ - V$	0.0	0.0	0.0
IA624	$r^+ - i^+$	0.0	0.0	0.0
IA679	$r^+ - i^+$	-0.30	-0.18	0.31
IA709	$r^+ - i^+$	-0.31	0.0	-0.13
IA738	$r^+ - i^+$	-0.14	0.08	-0.14
IA767	$i^+ - z$	0.0	0.25	-0.25
IA827	$i^+ - z$	-0.49	0.34	-0.20
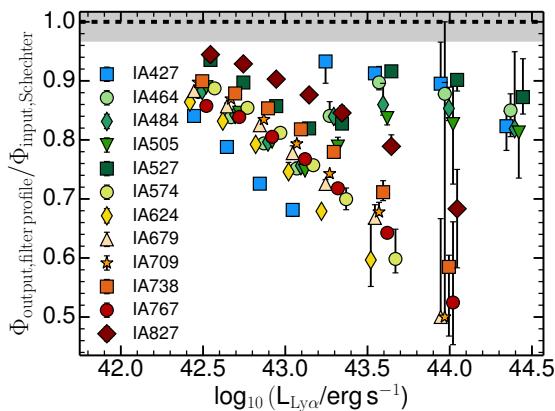
Figure A.2: The ratio between the observed Lyα luminosity function through the real filter profiles and input simulated sample of LAEs. We shift the bins by ±0.08 for visibility. For each filter, we distribute simulated sources over a large redshift range (wider than what the filter can detect), with a number density distribution given by the first pass/observed Lyα luminosity function. This includes both a Schechter component and a power-law component at the highest luminosities, for $2.2 < z < 3.5$, and a Schechter component only for $3.5 < z < 6$. Points are offset in luminosity for visibility. We find that the bright end of the Schechter component of the LF leads to a significant observed underestimation of the LF, while the power-law component is more easily recovered (see Section A.3.1 for full details).

A.3   Lyα Luminosity functions

A.3.1   Filter profile corrections

In order to evaluate the necessary potential corrections to the Lyα LF due to the use of different real filters in comparison to idealized top-hat versions of them, we follow the procedure fully described in Section 3.1.2.2 (see also Sobral et al., 2012). Here we provide the full results of our simulations, presented in Figure A.2 (see also Figure 3.1) and discuss them. We find that the number density of sources recovered by folding through a population of LAEs with a luminosity function described by a Schechter function is always underestimated, and strongly underestimated for the highest luminosities. This is a relatively easy effect to understand, and becomes particularly important for the state-of-the-art large volume surveys that can now probe significantly above L*, where the number counts may drop exponentially (contrarily to the behavior of the sub-L* component of the luminosity function, which behaves like a power-law).

Our results are a consequence of observed fluxes being the convolution between real input fluxes and the filter profile transmission. On average, this always results in a drop of flux except at the very peak of filter profile transmission. For medium-band filters there is still considerable volume under these conditions, while for narrow-bands such fraction is lower. For the evaluation of the luminosity function, this means that the observed number densities of sources at some luminosity L are always lower than reality, as most sources of that luminosity actually contribute to bins of fainter luminosities (they are observed to be fainter). The effect is not always extreme because it is partially compensated by sources at even higher luminosities that count towards a bin at luminosity L; this is why the intrinsic shape is crucial. The global result is that such corrections depend on the shape of the intrinsic luminosity function and how steep number density counts drop as a function of luminosity. This means that while for some shallow faint-end slopes the ratio of the number of sources scattering in from higher luminosities and those scattering out from a given bin is close to 1 (meaning the recovered number density is close to the input one), for steeper functions (and even more so for exponential declines), the effect starts to become very strong, as sources

move to lower luminosity bins and almost no other sources come from brighter bins to compensate (because they are simply too rare). The effect is therefore the strongest beyond $L^\star$ for an intrinsic Schechter LF distribution of sources being observed through any filter profile that is not a perfect top-hat.

Overall, our results show that for any reasonable Schechter function prior (the observed LF is a good prior which does not require any assumptions), the bright end of the Lyα LF must be corrected more than the faint end, and that such corrections are much larger with narrow-band filters than with medium band filters (see Figure 3.1). Moreover, once corrections are applied, narrow- and medium-band independent estimates agree. Our results also show that the exponential decline part of the LF becomes more and more underestimated the more Gaussian and the narrower a filter is (compared to assuming a top-hat transmission for the flux and volume). We note that another way to correct for the filter profile effects is to shift the luminosity function by some constant (to correct for the fact that a fraction of sources are not measured at full transmission). This is a relatively good way to do this in the case of a self-similar LF and/or when one is only measuring the power-law component of the Schechter function and when such a component is not tremendously steep. However, for large volumes that can trace the exponential decline, the corrections are simply not the same: for the power-law (faint) component there are sources coming in from higher luminosities and scattering to fainter luminosities, but in the exponential part there is much more migration away from the bin to lower L than there is migration into the bin from brighter sources.

We also find that if the decline of the number densities at high L is described by a relatively shallow power-law (such as the cases found at $z \sim 2 - 3$), then the corrections (Figure A.2) are close to unity (again, due to the same effect: there are brighter sources which are still numerous enough to make it into the bin and roughly compensate for sources that are observed to be fainter). Overall, our results show the importance of correcting for this effect specifically for Schechter-like functions, and less so for the case of a shallow power-law decline with increasing luminosity.

A.3.2 Luminosity functions: this study and the S-SC4K compilation/comparison

We provide the derived Lyα LFs (one example realisation; Table A.3), including the observed number of sources and number densities obtained after completeness and filter profile corrections (see Table A.3) and the full error propagation steps (see Section 3.1.2). Table A.4 presents the full S-SC4K compilation which we use to compare our results and to derive our synergy LF (S-SC4K). We also present the results from the 10,000 fits to each perturbed LF in Table A.5.

Table A.3: The global Lyα LF and for each of the medium-band filters in SC4K/this study (full LFs provided as a FITS catalogue). Here we present the first LF (global) with the first 13 entries in the table. We show the sample/filter name, followed by the Lyα luminosity bin. We also present the number of observed sources in each bin and the volume densities with the chain following the steps described in the work: observed, observed + perturbing selection, completeness corrected and filter profile corrected (final). In addition, we also show the full sequential error calculation/propagation (see full details in Section 3.1.2). We note that we set the error to 1.0 whenever it is not defined in log space (for the odd bins which are just populated by one source); for these bins the error propagation is not conducted.

Sample	$\log_{10} \mathrm{L_{Ly\alpha}}$ $(\mathrm{erg\,s^{-1}})$	Sources $(\#)$	Φ_{observed} $(\mathrm{Mpc^{-3}})$	$\Delta\Phi_{\mathrm{obs+pert}}$ $(\mathrm{Mpc^{-3}})$	$\Phi_{\mathrm{comp.corr}}$ $(\mathrm{Mpc^{-3}})$	Φ_{final} $(\mathrm{Mpc^{-3}})$	$\Delta\Phi_{\mathrm{final}}$ $(\mathrm{Mpc^{-3}})$
SC4K All MBs	42.60 ± 0.05	156 ± 12	$-3.41^{+0.03}_{-0.04}$	$^{+0.07}_{-0.07}$	$-3.02^{+0.08}_{-0.10}$	$-2.93^{+0.08}_{-0.11}$	$^{+0.09}_{-0.11}$
SC4K All MBs	42.70 ± 0.05	134 ± 11	$-3.47^{+0.04}_{-0.04}$	$^{+0.06}_{-0.06}$	$-3.23^{+0.06}_{-0.08}$	$-3.12^{+0.07}_{-0.09}$	$^{+0.07}_{-0.09}$
SC4K All MBs	42.80 ± 0.05	607 ± 24	$-3.45^{+0.02}_{-0.02}$	$^{+0.03}_{-0.03}$	$-3.12^{+0.04}_{-0.04}$	$-3.05^{+0.04}_{-0.04}$	$^{+0.05}_{-0.05}$
SC4K All MBs	42.90 ± 0.05	463 ± 21	$-3.68^{+0.02}_{-0.02}$	$^{+0.03}_{-0.03}$	$-3.47^{+0.04}_{-0.04}$	$-3.37^{+0.04}_{-0.04}$	$^{+0.05}_{-0.05}$
SC4K All MBs	43.00 ± 0.05	405 ± 20	$-3.89^{+0.02}_{-0.02}$	$^{+0.05}_{-0.05}$	$-3.70^{+0.05}_{-0.06}$	$-3.60^{+0.05}_{-0.06}$	$^{+0.06}_{-0.07}$
SC4K All MBs	43.10 ± 0.05	220 ± 14	$-4.22^{+0.03}_{-0.03}$	$^{+0.05}_{-0.05}$	$-4.12^{+0.06}_{-0.06}$	$-4.00^{+0.06}_{-0.07}$	$^{+0.07}_{-0.07}$
SC4K All MBs	43.20 ± 0.05	188 ± 13	$-4.35^{+0.03}_{-0.03}$	$^{+0.09}_{-0.09}$	$-4.22^{+0.10}_{-0.11}$	$-4.11^{+0.10}_{-0.12}$	$^{+0.11}_{-0.12}$
SC4K All MBs	43.30 ± 0.05	113 ± 10	$-4.57^{+0.04}_{-0.04}$	$^{+0.11}_{-0.11}$	$-4.48^{+0.11}_{-0.13}$	$-4.37^{+0.12}_{-0.14}$	$^{+0.12}_{-0.15}$
SC4K All MBs	43.40 ± 0.05	57 ± 7	$-4.92^{+0.05}_{-0.06}$	$^{+0.10}_{-0.10}$	$-4.80^{+0.10}_{-0.13}$	$-4.68^{+0.11}_{-0.14}$	$^{+0.11}_{-0.14}$
SC4K All MBs	43.50 ± 0.05	50 ± 7	$-5.06^{+0.06}_{-0.07}$	$^{+0.13}_{-0.13}$	$-4.93^{+0.14}_{-0.20}$	$-4.82^{+0.15}_{-0.21}$	$^{+0.15}_{-0.22}$
SC4K All MBs	43.60 ± 0.05	35 ± 5	$-5.21^{+0.07}_{-0.08}$	$^{+0.16}_{-0.16}$	$-5.14^{+0.16}_{-0.25}$	$-5.02^{+0.17}_{-0.27}$	$^{+0.18}_{-0.27}$
SC4K All MBs	43.75 ± 0.05	14 ± 3	$-5.61^{+0.10}_{-0.14}$	$^{+0.12}_{-0.12}$	$-5.58^{+0.12}_{-0.17}$	$-5.47^{+0.13}_{-0.19}$	$^{+0.13}_{-0.19}$
SC4K All MBs	44.00 ± 0.15	24 ± 4	$-5.86^{+0.08}_{-0.10}$	$^{+0.11}_{-0.11}$	$-5.85^{+0.11}_{-0.14}$	$-5.79^{+0.12}_{-0.15}$	$^{+0.12}_{-0.15}$
SC4K All MBs	44.30 ± 0.15	5 ± 2	$-6.54^{+0.16}_{-0.26}$	$^{+0.16}_{-0.16}$	$-6.54^{+0.16}_{-0.27}$	$-6.46^{+0.17}_{-0.28}$	$^{+0.17}_{-0.29}$
SC4K All MBs	44.60 ± 0.15	$1^{+1}_{-0.8}$	$-7.24^{+0.30}_{-1.00}$	$^{+0.30}_{-0.30}$	$-7.24^{+0.30}_{-1.00}$	$-7.24^{+0.30}_{-1.00}$	$^{+0.30}_{-1.00}$

Table A.4: A compilation of Lyα LFs used or compared with in this study, by alphabetical order. We provide references to the original papers and also references for LFs generated to make them more comparable with those we present in this work when appropriate (e.g. by correcting for potential contamination or by applying consistent filter profile corrections for comparison). We provide all these LFs as a FITS format catalogue. The redshifts are the average when studies have used redshift bins with the ± representing the maximum and minimum redshifts in the studies, and not the standard deviation. Note that for NB surveys this is given/rounded to 0.1, but typically the redshift range is smaller than that. The minimum and maximum luminosity bins probed by each study are given in $\log_{10}(L_{Ly\alpha}/\mathrm{erg\,s^{-1}})$. CC: correction for potential contamination by lower redshift emitters (see Sobral et al., 2017); FPC: correction for filter profile effects (see this study and Matthee et al., 2015; Santos et al., 2016).

Study # (This compilation)	Reference(s) (Original or w/ correction)	Technique/ Instrument	Redshift (z)	$L_{Ly\alpha,min}$ ($\log_{10}$)	$L_{Ly\alpha,max}$ ($\log_{10}$)
1	Bina et al. (2016)	IFU MUSE-All	$z = 4.8 \pm 1.8$	41.3	42.2
2.1	Cassata et al. (2011)	Slit VIMOS-bin	$z = 2.5 \pm 0.5$	41.3	42.9
2.2	Cassata et al. (2011)	Slit VIMOS-bin	$z = 3.8 \pm 0.8$	41.8	42.8
2.3	Cassata et al. (2011)	Slit VIMOS-bin	$z = 5.5 \pm 1.0$	42.1	43.3
3	Dawson et al. (2007)	NB Mosaic-CCD MT	$z = 4.5 \pm 0.1$	42.2	43.4
4	Drake et al. (2017a)	IFU MUSE-All	$z = 4.8 \pm 1.8$	41.9	42.9
5	Drake et al. (2017b)	IFU MUSE-All	$z = 4.7 \pm 1.9$	41.2	42.8
5.1	Drake et al. (2017b)	IFU MUSE-Bin	$z = 3.5 \pm 0.5$	41.6	42.8
5.2	Drake et al. (2017b)	IFU MUSE-Bin	$z = 4.5 \pm 0.5$	41.6	43.3
5.3	Drake et al. (2017b)	IFU MUSE-Bin	$z = 5.8 \pm 0.8$	41.6	43.2
6	Konno et al. (2016)	NB S-cam Subaru	$z = 2.2 \pm 0.1$	41.7	44.4
6.1	Konno et al. (2016); Sobral et al. (2017)	NB S-cam Subaru CC	$z = 2.2 \pm 0.1$	41.7	44.4
7	Konno et al. (2018)	NB HSC Subaru	$z = 5.7 \pm 0.1$	43.0	43.8
7.1	Konno et al. (2018); Santos et al. (2016)	NB HSC Subaru FPC	$z = 5.7 \pm 0.1$	43.0	43.8
8.1	Matthee et al. (2017b)	NB WFC INT	$z = 2.2 \pm 0.1$	42.8	43.5
8.2	Matthee et al. (2017b)	NB WFC INT	$z = 2.4 \pm 0.1$	43.4	44.7
8.3	Matthee et al. (2017b)	NB WFC INT	$z = 3.1 \pm 0.1$	43.0	43.6
9.1	Ouchi et al. (2008)	NB S-cam Subaru	$z = 3.1 \pm 0.1$	42.2	43.6
9.2	Ouchi et al. (2008)	NB S-cam Subaru	$z = 3.7 \pm 0.1$	42.7	43.5
9.3	Ouchi et al. (2008)	NB S-cam Subaru	$z = 5.7 \pm 0.1$	42.5	43.5
9.4	Ouchi et al. (2008); Santos et al. (2016)	NB S-cam Subaru FPC	$z = 5.7 \pm 0.1$	42.5	43.5
10	(Perez et al. in prep.)	NB S-cam Subaru	$z = 4.8 \pm 0.1$	43.1	43.5
11	Santos et al. (2016)	NB S-cam Subaru	$z = 5.7 \pm 0.1$	42.5	43.7
12	Sobral et al. (2017)	NB WFC INT	$z = 2.2 \pm 0.1$	42.3	43.5

A.4 The full SC4K catalogue with PSF photometry and all derived quantities

We provide the full catalogue of SC4K LAEs in electronic format (FITS format)[1] with PSF photometry and photometric errors in all bands, along with all the properties obtained in this work.

A.5 Additional plots and tables

In Figure A.3 we show $\mathrm{SFR_{Ly\alpha}}$ vs $\mathrm{SFR_{SED}}$ in 6 independent redshift intervals (see Section 6.1.6.3 for discussion). In Figure A.4 we show the evolution of median $\mathrm{EW_0}$ with redshift. We provide the full measurements of w_0 for different ranges of redshifts and galaxy properties ($\mathrm{M_\star}$ and $\mathrm{M_{UV}}$) in Table A.6.

---

[1]https://academic.oup.com/mnras/article/493/1/141/5704403#supplementary-data

Table A.5: The results of fitting different Lyα LFs 10,000 times with a Schechter function (and a single power-law, for comparison) at the appropriate luminosity range (* fitting only up to $10^{43.3}\,\mathrm{erg\,s^{-1}}$), when using SC4K only and when combining SC4K with deeper surveys (S-SC4K). As part of each fit we also integrate our Lyα LFs to obtain $\rho_{\mathrm{Ly}\alpha}$ (integral of the Schechter component, Sch), derived for different redshift bins, down to $1.75 \times 10^{41}\,\mathrm{erg\,s^{-1}}$, corresponding to $0.04\,L^{\star}_{z=3}$ from Gronwall et al. (2007); see Section 4.1. All errors are the 16th and 84th percentiles for all 10,000 realisations per LF estimation which, due to degeneracies in the parameters, can sometimes exaggerate the errors on individual parameters, so these can be seen as conservative. We also provide a comparison (ratio) between reduced χ^2 for Schechter and power-law fits ($\chi^2_{\mathrm{Sch}}/\chi^2_{\mathrm{PL}}$); values below 1 indicate that a Schechter fit performs better, while a large value indicates that a simple power-law fit provides a lower reduced χ^2.

| Redshift slice (S-)SC4K	α	$\log_{10} L^*_{\mathrm{Ly}\alpha}$ (erg s^{-1})	$\log_{10} \Phi^*_{\mathrm{Ly}\alpha}$ (Mpc^{-3})	$\rho_{\mathrm{Ly}\alpha}/10^{40}$ (Sch) (erg s^{-1} Mpc^{-3})	Power-law (PL) (A $\log_{10}$ L+B)	$\chi^2_{\mathrm{Sch}}/\chi^2_{\mathrm{PL}}$	Reference(s) (Table A.4)
$z = 2.2 \pm 0.1^*$	-1.8 ± 0.2 (fix)	$42.69^{+0.13}_{-0.11}$	$-3.33^{+0.21}_{-0.26}$	$0.48^{+0.04}_{-0.04}$	$-1.24^{+0.08}_{-0.09}$, $49.3^{+3.6}_{-3.6}$	0.3	2.1
$z = 2.5 \pm 0.1^*$	-1.8 ± 0.2 (fix)	$42.76^{+0.07}_{-0.07}$	$-3.23^{+0.14}_{-0.15}$	$0.73^{+0.18}_{-0.13}$	$-2.34^{+0.19}_{-0.20}$, $96.9^{+8.5}_{-8.1}$	2.3	SC4K only
$z = 2.8 \pm 0.1^*$	-1.8 ± 0.2 (fix)	$42.83^{+0.36}_{-0.19}$	$-3.27^{+0.58}_{-0.75}$	$0.84^{+1.12}_{-0.41}$	$-2.66^{+1.03}_{-1.01}$, $110.7^{+43.4}_{-44.2}$	1.0	SC4K only
$z = 3.0 \pm 0.1^*$	-1.8 ± 0.2 (fix)	$42.64^{+0.06}_{-0.05}$	$-2.54^{+0.16}_{-0.16}$	$2.54^{+0.87}_{-0.62}$	$-3.17^{+0.28}_{-0.29}$, $132.9^{+12.3}_{-12.2}$	1.1	SC4K only
$z = 3.2 \pm 0.1^*$	-1.8 ± 0.2 (fix)	$42.80^{+0.09}_{-0.07}$	$-3.01^{+0.16}_{-0.19}$	$1.35^{+0.41}_{-0.29}$	$-2.41^{+0.25}_{-0.27}$, $100.0^{+11.4}_{-10.6}$	0.6	SC4K only
$z = 3.3 \pm 0.1^*$	-1.8 ± 0.2 (fix)	$42.68^{+0.07}_{-0.06}$	$-2.70^{+0.16}_{-0.16}$	$1.95^{+0.64}_{-0.44}$	$-2.98^{+0.25}_{-0.26}$, $124.7^{+11.3}_{-10.6}$	1.5	SC4K only
$z = 3.7 \pm 0.1$	-1.8 ± 0.2 (fix)	$43.03^{+0.18}_{-0.15}$	$-4.09^{+0.41}_{-0.40}$	$0.21^{+0.17}_{-0.09}$	$-3.18^{+0.68}_{-0.85}$, $133.0^{+36.7}_{-29.3}$	2.4	SC4K only
$z = 4.1 \pm 0.1$	-1.8 ± 0.2 (fix)	$42.83^{+0.17}_{-0.15}$	$-3.49^{+0.46}_{-0.43}$	$0.49^{+0.49}_{-0.19}$	$-3.11^{+0.71}_{-0.85}$, $129.8^{+36.8}_{-30.7}$	0.8	SC4K only
$z = 4.6 \pm 0.1$	-1.8 ± 0.2 (fix)	$43.15^{+0.16}_{-0.15}$	$-3.92^{+0.37}_{-0.38}$	$0.42^{+0.32}_{-0.68}$	$-2.98^{+0.62}_{-0.68}$, $124.6^{+29.6}_{-27.0}$	1.5	SC4K only
$z = 4.8 \pm 0.1$	-1.8 ± 0.2 (fix)	$42.98^{+0.17}_{-0.14}$	$-3.62^{+0.48}_{-0.46}$	$0.56^{+0.65}_{-0.28}$	$-3.99^{+0.89}_{-1.01}$, $168.1^{+43.9}_{-38.5}$	1.4	SC4K only
$z = 5.1 \pm 0.1$	-1.8 ± 0.2 (fix)	$43.30^{+0.23}_{-0.19}$	$-4.36^{+0.59}_{-0.54}$	$0.24^{+0.33}_{-0.13}$	$-3.88^{+1.09}_{-1.51}$, $163.8^{+65.7}_{-47.4}$	2.0	SC4K only
$z = 5.3 \pm 0.1$	-1.8 ± 0.2 (fix)	$43.30^{+0.28}_{-0.20}$	$-4.22^{+0.71}_{-0.73}$	$0.33^{+0.75}_{-0.21}$	$-3.88^{+1.47}_{-1.68}$, $164.2^{+73.2}_{-64.0}$	0.8	SC4K only
$z = 5.8 \pm 0.1$	-1.8 ± 0.2 (fix)	$43.35^{+0.24}_{-0.19}$	$-4.19^{+0.66}_{-0.67}$	$0.39^{+0.83}_{-0.25}$	$-3.55^{+1.15}_{-1.49}$, $149.7^{+65.3}_{-50.2}$	0.9	SC4K only
$z = 2.2 \pm 0.1^*$	$-2.00^{+0.15}_{-0.15}$	$42.82^{+0.13}_{-0.11}$	$-3.59^{+0.22}_{-0.28}$	$0.52^{+0.05}_{-0.05}$	$-1.54^{+0.07}_{-0.07}$, $62.1^{+3.1}_{-3.0}$	0.6	2.1, 6.1, 12
$z = 2.5 \pm 0.1^*$	$-1.72^{+0.15}_{-0.15}$	$42.70^{+0.09}_{-0.08}$	$-3.10^{+0.17}_{-0.20}$	$0.74^{+0.08}_{-0.07}$	$-1.33^{+0.07}_{-0.07}$, $53.6^{+2.9}_{-2.9}$	0.6	2.1, 5.1
$z = 2.8 \pm 0.1^*$	$-1.73^{+0.20}_{-0.21}$	$42.78^{+0.16}_{-0.12}$	$-3.18^{+0.27}_{-0.35}$	$0.77^{+0.10}_{-0.09}$	$-1.28^{+0.08}_{-0.08}$, $51.3^{+3.5}_{-3.5}$	0.8	2.1, 5.1
$z = 3.0 \pm 0.1^*$	$-1.58^{+0.17}_{-0.17}$	$42.75^{+0.12}_{-0.09}$	$-3.00^{+0.21}_{-0.25}$	$0.88^{+0.10}_{-0.09}$	$-1.15^{+0.07}_{-0.07}$, $46.0^{+3.0}_{-2.8}$	0.6	2.1, 5.1
$z = 3.2 \pm 0.1^*$	$-1.70^{+0.17}_{-0.17}$	$42.85^{+0.15}_{-0.11}$	$-3.20^{+0.24}_{-0.31}$	$0.84^{+0.09}_{-0.09}$	$-1.15^{+0.07}_{-0.07}$, $45.9^{+2.9}_{-3.0}$	0.7	2.1, 5.1
$z = 3.3 \pm 0.1^*$	$-1.62^{+0.17}_{-0.17}$	$42.76^{+0.12}_{-0.10}$	$-3.05^{+0.22}_{-0.26}$	$0.85^{+0.10}_{-0.09}$	$-1.17^{+0.07}_{-0.07}$, $46.9^{+3.0}_{-2.9}$	0.6	2.1, 5.1
$z = 3.7 \pm 0.1$	$-2.57^{+0.23}_{-0.21}$	$43.23^{+0.37}_{-0.23}$	$-4.54^{+0.61}_{-0.91}$	$1.01^{+0.20}_{-0.16}$	$-2.01^{+0.12}_{-0.14}$, $82.2^{+5.7}_{-5.2}$	0.8	2.2, 5.1
$z = 4.1 \pm 0.1$	$-2.23^{+0.30}_{-0.24}$	$42.96^{+0.28}_{-0.22}$	$-3.79^{+0.53}_{-0.66}$	$0.87^{+0.15}_{-0.11}$	$-1.93^{+0.12}_{-0.14}$, $78.8^{+6.1}_{-5.2}$	0.9	2.2, 5.1
$z = 4.6 \pm 0.1$	$-2.38^{+0.20}_{-0.19}$	$43.32^{+0.24}_{-0.16}$	$-4.34^{+0.39}_{-0.57}$	$1.19^{+0.40}_{-0.30}$	$-1.80^{+0.10}_{-0.10}$, $73.6^{+4.1}_{-4.1}$	0.9	2.3, 3, 5.2, 10
$z = 4.8 \pm 0.1$	$-2.28^{+0.22}_{-0.22}$	$43.14^{+0.19}_{-0.15}$	$-3.98^{+0.36}_{-0.46}$	$1.12^{+0.37}_{-0.27}$	$-1.92^{+0.12}_{-0.12}$, $78.5^{+5.2}_{-4.9}$	0.8	2.3, 3, 5.2, 10
$z = 5.1 \pm 0.1$	$-2.46^{+0.22}_{-0.20}$	$43.41^{+0.28}_{-0.21}$	$-4.58^{+0.54}_{-0.67}$	$1.27^{+0.48}_{-0.32}$	$-2.00^{+0.13}_{-0.15}$, $82.1^{+6.3}_{-5.8}$	0.7	2.3, 3, 5.2, 10
$z = 5.3 \pm 0.1$	$-1.92^{+0.22}_{-0.19}$	$43.21^{+0.14}_{-0.13}$	$-3.70^{+0.30}_{-0.32}$	$1.08^{+0.21}_{-0.16}$	$-1.80^{+0.13}_{-0.14}$, $73.6^{+6.2}_{-5.3}$	0.2	5.3, 9.4, 11
$z = 5.8 \pm 0.1$	$-1.95^{+0.20}_{-0.18}$	$43.26^{+0.13}_{-0.13}$	$-3.78^{+0.29}_{-0.30}$	$1.10^{+0.21}_{-0.16}$	$-1.74^{+0.12}_{-0.13}$, $71.0^{+5.4}_{-5.0}$	0.2	5.3, 9.4, 11
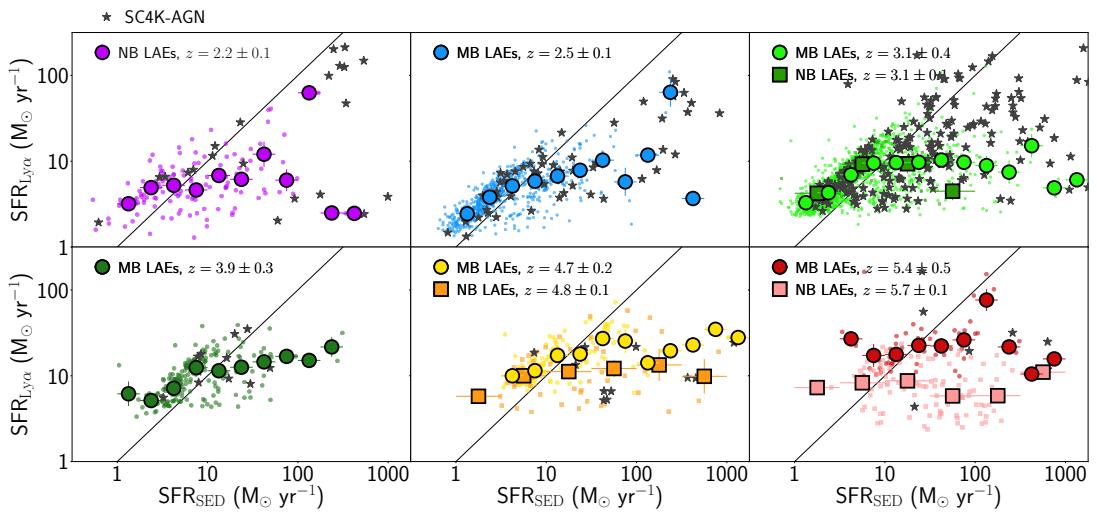
Figure A.3: Emission line-based SFR vs SED-fitting SFR for the full sample of LAEs at different redshift ranges. Coloured circles (squares) are the median bin for MBs (NBs) and individual points are plotted as scatter in the background. The black line is the 1-to-1 ratio. While the two approaches roughly follow the 1-to-1 ratio, there are some key differences. Similar to what is observed in Figure 6.3, median $SFR_{Ly\alpha}$ is slightly higher than SFR_{SED} for $SFR_{SED} < 10\,M_\odot\,yr^{-1}$. However, $SFR_{Ly\alpha}$ seems to saturate at median $SFR_{Ly\alpha} \approx 10-30\,M_\odot\,yr^{-1}$.
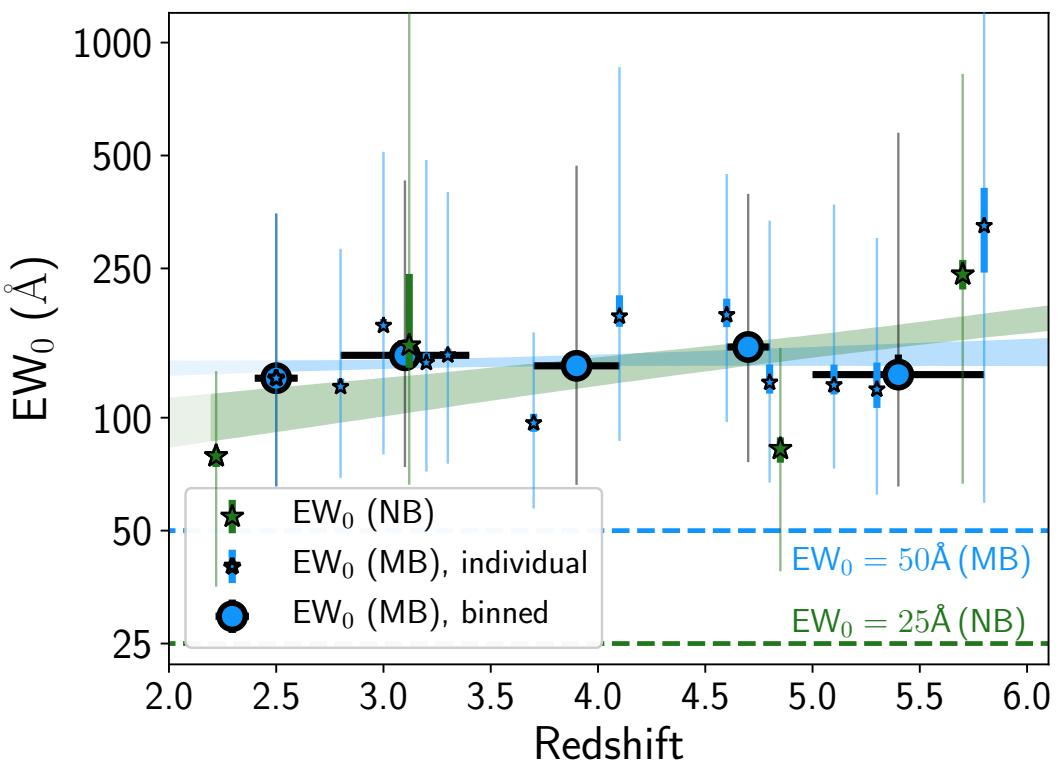
Figure A.4: Global median EW_0 evolution with redshift. The median EW_0 values for medium (narrow) bands are shown as blue circles (green stars). Blue stars are the measurements for individual MBs. The thin (thick) error bars are the 16th and 84th percentiles of the EW_0 distribution (divided by the Poissonian error $\sqrt{N}$). The median and errors of EW_0 can be found in Table 5.1. Blue (green) shaded region is the 1σ contour obtained by perturbing the EW_0 within the thick error bars for medium (narrow) band selected LAE. We find evidence of little EW_0 evolution with redshift for the global sample of LAEs, with the median EW_0 remaining roughly constant at $\sim 140\,\text{Å}$, although there is a tentative higher EW_0 at $z = 5.7$, albeit with large error bars.

Table A.6: EW_0 scale length (w_0) for different redshift bins, derived as fully detailed in Section 6.1.2.1. (1) Subset of SC4K; (2) Filter type (MB/NB) and whether this measurement is for LAEs selected from an individual filter or from multiple filters binned together; (3) w_0 derived directly from observed counts; (4) $\chi^2_{\rm red}$ of (3); (5) w_0 derived directly from observed counts [$EW_0 < 240$ Å]; (6) $\chi^2_{\rm red}$ of (5); (7) w_0 derived by perturbing EW_0; (8) $\chi^2_{\rm red}$ of (7); (9) w_0 derived by perturbing EW_0 [$EW_0 < 240$ Å]; (10) $\chi^2_{\rm red}$ of (9); $\star$Not enough sources to constrain w_0 (less than 3 bins with 5 sources). †Fit done for $EW_0 > 100$ Å as discussed in Section 6.1.2.1 as we are significantly incomplete for low EW_0, so we only fit exponential decay after the distribution peak at ~ 100 Å.

(1) Subset	(2) Filters	(3) w_0 (Å)	(4) $\chi^2_{\rm red}$	(5) $w_{0,[EW<240Å]}$ (Å)	(6) $\chi^2_{\rm red}$	(7) $w_{0,p}$ (Å)	(8) $\chi^2_{\rm red}$	(9) $w_{0,p,[EW<240Å]}$ (Å)	(10) $\chi^2_{\rm red}$
$z = 2.5 \pm 0.1$	MB, single	104^{+13}_{-13}	2.31	79^{+15}_{-15}	2.77	117^{+12}_{-12}	0.73	94^{+12}_{-12}	0.12
$z = 2.8 \pm 0.1$	MB, single	98^{+12}_{-12}	0.59	88^{+12}_{-12}	0.37	108^{+12}_{-13}	0.15	100^{+17}_{-15}	0.28
$z = 3.0 \pm 0.1$	MB, single	172^{+14}_{-14}	1.04	120^{+22}_{-22}	1.74	170^{+16}_{-16}	0.72	126^{+22}_{-17}	0.37
$z = 3.2 \pm 0.1$	MB, single	109^{+14}_{-14}	1.77	97^{+16}_{-16}	1.17	128^{+15}_{-14}	0.34	111^{+20}_{-18}	0.18
$z = 3.3 \pm 0.1$	MB, single	113^{+13}_{-13}	1.76	118^{+19}_{-19}	1.21	139^{+13}_{-14}	0.40	113^{+16}_{-14}	0.42
$z = 3.7 \pm 0.1$	MB, single	83^{+24}_{-24}	1.64	72^{+26}_{-26}	1.93	90^{+24}_{-19}	0.20	80^{+24}_{-17}	0.12
$z = 4.1 \pm 0.1$	MB, single	257^{+41}_{-41}	0.66	216^{+100}_{-100}	0.91	251^{+81}_{-65}	0.31	171^{+149}_{-54}	0.08
$z = 4.6 \pm 0.1$	MB, single	486^{+183}_{-183}	0.65	318^{+413}_{-413}	1.30	600^{+325}_{-223}	0.44	413^{+43491}_{-232}	0.08
$z = 4.8 \pm 0.1$	MB, single	93^{+17}_{-17}	0.42	82^{+14}_{-14}	0.18	119^{+41}_{-31}	0.21	94^{+30}_{-22}	0.12
$z = 5.0 \pm 0.1$	MB, single	108^{+26}_{-26}	0.75	85^{+24}_{-24}	0.75	143^{+84}_{-42}	0.40	103^{+39}_{-24}	0.19
$z = 5.3 \pm 0.1^\star$	MB, single	-	-	-	-	-	-	-	-
$z = 5.8 \pm 0.1^\star$	MB, single	-	-	-	-	-	-	-	-
$z = 2.2 \pm 0.1$	NB, single	174^{+95}_{-95}	4.95	174^{+95}_{-95}	4.95	143^{+22}_{-19}	0.52	131^{+26}_{-19}	0.62
$z = 3.1 \pm 0.1^\star$	NB, single	-	-	-	-	-	-	-	-
$z = 4.8 \pm 0.1$	NB, single	86^{+24}_{-24}	1.05	86^{+24}_{-24}	1.05	151^{+98}_{-50}	0.48	101^{+31}_{-26}	0.32
$z = 5.7 \pm 0.1$	NB, single	355^{+71}_{-71}	0.93	188^{+57}_{-57}	0.70	477^{+154}_{-106}	0.45	124^{+34}_{-27}	0.14
$z = 2.5 \pm 0.1$	MB, bin	104^{+13}_{-13}	2.31	79^{+15}_{-15}	2.77	117^{+12}_{-12}	0.73	94^{+12}_{-12}	0.12
$z = 3.1 \pm 0.4$	MB, bin	134^{+11}_{-11}	2.00	109^{+12}_{-12}	0.85	149^{+11}_{-11}	2.26	116^{+13}_{-12}	1.04
$z = 3.9 \pm 0.3$	MB, bin	118^{+18}_{-18}	1.29	90^{+21}_{-21}	1.69	120^{+18}_{-17}	0.20	103^{+21}_{-18}	0.15
$z = 4.7 \pm 0.2$	MB, bin	119^{+21}_{-21}	1.17	93^{+19}_{-19}	0.83	158^{+34}_{-27}	0.18	114^{+43}_{-23}	0.14
$z = 2.5 \pm 0.1$	MB, bin	95^{+24}_{-24}	1.66	70^{+21}_{-21}	1.97	125^{+40}_{-31}	0.24	90^{+22}_{-18}	0.21
Full sample	MB, bin	130^{+11}_{-11}	3.57	100^{+11}_{-11}	1.55	143^{+10}_{-11}	4.01	110^{+11}_{-11}	0.99
Full sample	NB, bin	109^{+13}_{-13}	1.53	102^{+14}_{-14}	0.98	151^{+18}_{-17}	0.45	102^{+15}_{-14}	0.15
Full sample	MB+NB, bin	129^{+11}_{-11}	4.19	99^{+11}_{-11}	1.22	147^{+11}_{-11}	4.46	109^{+11}_{-11}	0.96
$8< \log_{10} (M_\star/M_\odot < 9^\dagger$	MB+NB, bin	175^{+14}_{-14}	1.62	179^{+84}_{-84}	2.77	264^{+14}_{-16}	0.55	530^{+582}_{-163}	1.42
$9< \log_{10} (M_\star/M_\odot < 10$	MB+NB, bin	85^{+11}_{-11}	2.87	74^{+11}_{-11}	1.75	101^{+11}_{-10}	2.03	89^{+11}_{-11}	0.96
$10< \log_{10} (M_\star/M_\odot < 11$	MB+NB, bin	77^{+13}_{-13}	2.14	60^{+13}_{-13}	2.32	89^{+13}_{-13}	0.81	68^{+12}_{-11}	0.66
$-20< M_{UV} < -19^\dagger$	MB+NB, bin	182^{+13}_{-13}	1.35	253^{+103}_{-103}	1.29	263^{+15}_{-15}	0.50	470^{+212}_{-126}	0.97
$-21< M_{UV} < -20$	MB+NB, bin	77^{+10}_{-10}	1.52	71^{+11}_{-11}	1.00	93^{+10}_{-10}	1.09	90^{+11}_{-11}	1.27
$-22< M_{UV} < -21$	MB+NB, bin	55^{+11}_{-11}	1.78	50^{+11}_{-11}	1.91	63^{+11}_{-11}	0.39	58^{+11}_{-10}	0.07

Bibliography

Abbott T. M. C., et al., 2018, Phys. Rev. D, 98, 043526

Acquaviva V., Vargas C., Gawiser E., Guaita L., 2012, ApJL, 751, L26

Adams T. F., 1972, ApJ, 174, 439

Adams J. J., et al., 2011, ApJS, 192, 5

Agertz O., Kravtsov A. V., 2015, ApJ, 804, 18

Aird J., Nandra K., et al., 2010, MNRAS, 401, 2531

Alavi A., et al., 2016, ApJ, 832, 56

Amorín R., et al., 2017, Nature Astronomy, 1, 0052

Ando M., Ohta K., Iwata I., Akiyama M., Aoki K., Tamura N., 2006, ApJL, 645, L9

Arnouts S., et al., 2005, ApJL, 619, L43

Arrabal Haro P., et al., 2018, MNRAS, 478, 3740

Arrabal Haro P., et al., 2020, MNRAS, 495, 1807

Atek H., Kunth D., Hayes M., Östlin G., Mas-Hesse J. M., 2008, A&A, 488, 491

Atek H., Richard J., Kneib J.-P., et al., 2015, ApJ, 800, 18

Bacon R., et al., 2015, AAP, 575, A75

Bartelmann M., 2010, Classical and Quantum Gravity, 27, 233001

Bayliss M. B., Wuyts E., Sharon K., Gladders M. D., Hennawi J. F., Koester B. P., Dahle H., 2010, ApJ, 720, 1559

Bertin E., Arnouts S., 1996, AAPS, 117, 393

Bertin E., Mellier Y., Radovich M., Missonnier G., Didelon P., Morin B., 2002, in Bohlender D. A., Durand D., Handley T. H., eds, ASPCS Vol. 281, Astronomical Data Analysis Software and Systems XI. p. 228

Bielby R. M., et al., 2016, MNRAS, 456, 4061

Bina D., Pelló R., Richard J., et al., 2016, A&A, 590, A14

Blanc G. A., et al., 2011, ApJ, 736, 31

Bond J. R., Arnett W. D., Carr B. J., 1984, ApJ, 280, 825

Bond N. A., Gawiser E., Koekemoer A. M., 2011, ApJ, 729, 48

Bondi M., Ciliegi P., Schinnerer E., Smolčić V., Jahnke K., Carilli C., Zamorani G., 2008, ApJ, 681, 1129

Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, A&A, 622, A103

Borisova E., et al., 2016, ApJ, 831, 39

Boselli A., Gavazzi G., Donas J., Scodeggio M., 2001, AJ, 121, 753

Bouwens R. J., Illingworth G. D., Blakeslee J. P., Franx M., 2006, ApJ, 653, 53

Bouwens R. J., et al., 2010, ApJL, 709, L133

Bouwens R. J., et al., 2012, ApJ, 754, 83

Bouwens R. J., et al., 2014a, ApJ, 793, 115

Bouwens R. J., et al., 2014b, ApJ, 795, 126

Bouwens R. J., et al., 2015, ApJ, 803, 34

Bouwens R. J., Smit R., Labbé I., Franx M., Caruana J., Oesch P., Stefanon M., Rasappu N., 2016, ApJ, 831, 176

Bowler R. A. A., et al., 2014, MNRAS, 440, 2810

Bowler R. A. A., Dunlop J. S., McLure R. J., McLeod D. J., 2017, MNRAS, 466, 3612

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Brocklehurst M., 1971, MNRAS, 153, 471

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Bundy K., Ellis R. S., Conselice C. J., 2005, ApJ, 625, 621

Bunker A. J., Warren S. J., Hewett P. C., Clements D. L., 1995, MNRAS, 273, 513

Bunker A. J., Stanway E. R., Ellis R. S., McMahon R. G., 2004, MNRAS, 355, 374

Bunker A. J., et al., 2010, MNRAS, 409, 855

Calhau J., Sobral D., Stroe A., Best P., Smail I., Lehmer B., Harrison C., Thomson A., 2017, MNRAS, 464, 303

Calhau J., et al., 2020, MNRAS, 493, 3341

Cantalupo S., Porciani C., Lilly S. J., 2008, ApJ, 672, 48

Cantalupo S., Lilly S. J., Haehnelt M. G., 2012, MNRAS, 425, 1992

Capak P., et al., 2007, ApJS, 172, 99

Caputi K. I., Cirasuolo M., Dunlop J. S., McLure R. J., Farrah D., Almaini O., 2011, MNRAS, 413, 162

Caputi K. I., et al., 2015, ApJ, 810, 73

Carr B. J., Bond J. R., Arnett W. D., 1984, ApJ, 277, 445

Caruana J., Bunker A. J., Wilkins S. M., Stanway E. R., Lorenzoni S., Jarvis M. J., Ebert H., 2014, MNRAS, 443, 2831

Caruana J., Wisotzki L., Herenz E. C., et al., 2018, MNRAS, 473, 30

Cassata P., et al., 2011, A&A, 525, A143

Cassata P., Tasca L. A. M., Le Fèvre O., et al., 2015, A&A, 573, A24

Cen R., Haiman Z., 2000, ApJL, 542, L75

Chabrier G., 2003, Public. of the Astron. Soc. Pac., 115, 763

Charlot S., Fall S. M., 1993, ApJ, 415, 580

Charlot S., Fall S. M., 2000, ApJ, 539, 718

Ciardullo R., et al., 2012, ApJ, 744, 110

Ciardullo R., et al., 2014, ApJ, 796, 64

Civano F., et al., 2016, ApJ, 819, 62

Clowe D., Bradač M., Gonzalez A. H., Markevitch M., Randall S. W., Jones C., Zaritsky D., 2006, ApJL, 648, L109

Colbert J. W., Teplitz H., Atek H., et al., 2013, ApJ, 779, 34

Condon J. J., 1992, Annual Review of Astron and Astrophys, 30, 575

Cortzen I., et al., 2019, MNRAS, 482, 1618

Couchman H. M. P., Rees M. J., 1986, MNRAS, 221, 53

Cowie L. L., Hu E. M., 1998, AJ, 115, 1319

Cowie L. L., Barger A. J., Trouille L., 2009, ApJ, 692, 1476

Crain R. A., Schaye J., Bower R. G., et al., 2015, MNRAS, 450, 1937

Curtis-Lake E., et al., 2012a, MNRAS, 422, 1425

Curtis-Lake E., et al., 2012b, MNRAS, 422, 1425

Daddi E., et al., 2007, ApJ, 670, 156

Davidzon I., et al., 2017, A&A, 605, A70

Dawson S., Rhoads J. E., Malhotra S., Stern D., Wang J., Dey A., Spinrad H., Jannuzi B. T., 2007, ApJ, 671, 1227

De Barros S., et al., 2017, A&A, 608, A123

Delvecchio I., Gruppioni C., et al., 2014, MNRAS, 439, 2736

Dijkstra M., 2017, preprint, (arXiv:1704.03416)

Dijkstra M., Gronke M., Venkatesan A., 2016, ApJ, 828, 71

Drake A. B., et al., 2017a, MNRAS, 471, 267

Drake A. B., et al., 2017b, A&A, 608, A6

Dressler A., Henry A., Martin C. L., Sawicki M., McCarthy P., Villaneuva E., 2015, ApJ, 806, 19

Duncan K., Conselice C. J., 2015, MNRAS, 451, 2030

Duncan K., et al., 2014, MNRAS, 444, 2960

Einasto J., Saar E., Kaasik A., Chernin A. D., 1974, Nature, 252, 111

Ellis R. S., McLure R. J., Dunlop J. S., et al., 2013, ApJL, 763, L7

Elvis M., et al., 2009, ApJS, 184, 158

Erb D. K., et al., 2014, ApJ, 795, 33

Faisst A. L., 2016, ApJ, 829, 99

Faisst A. L., et al., 2016, ApJ, 821, 122

Fan X., et al., 2006, ApJ, 132, 117

Finkelstein S. L., 2016, PASA, 33, e037

Finkelstein S. L., Rhoads J. E., Malhotra S., Grogin N., 2009, ApJ, 691, 465

Finkelstein S. L., et al., 2012, ApJ, 756, 164

Finkelstein S. L., et al., 2015, ApJ, 810, 71

Fitzpatrick E. L., 1999, Public. of the Astron. Soc. Pac., 111, 63

Fitzpatrick E. L., Massa D., 2007, ApJ, 663, 320

Fontanot F., Cristiani S., Vanzella E., 2012, MNRAS, 425, 1413

Forrest B., Tran K.-V. H., et al., 2017, ApJL, 838, L12

Franck J. R., McGaugh S. S., 2016, ApJ, 833, 15

Fukugita M., Kawasaki M., 1994, MNRAS, 269, 563

Fumagalli M., Patel S. G., Franx M., others. 2012, ApJL, 757, L22

Garel T., Guiderdoni B., Blaizot J., 2016, MNRAS, 455, 3436

Garn T., Best P. N., 2010, MNRAS, 409, 421

Gawiser E., et al., 2006, ApJL, 642, L13

Gawiser E., Francke H., Lai K., et al., 2007, ApJ, 671, 278

Geach J. E., et al., 2014, Nature, 516, 68

Giallongo E., et al., 2015, A&A, 578, A83

Giavalisco M., Steidel C. C., Macchetto F. D., 1996, ApJ, 470, 189

González V., Labbé I., Bouwens R. J., Illingworth G., Franx M., Kriek M., 2011, ApJL, 735, L34

Grazian A., et al., 2015, A&A, 575, A96

Griffin M. J., et al., 2010, A&A, 518, L3

Gronwall C., Ciardullo R., et al., 2007, ApJ, 667, 79

Guaita L., et al., 2010, ApJ, 714, 255

Guaita L., Melinder J., Hayes M., et al., 2015, A&A, 576, A51

Guth A. H., 1981, Phys. Rev. D, 23, 347

Hagen A., Zeimann G. R., et al., 2016, ApJ, 817, 79

Harikane Y., et al., 2018, ApJ, 859, 84

Hashimoto T., et al., 2017, A&A, 608, A10

Hathi N. P., et al., 2016, A&A, 588, A26

Hayashi M., et al., 2018, Publications of the ASJ, 70, S17

Hayes M., et al., 2010a, Nature, 464, 562

Hayes M., Schaerer D., Östlin G., 2010b, AAP, 509, L5

Hayes M., Schaerer D., Östlin G., Mas-Hesse J. M., Atek H., Kunth D., 2011, ApJ, 730, 8

Hayes M. J., Runnholm A., Gronke M., Scarlata C., 2020, arXiv e-prints, p. arXiv:2006.03232

Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663

Henry A., Scarlata C., Martin C. L., Erb D., 2015, ApJ, 809, 19

Herenz E. C., Urrutia T., Wisotzki L., et al., 2017, A&A, 606, A12

Hildebrandt H., et al., 2017, MNRAS, 465, 1454

Hu E. M., Cowie L. L., Capak P., McMahon R. G., Hayashino T., Komiyama Y., 2004, AJ, 127, 563

Hu E. M., Cowie L. L., Barger A. J., Capak P., Kakazu Y., Trouille L., 2010, ApJ, 725, 394

Hu E. M., Cowie L. L., Songaila A., Barger A. J., Rosenwasser B., Wold I. G. B., 2016, ApJL, 825, L7

Ilbert O., et al., 2009, ApJ, 690, 1236

Ilbert O., et al., 2013, A&A, 556, A55

Iye M., et al., 2006, Nature, 443, 186

Izotov Y. I., Schaerer D., Thuan T. X., Worseck G., Guseva N. G., Orlitová I., Verhamme A., 2016, MNRAS, 461, 3683

Izotov Y. I., Guseva N. G., Fricke K. J., Henkel C., Schaerer D., 2017, MNRAS, 467, 4118

Izotov Y. I., Schaerer D., Worseck G., Guseva N. G., Thuan T. X., Verhamme A., Orlitová I., Fricke K. J., 2018, MNRAS, 474, 4514

Jaacks J., Nagamine K., Choi J. H., 2012, MNRAS, 427, 403

Jin S., et al., 2018, ApJ, 864, 56

Karim A., et al., 2011, ApJ, 730, 61

Karman W., et al., 2015, AAP, 574, A11

Kashikawa N., et al., 2011, ApJ, 734, 119

Kashikawa N., et al., 2012, ApJ, 761, 85

Kennicutt Jr. R. C., 1998, ARAA, 36, 189

Kennicutt Robert C. J., Tamblyn P., Congdon C. E., 1994, ApJ, 435, 22

Khostovan A. A., Sobral D., Mobasher B., Best P. N., Smail I., Stott J. P., Hemmati S., Nayyeri H., 2015, MNRAS, 452, 3948

Khostovan A. A., Sobral D., Mobasher B., Smail I., Darvish B., Nayyeri H., Hemmati S., Stott J. P., 2016, MNRAS, 463, 2363

Khostovan A., et al., 2018, MNRAS, submitted,

Khostovan A. A., et al., 2019, MNRAS, 489, 555

Konno A., et al., 2014, ApJ, 797, 16

Konno A., Ouchi M., Nakajima K., Duval F., Kusakabe H., Ono Y., Shimasaku K., 2016, ApJ, 823, 20

Konno A., et al., 2018, Publications of the ASJ, 70, S16

Koo D. C., Kron R. T., 1980, PASP, 92, 537

Kriek M., Shapley A. E., Reddy N. A., et al., 2015, ApJS, 218, 15

Kubo M., Uchimoto Y. K., et al., 2013, ApJ, 778, 170

Kusakabe H., et al., 2018, Publications of the ASJ, 70, 4

Kusakabe H., et al., 2020, arXiv e-prints, p. arXiv:2003.12083

Lai K., et al., 2008, ApJ, 674, 70

Laigle C., et al., 2015, MNRAS, 446, 2744

Laigle C., et al., 2016, ApJS, 224, 24

Laporte N., Nakajima K., Ellis R. S., Zitrin A., Stark D. P., Mainali R., Roberts-Borsani G. W., 2017, ApJ, 851, 40

Le Fèvre O., et al., 2015, A&A, 576, A79

Leclercq F., Bacon R., Wisotzki L., et al., 2017, A&A, 608, A8

Lehmer B. D., Alexander D. M., Geach J. E., Smail I., et al., 2009, ApJ, 691, 687

Lehmer B. D., et al., 2016, ApJ, 825, 7

Leitherer C., Heckman T. M., 1995, ApJS, 96, 9

Lilly S. J., Le Fevre O., Hammer F., Crampton D., 1996, ApJL, 460, L1

Lilly S. J., Le Fèvre O., et al., 2007, ApJS, 172, 70

Lowenthal J. D., Hogan C. J., Leach R. W., Schmidt G. D., Foltz C. B., 1990,
ApJ, 357, 3

Lowenthal J. D., Hogan C. J., Green R. F., Caulet A., Woodgate B. E., Brown
L., Foltz C. B., 1991, ApJL, 377, L73

Lutz D., et al., 2011, A&A, 532, A90

Madau P., 1995, ApJ, 441, 18

Madau P., Dickinson M., 2014, ARAA, 52, 415

Madau P., Ferguson H. C., Dickinson M. E., Giavalisco M., Steidel C. C., Fruchter
A., 1996, MNRAS, 283, 1388

Madau P., Pozzetti L., Dickinson M., 1998, ApJ, 498, 106

Malhotra S., Rhoads J. E., 2004, ApJL, 617, L5

Malhotra S., Rhoads J. E., Finkelstein S. L., Hathi N., Nilsson K., McLinden E.,
Pirzkal N., 2012, ApJL, 750, L36

Marino R. A., et al., 2018, ApJ, 859, 53

Markevitch M., Gonzalez A. H., David L., Vikhlinin A., Murray S., Forman W.,
Jones C., Tucker W., 2002, ApJL, 567, L27

Marques-Chaves R., et al., 2020, MNRAS, 492, 1257

Martin C. L., Sawicki M., 2004, ApJ, 603, 414

Maseda M. V., et al., 2018, ApJL, 865, L1

Maseda M. V., et al., 2020, MNRAS, 493, 5120

Mason C. A., Gronke M., 2020, MNRAS, 499, 1395

Matsuda Y., et al., 2004, AJ, 128, 569

Matsuda Y., Smail I., Geach J. E., et al., 2011, MNRAS, 416, 2041

Matthee J., Schaye J., 2019, MNRAS, 484, 915

Matthee J., Sobral D., 2020, in da Cunha E., Hodge J., Afonso J., Pentericci L., Sobral D., eds, IAU Symposium Vol. 352, IAU Symposium. pp 21–25 (arXiv:1911.04774), doi:10.1017/S1743921319009451

Matthee J. J. A., et al., 2014, MNRAS, 440, 2375

Matthee J., Sobral D., Santos S., Röttgering H., Darvish B., Mobasher B., 2015, MNRAS, 451, 400

Matthee J., Sobral D., Oteo I., Best P., Smail I., Röttgering H., Paulino-Afonso A., 2016, MNRAS, 458, 449

Matthee J., Sobral D., Best P., Khostovan A. A., Oteo I., Bouwens R., Röttgering H., 2017a, MNRAS, 465, 3637

Matthee J., Sobral D., Best P., Smail I., Bian F., Darvish B., Röttgering H., Fan X., 2017b, MNRAS, 471, 629

Matthee J., Sobral D., Darvish B., Santos S., Mobasher B., Paulino-Afonso A., Röttgering H., Alegre L., 2017c, MNRAS, 472, 772

Matthee J., et al., 2017d, ApJ, 851, 145

Matthee J., Sobral D., Gronke M., Paulino-Afonso A., Stefanon M., Röttgering H., 2018, A&A, 619, A136

Matthee J., et al., 2019, ApJ, 881, 124

McCracken H. J., et al., 2012, A&A, 544, A156

McGreer I. D., et al., 2013, ApJ, 768, 105

McLure R. J., Cirasuolo M., Dunlop J. S., Foucaud S., Almaini O., 2009, MNRAS, 395, 2196

McLure R. J., Dunlop J. S., Cirasuolo M., Koekemoer A. M., Sabbi E., Stark D. P., Targett T. A., Ellis R. S., 2010, MNRAS, 403, 960

Mehta V., et al., 2017, ApJ, 838, 29

Meurer G. R., Heckman T. M., Calzetti D., 1999, ApJ, 521, 64

Miley G., De Breuck C., 2008, AAPR, 15, 67

Miyazaki S., et al., 2002, PASJ, 54, 833

Møller P., Warren S. J., 1993, A&A, 270, 43

Momose R., et al., 2014, MNRAS, 442, 110

Mortlock A., Conselice C. J., Bluck A. F. L., Bauer A. E., Grützbauch R., Buitrago F., Ownsworth J., 2011, MNRAS, 413, 2845

Mortlock A., et al., 2015, MNRAS, 447, 2

Muldrew S. I., Hatch N. A., Cooke E. A., 2018, MNRAS, 473, 2335

Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, MNRAS, 454, 2691

Murayama T., Taniguchi Y., et al., 2007, ApJs, 172, 523

Muzzin A., et al., 2013, ApJS, 206, 8

Nakajima K., Ouchi M., Shimasaku K., et al., 2012, ApJ, 745, 12

Nakajima K., Ellis R. S., Iwata I., Inoue A. K., Kusakabe H., Ouchi M., Robertson B. E., 2016, ApJL, 831, L9

Nakamura F., Umemura M., 2001, ApJ, 548, 19

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Neufeld D. A., 1990, ApJ, 350, 216

Neufeld D. A., 1991, ApJL, 370, L85

Nilsson K. K., et al., 2007, A&A, 471, 71

Nilsson K. K., Tapken C., Møller P., Freudling W., Fynbo J. P. U., Meisenheimer K., Laursen P., Östlin G., 2009, A&A, 498, 13

Noeske K. G., et al., 2007, ApJL, 660, L43

Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713

Oliver S. J., et al., 2012, MNRAS, 424, 1614

Ono Y., et al., 2010, MNRAS, 402, 1580

Ono Y., et al., 2012, ApJ, 744, 83

Ono Y., et al., 2018, Publications of the ASJ, 70, S10

Östlin G., et al., 2014, ApJ, 797, 11

Ostriker J. P., Gnedin N. Y., 1996, ApJL, 472, L63

Oteo I., Sobral D., Ivison R. J., Smail I., Best P. N., Cepa J., Pérez-García A. M., 2015, MNRAS, 452, 2018

Ouchi M., et al., 2008, ApJs, 176, 301

Ouchi M., et al., 2010, ApJ, 723, 869

Ouchi M., et al., 2013, ApJ, 778, 102

Overzier R. A., 2016, A&AR, 24, 14

Oyarzún G. A., Blanc G. A., González V., Mateo M., Bailey III J. I., 2017, ApJ, 843, 133

Palla F., Salpeter E. E., Stahler S. W., 1983, ApJ, 271, 632

Partridge R. B., Peebles P. J. E., 1967, ApJ, 147, 868

Paulino-Afonso A., Sobral D., Buitrago F., Afonso J., 2017, MNRAS, 465, 2717

Paulino-Afonso A., et al., 2018, MNRAS, 476, 5479

Pentericci L., Grazian A., Fontana A., Salimbeni S., Santini P., de Santis C., Gallozzi S., Giallongo E., 2007, A&A, 471, 433

Pentericci L., et al., 2011, ApJ, 743, 132

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S., et al., 1999, ApJ, 517, 565

Pilbratt G. L., et al., 2010, A&A, 518, L1

Planck Collaboration et al., 2018, arXiv e-prints, p. arXiv:1807.06209

Popping G., Puglisi A., Norman C. A., 2017, MNRAS, 472, 2315

Pozzetti L., et al., 2010, A&A, 523, A13

Pritchet C. J., 1994, PASP, 106, 1052

Pritchet C. J., Hartwick F. D. A., 1987, ApJ, 320, 464

Puccetti S., et al., 2009, ApJS, 185, 586

Raiter A., Schaerer D., Fosbury R. A. E., 2010, A&A, 523, A64

Rauch M., et al., 2008, ApJ, 681, 856

Reddy N. A., Steidel C. C., 2009, ApJ, 692, 778

Reddy N., et al., 2012, ApJ, 744, 154

Reddy N. A., et al., 2018, ApJ, 853, 56

Rhoads J. E., Malhotra S., Dey A., Stern D., Spinrad H., Jannuzi B. T., 2000, ApJL, 545, L85

Rhoads J. E., et al., 2003, AJ, 125, 1006

Ribeiro B., Le Fèvre O., et al., 2016, A&A, 593, A22

Richards G. T., et al., 2006, AJ, 131, 2766

Riess A. G., et al., 1998, AJ, 116, 1009

Rosdahl J., Blaizot J., 2012, MNRAS, 423, 344

Rubin V. C., Ford W. K. J., Thonnard N., 1980, ApJ, 238, 471

Salim S., et al., 2009, ApJ, 700, 161

Salmon B., et al., 2015, ApJ, 799, 183

Salpeter E. E., 1955, ApJ, 121, 161

Sanders D. B., et al., 2007, ApJS, 172, 86

Santini P., et al., 2012, A&A, 538, A33

Santos S., Sobral D., Matthee J., 2016, MNRAS, 463, 1678

Santos S., et al., 2020, MNRAS, p. 92

Sawicki M., Thompson D., 2006, ApJ, 642, 653

Sawicki M., et al., 2008, ApJ, 687, 884

Schaerer D., 2003, A&A, 397, 527

Schaye J., Crain R. A., et al., 2015, MNRAS, 446, 521

Schechter P., 1976, ApJ, 203, 297

Schenker M. A., Ellis R. S., Konidaris N. P., Stark D. P., 2014, ApJ, 795, 20

Schinnerer E., et al., 2004, AJ, 128, 1974

Schinnerer E., et al., 2007, ApJS, 172, 46

Schinnerer E., et al., 2010, ApJS, 188, 384

Schreiber C., et al., 2015, A&A, 575, A74

Scoville N., et al., 2007, ApJS, 172, 1

Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, ApJ, 690, 20

Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65

Shibuya T., Kashikawa N., Ota K., Iye M., Ouchi M., Furusawa H., Shimasaku
    K., Hattori T., 2012, ApJ, 752, 114

Shibuya T., et al., 2014, ApJ, 788, 74

Shibuya T., et al., 2018, Publications of the ASJ, 70, S15

Shibuya T., Ouchi M., Harikane Y., Nakajima K., 2019, ApJ, 871, 164

Shimakawa R., et al., 2017, MNRAS, 468, 1123

Shimasaku K., et al., 2006, Publications of the ASJ, 58, 313

Shimizu I., Yoshida N., Okamoto T., 2011, MNRAS, 418, 2273

Shimizu I., Todoroki K., Yajima H., Nagamine K., 2019, MNRAS, 484, 2632

Shioya Y., Taniguchi Y., et al., 2009, ApJ, 696, 546

Shivaei I., et al., 2018, ApJ, 855, 42

Silk J., 1977, ApJ, 211, 638

Skrutskie M. F., et al., 2006, AJ, 131, 1163

Smit R., Bouwens R. J., Labbé I., Franx M., Wilkins S. M., Oesch P. A., 2016,
    ApJ, 833, 254

Smolčić V., et al., 2017, A&A, 602, A1

Sobral D., Matthee J., 2019, A&A, 623, A157

Sobral D., et al., 2009, MNRAS, 398, L68

Sobral D., Best P. N., Matsuda Y., Smail I., Geach J. E., Cirasuolo M., 2012, MNRAS, 420, 1926

Sobral D., Smail I., Best P. N., Geach J. E., Matsuda Y., Stott J. P., Cirasuolo M., Kurk J., 2013, MNRAS, 428, 1128

Sobral D., Best P. N., Smail I., Mobasher B., Stott J., Nisbet D., 2014, MNRAS, 437, 3516

Sobral D., Matthee J., Darvish B., Schaerer D., Mobasher B., Röttgering H. J. A., Santos S., Hemmati S., 2015, ApJ, 808, 139

Sobral D., et al., 2017, MNRAS, 466, 1242

Sobral D., Santos S., Matthee J., Paulino-Afonso A., Ribeiro B., Calhau J., Khostovan A. A., 2018a, MNRAS, 476, 4725

Sobral D., et al., 2018b, MNRAS, 477, 2817

Sobral D., et al., 2019, MNRAS, 482, 2422

Song M., et al., 2014, ApJ, 791, 3

Song M., et al., 2016, ApJ, 825, 5

Spergel D. N., et al., 2003, ApJS, 148, 175

Stanway E. R., Bunker A. J., McMahon R. G., 2003, MNRAS, 342, 439

Stanway E. R., et al., 2004, ApJL, 604, L13

Stanway E. R., McMahon R. G., Bunker A. J., 2005, MNRAS, 359, 1184

Stark D. P., Ellis R. S., Richard J., Kneib J.-P., Smith G. P., Santos M. R., 2007, ApJj, 663, 10

Stark D. P., Ellis R. S., Chiu K., Ouchi M., Bunker A., 2010, MNRAS, 408, 1628

Stark D. P., et al., 2017, MNRAS, 464, 469

Steidel C. C., Hamilton D., 1993, AJ, 105, 2017

Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996a, ApJL, 462, L17

Steidel C. C., Giavalisco M., Pettini M., Dickinson M., Adelberger K. L., 1996b, ApJL, 462, L17

Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1

Steinhardt C. L., et al., 2014, ApJL, 791, L25

Stroe A., Sobral D., Matthee J., Calhau J., Oteo I., 2017a, MNRAS, 471, 2558

Stroe A., Sobral D., Matthee J., Calhau J., Oteo I., 2017b, MNRAS, 471, 2575

Sutherland W., et al., 2015, A&A, 575, A25

Taniguchi Y., et al., 2005, PASJ, 57, 165

Taniguchi Y., et al., 2007, ApJS, 172, 9

Taniguchi Y., et al., 2015a, Publications of the ASJ, 67, 104

Taniguchi Y., et al., 2015b, ApJL, 809, L7

Tasca L. A. M., et al., 2015, A&A, 581, A54

Taylor A. J., Barger A. J., Cowie L. L., Hu E. M., Songaila A., 2020, arXiv e-prints, p. arXiv:2004.09510

Thomas D., et al., 2013, MNRAS, 431, 1383

Thompson D., Djorgovski S., Trauger J., 1995, AJ, 110, 963

Tomczak A. R., et al., 2014, ApJ, 783, 85

Tormen G., Bouchet F. R., White S. D. M., 1997, MNRAS, 286, 865

Trainor R. F., Steidel C. C., Strom A. L., Rudie G. C., 2015, ApJ, 809, 89

Vanzella E., De Barros S., et al., 2016, ApJL, 821, L27

Venemans B. P., Röttgering H. J. A., Miley G. K., et al., 2007, A&A, 461, 823

Verhamme A., Schaerer D., Maselli A., 2006, A&A, 460, 397

Verhamme A., Schaerer D., Atek H., Tapken C., 2008, A&A, 491, 89

Verhamme A., Orlitová I., Schaerer D., Hayes M., 2015, A&A, 578, A7

Verhamme A., Orlitová I., Schaerer D., Izotov Y., Worseck G., Thuan T. X., Guseva N., 2017, A&A, 597, A13

Vilella-Rojo G., et al., 2015, A&A, 580, A47

Wardlow J. L., et al., 2014, ApJ, 787, 9

Westra E., et al., 2006, AAP, 455, 61

Whitaker K. E., et al., 2014, ApJ, 795, 104

Whitaker K. E., Pope A., Cybulski R., Casey C. M., Popping G., Yun M. S., 2017, ApJ, 850, 208

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Wisotzki L., Bacon R., et al., 2016, A&A, 587, A98

Wisotzki L., et al., 2018, Nature, 562, 229

Wold I. G. B., Barger A. J., Cowie L. L., 2014, ApJ, 783, 119

Wold I. G. B., Finkelstein S. L., Barger A. J., Cowie L. L., Rosenwasser B., 2017, ApJ, 848, 108

Yajima H., Li Y., Zhu Q., Abel T., Gronwall C., Ciardullo R., 2012, ApJ, 754, 118

Yamada T., Nakamura Y., Matsuda Y., Hayashino T., Yamauchi R., Morimoto N., Kousai K., Umemura M., 2012, AJ, 143, 79

Yang H., et al., 2017a, ApJ, 844, 171

Yang H., Malhotra S., Rhoads J. E., Wang J., 2017b, ApJ, 847, 38

Zheng Z.-Y., Wang J.-X., Malhotra S., Rhoads J. E., Finkelstein S. L., Finkelstein K., 2014, MNRAS, 439, 1101

Zheng Z.-Y., et al., 2017, ApJL, 842, L22

Zitrin A., et al., 2015, ApJL, 810, L12

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595

da Cunha E., Charlot S., Dunne L., Smith D., Rowlands K., 2012, in Tuffs R. J., Popescu C. C., eds, IAU Symposium Vol. 284, The Spectral Energy Distribution of Galaxies - SED 2011. pp 292–296 (arXiv:1111.3961), doi:10.1017/S1743921312009283

da Cunha E., et al., 2015, ApJ, 806, 110

van Breukelen C., Jarvis M. J., Venemans B. P., 2005, MNRAS, 359, 895

van der Wel A., Franx M., et al., 2014, ApJ, 788, 28